

LETTER • OPEN ACCESS

Analysis of the trade-off between high crop yield and low yield instability at the global scale

To cite this article: Tamara Ben-Ari and David Makowski 2016 *Environ. Res. Lett.* **11** 104005

View the [article online](#) for updates and enhancements.

You may also like

- [Disentangling the separate and confounding effects of temperature and precipitation on global maize yield using machine learning, statistical and process crop models](#)
Xiaomeng Yin, Guoyong Leng and Linfei Yu
- [Seasonal tropospheric cooling in Northeast China associated with cropland expansion](#)
Yaqian He, Eungul Lee and Justin S Mankin
- [Decomposing global crop yield variability](#)
Tamara Ben-Ari and David Makowski

Environmental Research Letters



LETTER

Analysis of the trade-off between high crop yield and low yield instability at the global scale

OPEN ACCESS

RECEIVED

25 February 2016

REVISED

17 June 2016

ACCEPTED FOR PUBLICATION

4 July 2016

PUBLISHED

3 October 2016

Original content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Tamara Ben-Ari¹ and David Makowski

UMR 211 Agronomie INRA, Agroparistech, Université Paris-Saclay, 78850 Thiverval-Grignon, France

¹ Author to whom any correspondence should be addressed.E-mail: tamara.ben-ari@grignon.inra.fr**Keywords:** land use, yield interannual variability, quadratic optimizationSupplementary material for this article is available [online](#)

Abstract

Yield dynamics of major crops species vary remarkably among continents. Worldwide distribution of cropland influences both the expected levels and the interannual variability of global yields. An expansion of cultivated land in the most productive areas could theoretically increase global production, but also increase global yield instability if the most productive regions are characterized by high interannual yield variability. In this letter, we use portfolio analysis to quantify the tradeoff between the expected values and the interannual variance of global yield. We compute optimal frontiers for four crop species i.e., maize, rice, soybean and wheat and show how the distribution of cropland among large world regions can be optimized to either increase expected global crop production or decrease its interannual variability. We also show that a preferential allocation of cropland in the most productive regions can increase global expected yield at the expense of yield stability. Theoretically, optimizing the distribution of a small fraction of total cultivated areas can help find a good compromise between low instability and high crop yields at the global scale.

Introduction

The last fifty years have seen dramatic increase in crop yields for several species including maize, rice, soybean, and wheat in large producing regions [1]. Yield increases (in quantity per unit of land) are tied to several factors including sharp increases in the use of mineral fertilizer, pesticides, irrigation, and investments in crop breeding. These practices and technologies have enabled a reduction of estimated yield gaps [2] in high income countries of Western Europe, the Americas, and Southern and Eastern Asia for example [3, 4] whereas in other parts of the world estimated yield gaps remain large (e.g., Africa [2] or in Eastern Europe [3, 5]). Prospects of further yield increases in the near future seem to rely on the evermore providing of resources (e.g., water [6], nitrogen [7] or phosphorus [8]) and technological improvements. Hence, the adoption of more intensive practices may increase crop yields in several regions of the world [7], but also generate more adverse environmental effects. Several studies cast doubt on crop production growth rates'

ability to keep pace with past accelerations [9, 10]. Concerns for yield stagnation in key producing countries are indeed rising in the food security debate [11–13] particularly considering how climate change [14, 15] or resource access and limitations [16] may limit further yield improvements or affect vulnerability to extremes in various key producing regions [17]. Besides intensifying production on existing land, increasing future crop production relies on an expansion onto remaining cultivable lands. Yet, because of urbanization [22], land degradation [23] or biomass for energy production [24], arable land is undoubtedly becoming a scarce resource. Substitution of natural ecosystems (such as forests, permanent grasslands or wetlands) for agricultural activities is notwithstanding associated with livelihood [26] and environmental deteriorations [25] and acute biodiversity loss [26]. The effectiveness of an increase of regional yields and/or of global cultivated areas for increasing global crop production thus remains uncertain.

Theoretically, it is possible to increase global crop production by allocating larger proportions of

cropland to the most productive regions, without the need of yield gap reduction at the regional scale and without increasing global cropland areas. However, this may also have negative consequences if the most productive regions are characterized by high inter-annual yield variability or if their yields are positively correlated. Increasing the proportions of total cropland in such regions can increase global yield variance and, consequently, increase global production instability [18]. A strong focus on the maximization of expected yield may therefore lead to an increase of the instability of global crop production. Cropland allocation strategies designed for increasing global crop production thus need to be evaluated by considering both the expected levels and interannual variance of crop yields. In fact, understanding the duality between these two dimensions is essential to address the first and fourth pillars of global food security (i.e., availability and stability) [19] altogether.

Our analysis helps quantify the tradeoff between these two dimensions for the four most cropped species (maize, rice, soybean, and wheat). We present a theoretical framing of the relationship between expected global crop yield and its interannual variance. Quadratic optimization helps finding compromises in which both global production levels and stability are improved. We identify theoretical contrasted cropland distributions leading to higher expected production and lower instability levels compared to the current situation.

The main questions addressed in this study are: (i) what is the trade-off between high crop yield—and low yield instability at the global scale for maize, rice, soybean, and wheat? (ii) Is current cropland distribution, for each of the four crop species considered here, optimal *vis a vis* global expected yield level or inter-annual variance? (iii) How much could expected production theoretically increase without increasing yield variance, or how much could yield variance decrease without decreasing expected production (i.e., stability improved for a fixed production level)? Based on annual yield statistics collected for the 1961–2013 period, our results show that, it is theoretically possible to either increase expected crop production or to decrease its instability, and to find compromises between these two objectives. The strength of this compromise depends on the crop species studied.

Material and methods

Data

Our dataset includes yield and harvested area time series extracted from the UN Food and Agriculture Organisation FAOSTAT database [1] for four crops species (maize, rice, soybean, and wheat). The time series extend from 1961 to 2013, except in Central Asia (1992–2012). We select FAO world regions totalizing at least 90% of global production over the studied time

period (i.e., a total of 12 regions are considered in the study for the four crop species). Three regions, all in Asia (i.e., Eastern, Southern and Southeastern Asia), amounted to about 91% of the global rice production. Since 1961, 93% of soybean has also been produced in three regions (i.e., Eastern Asia, Northern and South America). On the other hand, nine regions produced at least 92% of global wheat production (Northern America, Central Asia, Eastern Asia, Southern Asia, Western Asia, Eastern Europe, Northern Europe, Southern Europe and Western Europe) and 90.3% of maize global production (Central America, Northern America, South America, Eastern Asia, South-Eastern Asia, Southern Asia, Eastern Europe, Southern Europe and Western Europe for maize).

Throughout the letter, we refer to total yearly harvested areas in each region as cropland.

Global expected yield

For each crop species we calculate a yield residual $r_{i,t}$ at year t in region i as:

$$r_{i,t} = Y_{i,t} - \mu_{i,t} \quad (1)$$

where $Y_{i,t}$ is the yield and $\mu_{i,t}$ is the expected yield value at year t for a given crop in the i th region. Values of $\mu_{i,t}$ are estimated using linear, quadratic or cubic regression models [20, 21]. Polynomial regressions are based on the following equation

$$Y_{i,t} = a_i t^3 + b_i t^2 + c_i t + d_i + \varepsilon_{it} \quad (2)$$

where a_i , b_i , c_i , and d_i are the regression coefficients in the i th region for each crop species. Three variants are defined from equation (2), linear regression ($a_i = b_i = 0$), quadratic regression ($a_i = 0$), and cubic regression (all coefficients are non-zero). These models are fitted by ordinary least squares, using the R function `lm`. The best model (i.e., according to the Akaike Information Criteria (AIC)) is selected to detrend yield time series. We check for absence of autocorrelation in the residuals.

Let S_t be the total cultivated area in all regions at time t for a given crop: $S_t = \sum_{i=1}^N S_{i,t}$ where $S_{i,t}$ is the area harvested in the i th region, $i = 1, \dots, N$ with N the total number of regions considered for each crop. Let $\omega_{i,t}$ be the proportion of the total cultivated area allocated to the i th region at time t :

$$\omega_{i,t} = S_{i,t} / S_t \quad (3)$$

Global yield can be expressed as the weighted average of regional yields; $\bar{Y}_t = \sum_{i=1}^N \omega_{i,t} Y_{i,t}$ and similarly, expected global yield is:

$$\bar{\mu}_t = \sum_{i=1}^N \omega_{i,t} \mu_{i,t} \quad (4)$$

Standard deviation, variance and covariance of regional yields

Let C be the yield residual variance covariance matrix:

$$C = \begin{pmatrix} c_{1,1} & c_{1,2} & \dots & c_{1,N} \\ c_{2,1} & c_{2,2} & \dots & c_{2,N} \\ c_{i,1} & \dots & c_{i,i} & c_{i,N} \\ c_{N,1} & \dots & c_{N,N-1} & c_{N,N} \end{pmatrix}$$

with $c_{i,i}$ is the variance in region i and $c_{i,j}$ is the covariance between yield residuals in regions i and j calculated as:

$$c_{i,j} = 1/T \sum_{t=1}^T r_{i,t} r_{j,t} \quad (5)$$

Global yield variance is expressed as:

$$V(\bar{Y}_t) = \omega'_t C \omega_t \quad (6)$$

with ω'_t the transposed vector of proportions $\omega_{i,t}$, $i = 1, \dots, N$ at year t . Note that the variance-covariance matrix C is assumed constant across time.

Quadratic programming for optimizing cropland distribution

For each crop species independently, we calculate optimal proportions of cropland $\omega_{i,t}$, $i = 1, \dots, N$, at year t and for a given set of N regions by minimizing the following expression

$$\omega'_t C \omega_t - \lambda \sum_{i=1}^N \omega_{i,t} \mu_{i,t} = V(\bar{Y}_t) - \lambda \bar{\mu}_t \quad (7)$$

under the constraints $\sum_{i=1}^N \omega_{i,t} = 1$ (the sum of cropland proportions is equal to 1 every year) and $\omega_{i,t} \geq 0 \forall i$ (proportions are positive, for all crop species and all years). λ is an optimization parameter. Equation (7) can be minimized by quadratic programming for any λ . Low values of λ give more weight to minimizing global yield variance and high values to maximizing expected global yield. The parameter λ thus corresponds to a tolerance to yield instability. When $\lambda = 0$, the optimal solution found by quadratic programming gives the lowest global yield variance, for a given set of N regions. When $\lambda \rightarrow +\infty$, the optimal solution gives the highest possible expected global yield, for a given set of N regions. Intermediate values of λ correspond to a continuum of compromises between expected global yield and yield variance. Here, optimal solutions are generated for the full range of values of λ using the function `solve.QP` of the R package `quadprog` [22]. An optimal frontier is drawn for each crop species. Any point on this frontier gives, for each year, the lowest achievable global variance for a given expected global yield value. In essence, this approach is similar to modern portfolio theory that attempts to maximize portfolio expected return for a given amount of portfolio risk and a given set of assets [23].

For each crop species, two specific optimal solutions are analyzed in detail for 2013 (the most recent year of our time series); the production optimum defined as the solution maximizing global expected yield without increasing global yield variance

compared to 2013, and the variance optimum defined as the solution minimizing global yield variance without decreasing global expected yield compared to 2013.

Consequences on production and variance of optimizing a fraction of croplands

We calculate crop production resulting from an optimization of a fraction (noted α , $\alpha = 0, \dots, 0.4$) of cropland to maximize expected yields without increasing global yield variance (production optimum). In this scenario, optimized proportions $\omega_{i,t}^{\text{opt}}$, $i = 1, \dots, N$, are applied on a fraction α of the total cropland area, whereas observed cropland proportions $\omega_{i,t}$, $i = 1, \dots, N$, are applied in the remaining fraction $1 - \alpha$. The resulting crop production, noted $P_{\text{opt}}^{\text{exp}}$, is expressed as:

$$P_{\text{opt}}^{\text{exp}} = (1 - \alpha) \left(S_t \sum_{i=1}^N \omega_{i,t} \mu_{i,t} \right) + \alpha \left(S_t \sum_{i=1}^N \omega_{i,t}^{\text{opt}} \mu_{i,t} \right) \quad (8)$$

We then calculate the relative difference ΔP between the theoretical production $P_{\text{opt}}^{\text{exp}}$ and the expected global production P^{exp} obtained by applying observed cropland proportions on the whole cropland area: $\Delta P = (P_{\text{opt}}^{\text{exp}} - P^{\text{exp}})/P^{\text{exp}}$, where the expected global production P^{exp} is defined as $P^{\text{exp}} = S_t \sum_{i=1}^N \omega_{i,t} \mu_{i,t}$.

Following a similar reasoning, we calculate the reduction in global production variance resulting from an optimization of a fraction α of cropland based on the variance optimum. If Q is the vector including the quantity $(1 - \alpha)(S_t \omega_{i,t}) + \alpha(S_t \omega_{i,t}^{\text{opt}})$, $i = 1, \dots, N$, the production variance obtained when a fraction α of the cropland is optimized is defined by

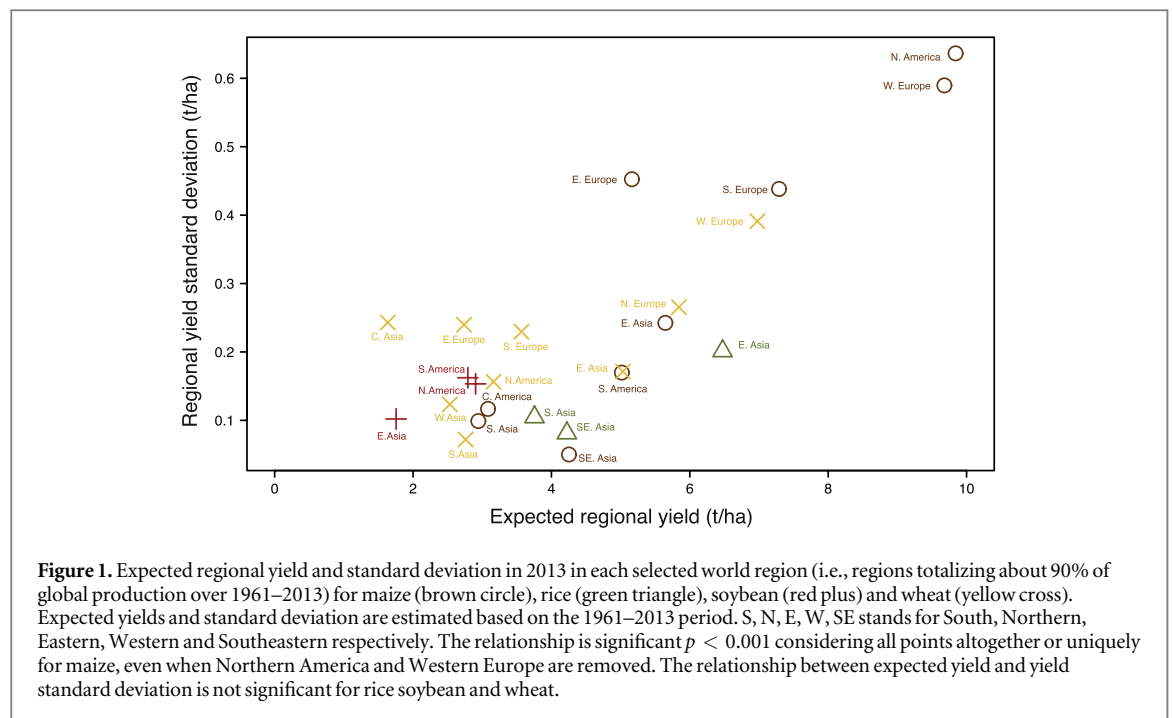
$$V_{\text{opt}}^{\text{exp}} = Q' C Q. \quad (9)$$

The resulting relative global variance decrease is expressed as: $\Delta V = (V^{\text{exp}} - V_{\text{opt}}^{\text{exp}})/V^{\text{exp}}$ with $V^{\text{exp}} = Q_o' C Q_o$ and Q_o is equal to Q when $\alpha = 0$.

In this approach, 10% of global cropland is optimized when alpha is equal to 0.1. In this case, results of the optimization concern 10% of global cultivated area. For illustration, relative production increase ΔP and variance decrease ΔV are computed for values of α ranging from zero to 0.4. To assess the impact of optimizing a fraction of global cropland, we also present the results obtained for a specific value of α equal to the observed relative increase of cultivated area in 2013 compared to 2000 i.e., $\alpha = (S_{2013} - S_{2000})/S_{2013}$. This fraction ranges from 0.01 to 0.3, depending on the crop species.

Results

We base our analysis on yields and harvested areas time series at the scale of UN large world regions for the 1961–2013 period [1]. Figure 1 shows expected yields for year 2013 and yield standard deviations



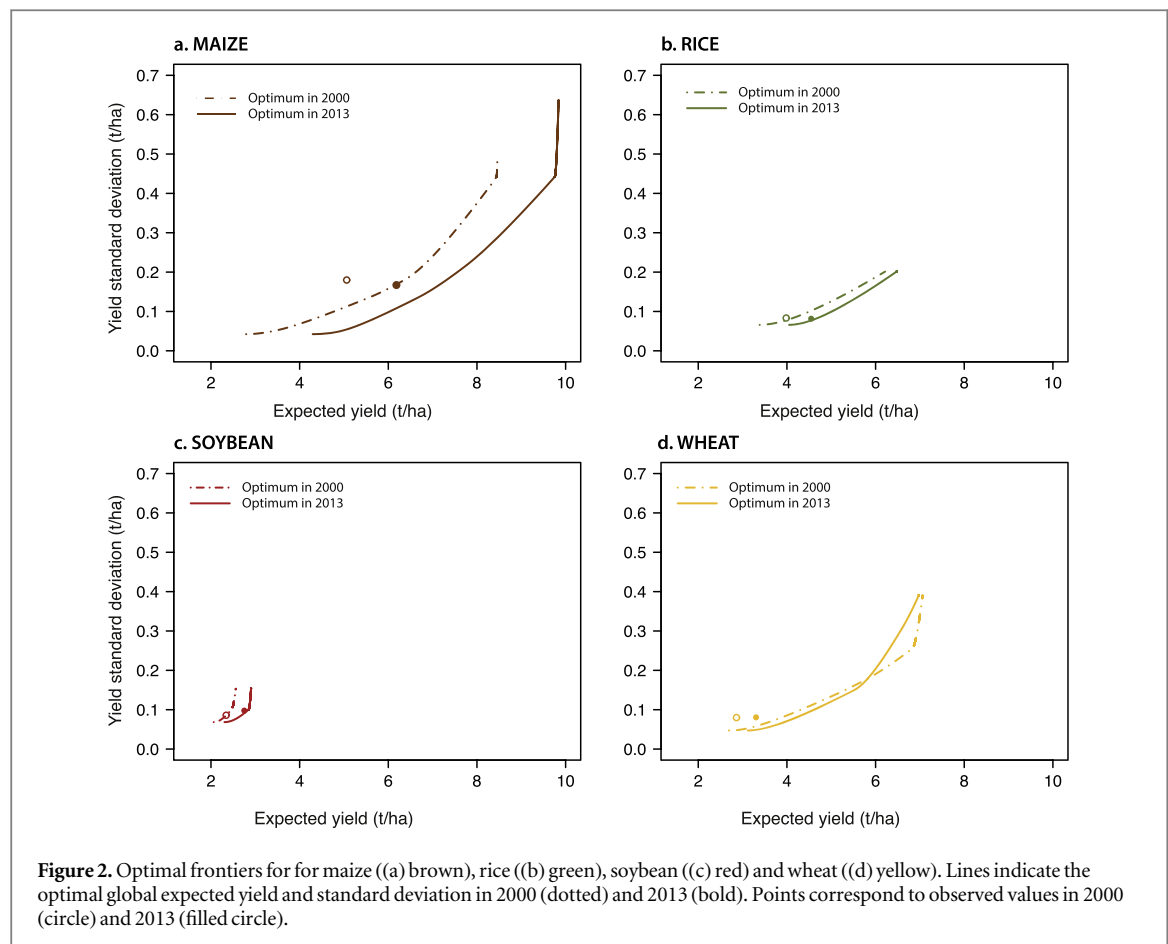
(assumed constant over time) for each crop species in each selected region. World regions characterized by the highest expected yields never coincide with the regions showing the lowest yield standard deviation. For three crop species (i.e., maize, wheat and rice), the regions with the highest expected yields are those showing the highest yield standard deviation. We find a significant linear relationship between expected yield and yield standard deviation ($p < 0.001$) for all four crops considered together and for maize alone; but this relation is not significant for rice, soybean and wheat. For all species though, comparable expected yield levels can be associated with contrasted standard deviations (e.g., South American and Eastern Europe expected maize yield are 5.01 and 5.16 t ha⁻¹ but correspond to standard deviations of 0.17 and 0.45 t ha⁻¹, respectively). Similarly, comparable instability levels can be associated with very different expected yields (e.g., wheat yield standard deviations of 0.26 and 0.24 t ha⁻¹ for 5.84 and 1.63 t ha⁻¹ expected yields in Northern Europe and Central Asia respectively). Both the expected and standard deviations of maize yields and—to a smaller extent—wheat yields, show much wider amplitudes than those observed for rice and soybean for the considered regions; i.e., there are larger inter-regional differences for maize and wheat than for rice and soybean in the regions considered here.

We calculate cropland proportions ($\omega_{i,t}$) minimizing a linear combination of global yield variance and expected yield values by quadratic optimization (equation (7)). This enables us to draw an optimal frontier for each crop species (figures 2, 3 and S3). Any point on this frontier gives, for each year, the lowest achievable global variance for a given expected global yield value (figure S1). Each solution corresponds to a

compromise between the production and stability criteria. Solutions located on the top-right corner of the frontier give more weight to increasing global yield and those located on the bottom-left give more weight to decreasing yield variance. Each value of the tolerance to instability parameter (λ) is associated with one global cropland distribution. Scanning all values for this parameter thus shows all theoretical cropland proportion distributions between the lowest possible global yield instability and the highest possible expected yield (figure 2). Obviously, exclusively maximizing production is equivalent to allocating any additional hectare to the region of the world with the highest yield (i.e., North America for maize and soybean, Eastern Asia for rice and Western Europe for wheat). An application of this solution to the whole cropping area would lead to an increase of yield variance by a factor of the order of 2 to 4 according to the crop species (figure 3).

Observed cropland distributions for the years 2000 and 2013 result in suboptimal situations for maize and wheat, but are close to the optimal frontier for soybean and rice (figure 2). Note that the number of regions considered is lower in the latter. The distance between the actual yield-variance tradeoff and the optimal frontier remains stable during the 2000–2013 (figure 2). This distance is in fact fairly constant since 1961 (figure S1).

For each crop, we focus on two optimal solutions, presented for the last year of our time series (i.e., 2013). These two optima are particular solutions on the continuum represented by the optimal frontier (figure 3), and correspond to different cropland proportions. The first solution corresponds to an optimal distribution of global cropland decreasing global yield



interannual variance without decreasing global expected yield compared to 2013 levels. The second solution allows increasing global expected yield without increasing yield variance. These two optima are obtained by drawing a vertical (alt. horizontal) line between the sub-optimal observed point and the frontier (figure 3). The first solution can be considered as optimal in stability and the second as optimal in expected yield (henceforth referred to as the stability and production optima). All solutions located between these two optima improve both yield levels and stability compared to the situation observed in 2013.

For maize, the two optima considered here are associated with a preferential distribution of cropland fraction to Eastern Asia (in particular to increase average production, i.e., from about 24% of considered cropland in 2013 to about 40%) and a decrease of Northern America cropland proportions (from about 25% in 2013 to 17 and 10% for the production and stability optima respectively) (figure 4). The 2013 proportion of maize cropland in South America (16%) is nearly optimal in terms of expected production levels but should be increased to 24% to improve its stability. To improve maize global expected yield without increasing its instability, about 16% of considered maize cropland should also be distributed in Western Europe (figure 4). Similarly, global maize yield stability

would benefit from an increase of area proportions in Southeastern Asia (+12%).

The rice cropland distribution in 2013 is close to the optimal frontier (figures 2 and 3). Note though that the optimal frontier also depends on the set of selected regions, a different result can be expected from a larger set of regions. Two of the most important rice producing regions are characterized by very similar expected yield and variance (i.e., South and Southeastern Asia, but with the latter characterized by a slightly higher expected yield and lower yield standard deviation, figure 1). Eastern Asia on the other hand takes higher values for both criteria. Both the production and stability optima are associated with an increased proportion of cropland allocated to Southeastern Asia compared to 2013 (i.e., from 35% to 56% or 58% for the production and stability optima respectively). These gains are compensated by a halving of cropland fraction in Southern Asia and a slight decrease in Eastern Asia for the production optimum only.

In 2013, the bulk of soybean production is located in two regions with comparable regional yield average and standard deviation levels: North and South America. The two optima obtained for soybean are both associated with increased proportions in Northern America (up to about 53% of global areas) and a decrease in South America (from 57 to above 40%,

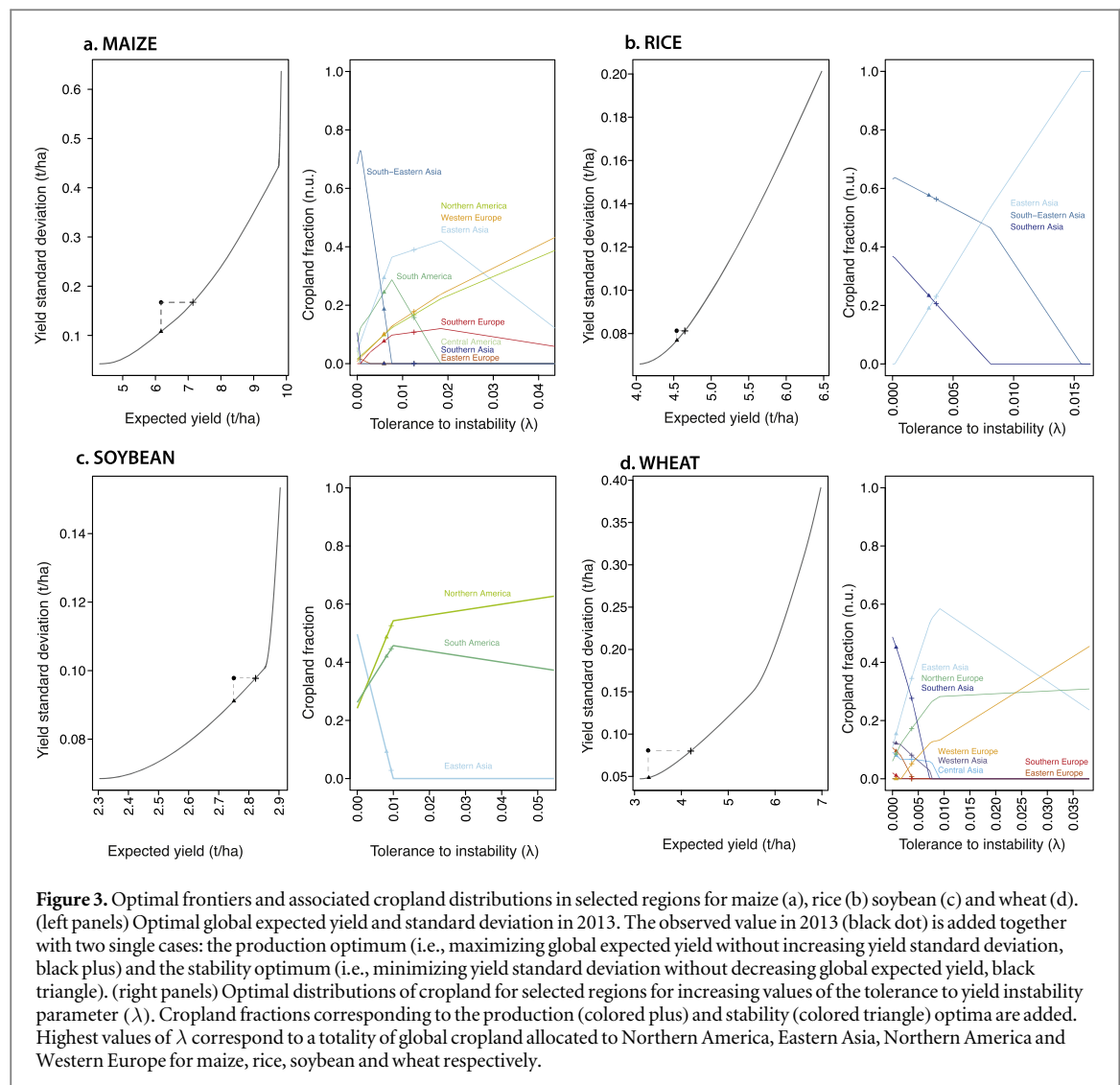


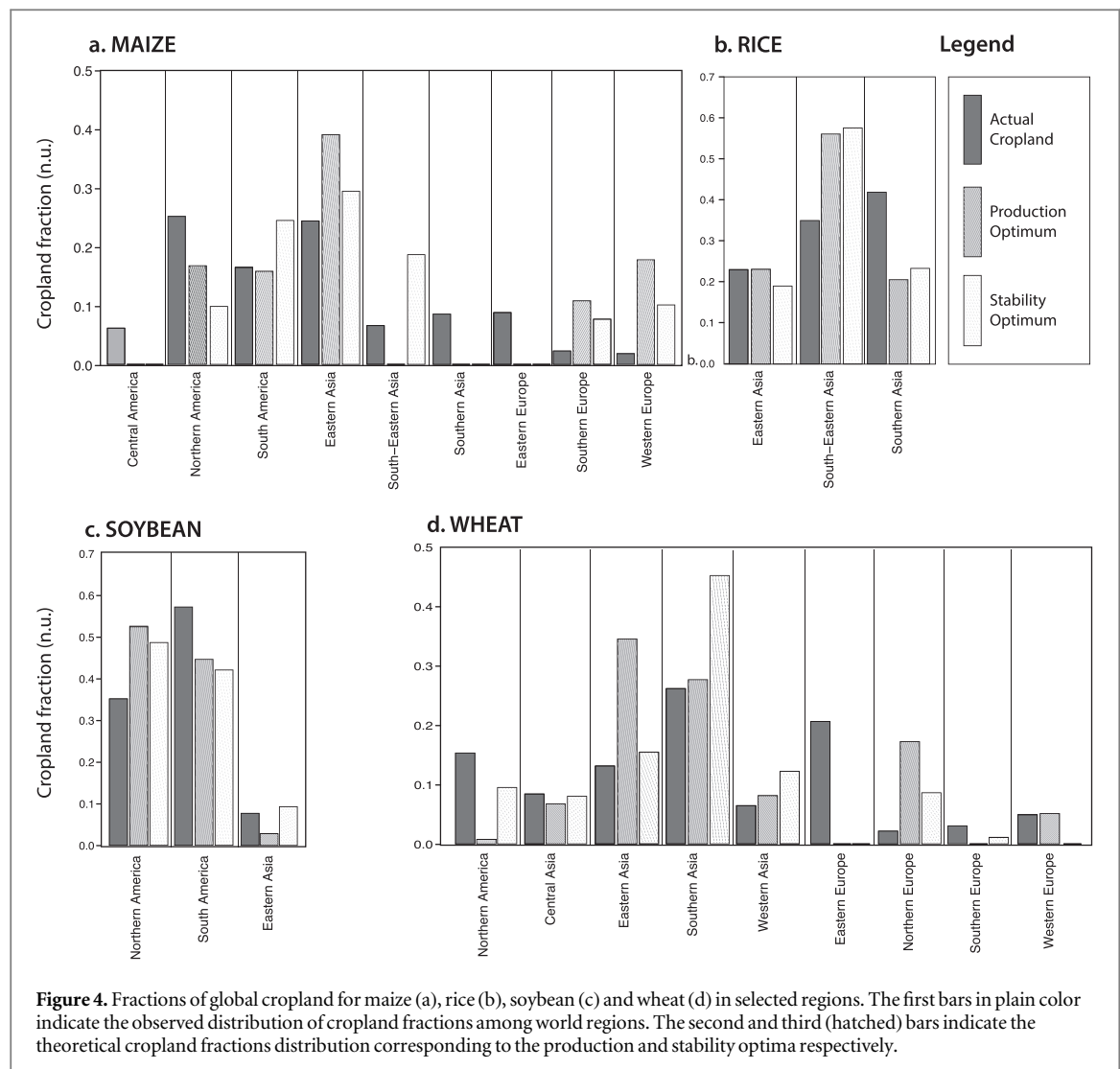
figure 4). Note that yield variance in South America is slightly higher when estimated on a more recent time period (although with more uncertainty due to a lower number of data).

Finally, an increase of wheat cropland proportions in Eastern Asia and Northern Europe together with a large decrease in North American would contribute to increasing global expected yield. A large decrease of the proportion of global cropland in Eastern Europe (from above 20% to none) associated with a large increase in Southern Asia would improve wheat production stability. For both optima, cropland distributions are less dispersed than observed in 2013 (i.e., with one or two dominant world regions in both optima and a strong decrease elsewhere).

To grasp the magnitude of changes induced by applying optimal distributions we estimate (i) the gain in production obtained from applying the production optimum on an increasing fraction of global cropland (equation (8)) and (ii) the decrease in production variance obtained from applying the stability optimum on a fraction of global cropland (equation (9)). The impact of land use optimization on global expected

production is proportional to the fraction of area optimized. Figure 5 shows the levels of average production gain and the levels of variance reduction that would result from an optimal allocation of a fraction (from 0 to 0.4) of global areas. During the 2000–2013 period, global cropland of maize, rice, soybean and wheat net increase amounts to about 36, 8, 28 and 2 million hectares (as measured by $\Delta S = S_{2013} - S_{2000}$) respectively. In percentage of 2013 expected production, an additional 3.8%, 0.12%, 0.78% and 0.3% equivalent respectively to 35 million tons for maize, 0.8 million tons for rice and about 2 million tons for both soybeans and wheat would have been added by an optimal distribution of cropland proportions on ΔS . During the same time period, should the totality of ΔS be optimally distributed, global production variance decrease by 23.3, 1.1, 6.7 and 1.3% for maize, rice, soybean, and wheat, respectively.

In our analyses, we measure yield interannual variability from estimated yield variance (or standard deviation in tons per hectare) calculated as the average of squared yield residuals (i.e., the distance to fitted expected yield, equation (1)). Note that one single



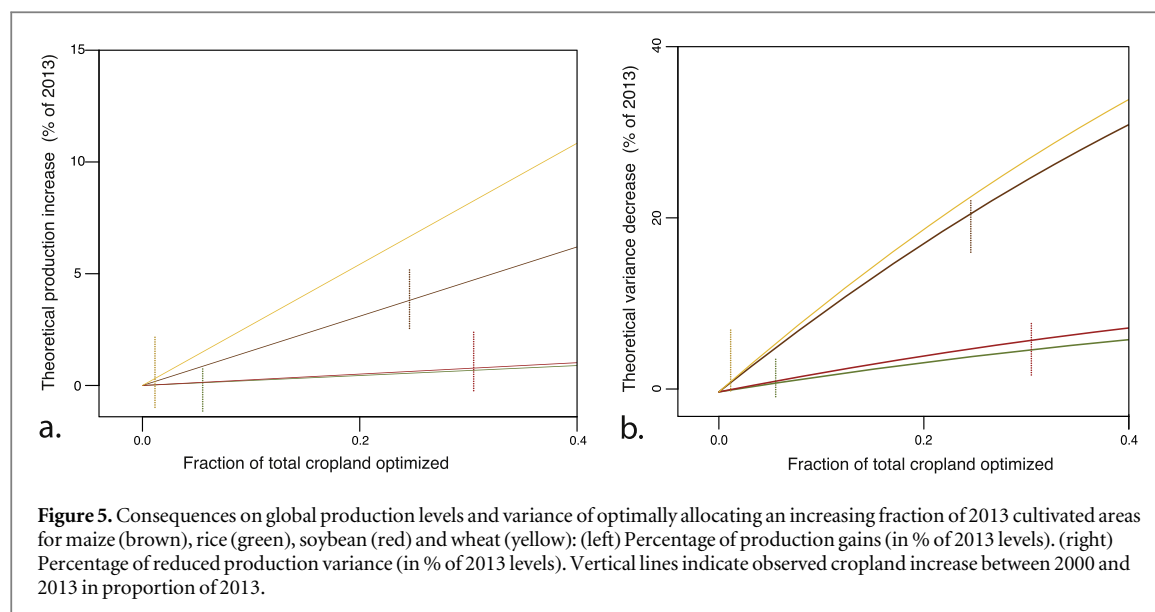
variance-covariance matrix C is computed for the whole time period for each crop*region combination. However, we found that the results were similar when the matrix C was estimated using a shorter time period corresponding to the earliest part of the time series (1990–2013) (figures S3 and S4).

Discussion

Both the average and interannual variance of global production is influenced by cropland distribution among the selected world regions. Largest yields are found in rain-fed wheat systems of Northwestern Europe and Eastern Asia and in rain-fed maize systems in North America with favorable soil quality and climate [24]. Irrigated annual multiple-cropping system in Eastern Asian lowlands are characterized by the highest rice yields [24]. Southeastern Asia is characterized by relatively low rice yields compared to its estimated potential and so does for example, rainfed wheat in Eastern Europe [24]. Obviously, expected yield differences can be explained by contrasts in production systems and intensification levels:

agronomic practices, access to irrigation, inputs, machinery and agricultural labor [25] but also landscape heterogeneity such as topography [26] or soil erosion [27, 28], although it is argued that these differences may wane in the future as yield asymptotically approach their ceilings [10, 13]. Climate variability (in particular temperatures, precipitation or soil moisture) also explains both yield trends and inter-annual variability [15, 29–31]. Note that, the regions of origin for each crop species appear to be among the least variable regions considered here (i.e., Central America for maize, Southern Asia for rice, Eastern Asia for soybean and Central and Southern Asia for wheat [32]).

It is often speculated that yield instability, at least measured in absolute terms, should increase as a result of average yield increase. Such an increase has for example been shown for wheat in many countries spanning five continents [13] or earlier in South Asia and Northern America [33]. The reasons for this relationship are multiple and probably intertwined. Proposed explanations mention the adoption of cultivars more responsive to environmental variation [13]; the



vulnerability of intensive agronomic technologies and practices or to fluctuating prices [33] could also explain an increase of yield variance. Yield interannual variability could also be higher in small than large regions simply due to a lack of risk-pooling effects [34]. A portion of the hypothetical increases of large-scale interannual yield variability could also be due to concomitancy between recent yield increases and climate-change (e.g., increased frequency or intensity of yield impacting out-of-normal events [35]).

Estimations of the time-dynamics of yield variance may convey important uncertainties as they usually rely on the comparison of relatively small data samples [13, 36]. Results can thus be sensitive to the occurrence of a few out-of-normal events and reveal transient variance increase rather than structural changes [34]. As a first step, we hypothesize that the variance of yields, at the scale of large world regions, is constant during the studied period and thus does not contribute to the temporal dynamics in global yield variance (equation (6)). We tested the sensitivity of our results to a change in regional yield variance by computing the variance-covariance matrix over a more recent period (1990–2013). Only the frontiers obtained for rice are affected by the use of a shorter time period; the tradeoff between mean yield and variance is sharper when derived from the shorter time series (figure S4). The empirical variance of soybean is also slightly higher when estimated over 1990–2013, but this has no effect on the optimal frontier (figure S4). No noticeable difference was found for wheat and maize (figure S4).

Global yield variance increases if croplands are located in world regions with the highest levels of yield interannual variability (e.g., maize in Northern America; rice in Eastern Asia, soybean in South America and wheat in Western Europe, figure 1). The local slope of the frontier curve gives the rate at which global yield

variance increases for any gain in expected yield (i.e., corresponding to a preferential allocation of additional acreage in high yielding regions, figures 2 and 3.). The slopes are not monotonous: for soybean, maize and—to a lesser extent—wheat, there is an expected yield threshold above which global yield instability increases sharply.

Main rice producing regions are characterized by below average variance and small yield covariation [18]. Since 1961, the largest estimated yield gains obtained from optimally rearranging rice spatial distribution is of about 0.15 t.ha^{-1} , i.e., the distribution is nearly optimal for the regions considered (figure S2). Both the rice yield variance and expected levels distance to optima show a slight increase in the decade leading to 1990 and a decrease after 2000 but with an overall negligible amplitude (distance to optima are smaller than 0.15 t.ha^{-1} and 0.004 t.ha^{-1} for expected yield and standard deviation respectively, figure S2). Computed optima suggest that both the global rice expected yield and its interannual variability would have been only slightly improved from additional acreage in Eastern (China, Japan, Korea) and Southern Asia (India, Bangladesh, etc).

Soybean production is also very close to the optimal frontier. Aggregated expected yield and yield variance considering the largest three regions were the closest to optimum during 1980–2000 (figure S2). Expected yields and standard deviation are in fact similar in Northern and South America, but with slightly lower average yields and higher instability in the latter (i.e., expected yield is about 2.90 versus 2.79 t ha^{-1} and standard deviation 0.15 versus 0.16 t ha^{-1} in Northern and South America respectively, figure 1). Both the production and variance optima are associated with larger acreage fractions in Northern America. The shape of the optimal frontier indicates that

further gains in expected yield would strongly increase soybean instability.

Above 90% of Maize production is distributed among nine main regions totalizing an expected global yield of about $4 \text{ t} \cdot \text{ha}^{-1}$ over the studied period. It is the crop species with the largest inter-regional differences. Because maize cropland areas are located in regions with high levels of interannual variability (i.e., in particular Northern America and Western Europe), the current cropland allocation is somewhat far from the optimal frontier. Based on this result, maize appears to be the crop species that would have most benefited from an optimization of a small fraction of global acreage (equations (8)–(9); figure 5). Soybean and maize tend to be cropped in similar bioclimatic zones, a finer scale analysis would be useful to distribute acreage between these two crops within the main producing regions (e.g., in the Corn belt using dataset from [17]).

Wheat production is fairly spread out, although less than maize, when global cropland distribution is analyzed using an heterogeneity index [18]. Contrarily to maize, the region with the highest yield and instability levels (i.e., Western Europe) is not dominant in terms of acreage. But, wheat production could also benefit from optimizing a fraction of existing or additional global cropland (figures 3 and S5). The closest optimal solutions to the ones observed include increase acreage proportions in Eastern Asia to improve average yields and in Southern Asia to improve stability.

Overall, our results show that a preferential allocation of croplands in the most productive regions can increase global expected yield at the expense of yield stability, in particular for maize. But, our results also suggest that there is space for increasing global average production without increasing its interannual variability and without increasing regional average yields (i.e., without intensifying production). This can theoretically be achieved by optimizing cropland allocation among the main producing regions. Our calculations suggest that, in the last decades, additional cultivated land have not been distributed in a way that alleviates global yield loss risks at least for maize and wheat, i.e., by bringing actual croplands closer to the optimal frontiers (figures S1 and S2). Incidentally, incentives for the cultivation of new land can be very diverse and possibly more regional than global (e.g., political or economical incentives to support staples, feed crops or biofuels). The tension between increasing production and decreasing global yield variance will presumably be exacerbated as climate change unfolds during the coming decades. Our framework could be use to deal with such issue, for example by computing anticipated optimal distribution based on yield projections. The most relevant scale to achieve this may be not the one considered here. Indeed, it is important to keep in mind, that large-scale crop yields dynamics (FAO regions are typically about 1000 km wide) may hide very important local disparities [4].

A few methodological and theoretical limitations should be outlined. Results from quadratic optimization are dependent on the set of selected regions. Our framework can be adapted to any other set of regions and to different geographical scales. Our optimization framework could be easily adapted to optimize cropland allocation within these regions based on finer scale datasets [37]. Providing instability measures of yield or production requires making a series of choices. The first one concerns the definition of instability. Yield variance and standard deviation are arguably the most straightforward rationales to measure risk from the point of view of total production. Coefficient of variation, distribution percentiles or expected shortfall are other routinely used risk indicators in various contexts [34, 38]. All these measurements are based on an estimation of a distance between observed and expected values, i.e., a trend. Yield trends can be inferred from time series using a variety of methods such as standard or local regression, moving average, smoothing filters, or dynamic linear modeling [12, 13, 18, 20, 21, 30, 31, 38, 39]. As each method has advantages and limitations, their preference mostly depends on the habits in the concerned field of research. But, although some level of uncertainty is associated with this first step, previous work shows low sensitivity of instability measures to a change of detrending methodology [21, 38]. Finally, our study also somehow relies on the premise that additional areas would be characterized by yield average and variance equivalent to the values obtained in existing cropland. This is uncertain, as additional croplands may be gained on degraded, marginal lands or permanent vegetation.

This study is based on a theoretical framework and as such, it conveys important agronomic and political impediments. First and foremost, relocating large parts of the global production is neither necessarily practical nor desirable. Our results suggest that an optimization of a small fraction of cropland could be sufficient to increase production and improve stability. By no means does our study advocates for drastic changes in worldwide cropland distribution. Undoubtedly, other criteria than expected yield and variance should be taken into account: national food (and feed) autonomy, agrobiodiversity or the sparing of land. However, as it has been shown in other contexts, optimization is a useful tool to explore land use patterns that are optimal for a wide selection of criteria [40, 41]. In fact, there is yet no international coordination supporting acreage stewardship for example in the form of a global common agricultural policy. Price-based mechanisms or market integration seem to have not succeeded to create a close-to-optimum situation.

Quadratic optimization on large-scale production and harvested area datasets indicate that strong trade-offs exist between yield levels and instability for the most cropped species. Large gains are theoretically

possible for both the levels and stability of global yields. Our framework could be applied to help deal with such tradeoffs at a scale compatible with decision making, e.g., at the scale of small administrative units.

References

- [1] FAO 2015 FAOSTAT Database on Agriculture (<http://faostat.fao.org/>) (accessed 1 September 2015)
- [2] Van Ittersum M K, Cassman K G, Grassini P, Wolf J, Tittonell P and Hochman Z 2013 Field crops research yield gap analysis with local to global relevance—a review *F Crop Res* **143** 4–17
- [3] Licker R *et al* 2010 Mind the gap: how do climate and agricultural management explain the ‘yield gap’ of croplands around the world? *PLoS One* **5** 1769–82
- [4] Mueller N D, Gerber J S, Johnston M, Ray D K, Ramankutty N and Foley J A 2012 Closing yield gaps through nutrient and water management *Nature* **490** 254–7
- [5] Schierhorn F, Faramarzi M, Prishchepov A V, Koch F J and Müller D 2014 Quantifying yield gaps in wheat production in Russia *Environ. Res. Lett.* **9** 084017
- [6] Cassman K G and Grassini P 2013 Can there be a green revolution in Sub-Saharan Africa without large expansion of irrigated crop production? *Glob. Food Sec.* **2** 203–9
- [7] Mueller N D, West P C, Gerber J S, MacDonald G K, Polasky S and Foley J A 2014 A tradeoff frontier for global nitrogen use and cereal production *Environ. Res. Lett.* **9** 054002
- [8] Cordell D, Drangert J-O and White S 2009 The story of phosphorus: global food security and food for thought *Glob. Environ. Change* **19** 292–305
- [9] Brown L R 1997 Facing the challenge of food scarcity: can we raise grain yields fast enough? *Plant Nutrition for Sustainable Food Production and Environment (Developments in Plant and Soil Sciences vol 78)* pp 15–24
- [10] Grassini P, Eskridge K M and Cassman K G 2013 Distinguishing between yield advances and yield plateaus in historical crop production trends *Nat. Commun.* **4** 2918
- [11] Brisson N, Gate P, Gouache D, Charmet G, Oury F-X and Huard F 2010 Why are wheat yields stagnating in Europe? A comprehensive data analysis for France *F Crop. Res.* **119** 201–12
- [12] Michel L and Makowski D 2013 Comparison of statistical models for analyzing wheat yield time series *PLoS One* **8** e78615
- [13] Calderini D F and Slafer G A 1998 Changes in yield and yield stability in wheat during the 20th century *F Crop. Res.* **57** 335–47
- [14] Field C B *et al* 2014 Food security and food production systems *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel of Climate Change* (Cambridge: Cambridge University Press) pp 485–533
- [15] Lobell D B, Schlenker W and Costa-Roberts J 2011 Climate trends and global crop production since 1980 *Science* **333** 616–20
- [16] Lobell D B and Tebaldi C 2014 Getting caught with our plants down: the risks of a global crop yield slowdown from climate trends in the next two decades *Environ. Res. Lett.* **9** 074003
- [17] Lobell D B *et al* 2014 Greater sensitivity to drought accompanies maize yield increase in the US Midwest *Science* **344** 516–9
- [18] Ben-Ari T and Makowski D 2014 Decomposing global crop yield variability *Environ. Res. Lett.* **9** 114011
- [19] Food and Agriculture Organisation (FAO) of the United Nations. Declaration of the World Summit on Food Security (www.fao.org/fileadmin/templates/wsfs/Summit/Docs/Final_Declaration/WSFS09_Declaration.pdf)
- [20] Ray D K, Gerber J S, MacDonald G K and West P C 2015 Climate variation explains a third of global crop yield variability *Nat. Commun.* **6** 5989
- [21] Ben-Ari T, Adrian J, Klein T, Calanca P, Van der Velde M and Makowski D 2016 Identifying indicators for extreme wheat and maize yield losses *Agric. For Meteorol.* **220** 130–40
- [22] Turlach B A and Weingessel A 2013 quadprog: Functions to Solve Quadratic Programming Problems (<http://cran.r-project.org/package=quadprog>)
- [23] Markowitz H M 1989 Mean—variance analysis *Finance* ed J Eatwell, M Milgate and P Newman (London: Palgrave Macmillan) pp 194–8
- [24] Cassman K G 1999 Ecological intensification of cereal production systems: yield potential, soil quality, and precision agriculture *Proc. Natl Acad. Sci.* **96** 5952–9
- [25] Neumann K, Verburg P H, Stehfest E and Müller C 2010 The yield gap of global grain production: a spatial analysis *Agric. Syst.* **103** 316–26
- [26] Si B C and Farrell R E 2004 Scale-dependent relationship between wheat yield and topographic indices *Soil. Sci. Soc. Am. J.* **68** 577
- [27] Amundson R, Berhe A A, Hopmans J W, Olson C, Szein A E and Sparks D L 2015 Soil and human security in the 21st century *Science* **348** 1261071
- [28] Lal R 2014 Desertification and soil erosion *Global Environmental Change* (Netherlands: Springer) pp 369–78
- [29] Iizumi T *et al* 2013 Prediction of seasonal climate-induced variations in global food production *Nat. Clim. Change* **3** 904–8
- [30] Lobell D B and Field C B 2007 Global scale climate–crop yield relationships and the impacts of recent warming *Environ. Res. Lett.* **2** 014002
- [31] Iizumi T *et al* 2014 Impacts of El Niño Southern Oscillation on the global yields of major crops *Nat. Commun.* **5** 3712
- [32] Beddow J M, Pardey P G, Koo J and Wood S 2010 The changing landscape of global agriculture *The Shifting Patterns of Agricultural Production and Productivity Worldwide* ed J M Alston, B A Babcock and P G Pardey (Ames, IA: The Midwest Agribusiness Trade Research and Information Center) pp 8–38
- [33] Hazell P B R 1985 Sources of increased variability in world cereal production since the 1960s *J. Agric. Econ.* **36** 145–59
- [34] Anderson J R, Hazell P B R and Lloyd T 1987 Variability of cereal yields for agricultural research and policy *Food Policy* **12** 199–212
- [35] Lesk C, Rowhani P and Ramankutty N 2016 Influence of extreme weather disasters on global crop production *Nature* **529** 84–7
- [36] Osborne T M and Wheeler T R 2013 Evidence for a climate signal in trends of global crop yield variability over the past 50 years *Environ. Res. Lett.* **8** 024001
- [37] Iizumi T *et al* 2014 Historical changes in global yields: major cereal and legume crops from 1982 to 2006 *Glob. Ecol. Biogeogr.* **23** 346–57
- [38] Cernay C, Ben-Ari T, Pelzer E, Meynard J and Makowski D 2015 Estimating variability in grain legume yields across Europe and the Americas *Sci. Rep.* **5** 11171
- [39] Licker R, Kucharik C J, Doré T, Lindeman M J and Makowski D 2013 Climatic impacts on winter wheat yields in Picardy, France and Rostov, Russia: 1973–2010 *Agric. For Meteorol.* **176** 25–37
- [40] Seppelt R, Lautenbach S and Volk M 2013 Identifying trade-offs between ecosystem services, land use, and biodiversity: a plea for combining scenario analysis and optimization on different spatial scales *Curr. Opin. Environ. Sustain.* **5** 458–63
- [41] Seppelt R and Voinov A 2003 Optimization methodology for land use patterns—evaluation based on multiscale habitat pattern comparison *Ecol. Modell.* **168** 217–31