LETTER

doi:10.1038/nature11069

Comparing the yields of organic and conventional agriculture

Verena Seufert[1], Navin Ramankutty[1] & Jonathan A. Foley[2]

Numerous reports have emphasized the need for major changes in the global food system: agriculture must meet the twin challenge of feeding a growing population, with rising demand for meat and high-calorie diets, while simultaneously minimizing its global environmental impacts[1,2]. Organic farming—a system aimed at producing food with minimal harm to ecosystems, animals or humans—is often proposed as a solution[3,4]. However, critics argue that organic agriculture may have lower yields and would therefore need more land to produce the same amount of food as conventional farms, resulting in more widespread deforestation and biodiversity loss, and thus undermining the environmental benefits of organic practices[5]. Here we use a comprehensive meta-analysis to examine the relative yield performance of organic and conventional farming systems globally. Our analysis of available data shows that, overall, organic yields are typically lower than conventional yields. But these yield differences are highly contextual, depending on system and site characteristics, and range from 5% lower organic yields (rain-fed legumes and perennials on weak-acidic to weak-alkaline soils), 13% lower yields (when best organic practices are used), to 34% lower yields (when the conventional and organic systems are most comparable). Under certain conditions—that is, with good management practices, particular crop types and growing conditions—organic systems can thus nearly match conventional yields, whereas under others it at present cannot. To establish organic agriculture as an important tool in sustainable food production, the factors limiting organic yields need to be more fully understood, alongside assessments of the many social, environmental and economic benefits of organic farming systems.

Although yields are only part of a range of ecological, social and economic benefits delivered by farming systems, it is widely accepted that high yields are central to sustainable food security on a finite land basis[1,2]. Numerous individual studies have compared the yields of organic and conventional farms, but few have attempted to synthesize this information on a global scale. A first study of this kind[6] concluded that organic agriculture matched, or even exceeded, conventional yields, and could provide sufficient food on current agricultural land. However, this study was contested by a number of authors; the criticisms included their use of data from crops not truly under organic management and inappropriate yield comparisons[7,8].

We performed a comprehensive synthesis of the current scientific literature on organic-to-conventional yield comparisons using formal meta-analysis techniques. To address the criticisms of the previous study[6] we used several selection criteria: (1) we restricted our analysis to studies of 'truly' organic systems, defined as those with certified organic management or non-certified organic management, following the standards of organic certification bodies (see Supplementary Information); (2) we only included studies with comparable spatial and temporal scales for both organic and conventional systems (see Methods); and (3) we only included studies reporting (or from which we could estimate) sample size and error. Conventional systems were either high- or low-input commercial systems, or subsistence agriculture. Sixty-six studies met these criteria, representing 62 study sites, and reporting 316 organic-to-conventional yield comparisons on 34 different crop species (Supplementary Table 4).

The average organic-to-conventional yield ratio from our meta-analysis is 0.75 (with a 95% confidence interval of 0.71 to 0.79); that is, overall, organic yields are 25% lower than conventional (Fig. 1a). This result only changes slightly (to a yield ratio of 0.74) when the analysis is limited to studies following high scientific quality standards (Fig. 2). When comparing organic and conventional yields it is important

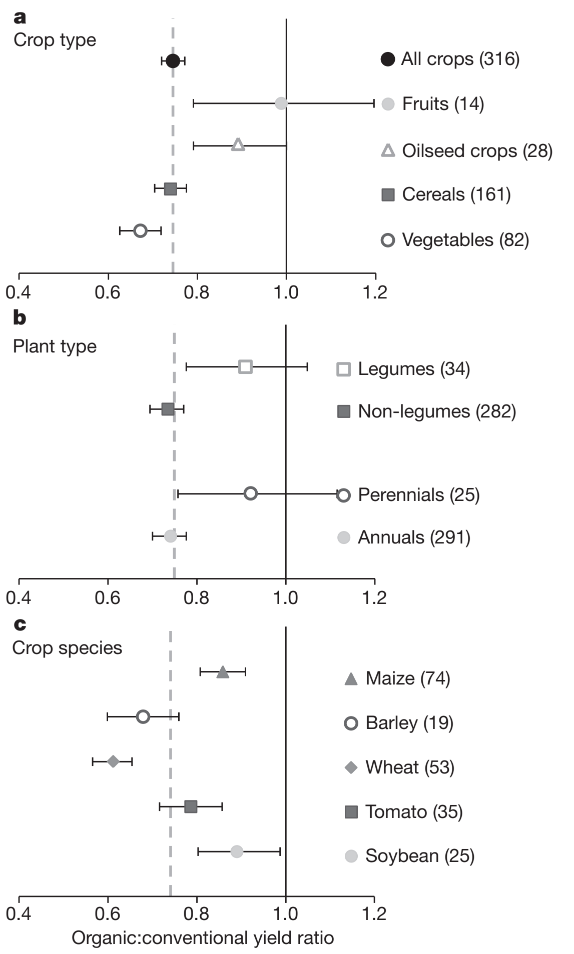

Figure 1 | Influence of different crop types, plant types and species on organic-to-conventional yield ratios. **a–c**, Influence of crop type (**a**), plant type (**b**) and crop species (**c**) on organic-to-conventional yield ratios. Only those crop types and crop species that were represented by at least ten observations and two studies are shown. Values are mean effect sizes with 95% confidence intervals. The number of observations in each class is shown in parentheses. The dotted line indicates the cumulative effect size across all classes.

[1]Department of Geography and Global Environmental and Climate Change Center, McGill University, Montreal, Quebec H2T 3A3, Canada. [2]Institute on the Environment (IonE), University of Minnesota, 1954 Buford Avenue, St Paul, Minnesota 55108, USA.

to consider the food output per unit area and time, as organic rotations often use more non-food crops like leguminous forage crops in their rotations[7]. However, the meta-analysis suggests that studies using longer periods of non-food crops in the organic rotation than conventional systems do not differ in their yield ratio from studies using similar periods of non-food crops (Fig. 2 and Supplementary Table 5). It thus appears that organic rotations do not require longer periods of non-food crops, which is also corroborated by the fact that the majority of studies (that is, 76%) use similar lengths of non-food crops in the organic and conventional systems.

The performance of organic systems varies substantially across crop types and species (Fig. 1a–c; see Supplementary Table 5 for details on categorical analysis). For example, yields of organic fruits and oilseed crops show a small (−3% and −11% respectively), but not statistically significant, difference to conventional crops, whereas organic cereals and vegetables have significantly lower yields than conventional crops (−26% and −33% respectively) (Fig. 1a).

These differences seem to be related to the better organic performance (referring to the relative yield of organic to conventional systems) of perennial over annual crops and legumes over non-legumes (Fig. 1b). However, note that although legumes and perennials (and fruits and oilseed crops) show statistically insignificant organic-to-conventional yield differences, this is owing to the large uncertainty range resulting from their relatively small sample size ($n = 34$ for legumes, $n = 25$ for perennials, $n = 14$ for fruits and $n = 28$ for oilseed crops; Fig. 1), and combining legumes and perennials reveals a significant, but small, yield difference (Fig. 2).

Part of these yield responses can be explained by differences in the amount of nitrogen (N) input received by the two systems (Fig. 3a). When organic systems receive higher quantities of N than conventional systems, organic performance improves, whereas conventional systems do not benefit from more N. In other words, organic systems appear to be N limited, whereas conventional systems are not. Indeed, N availability has been found to be a major yield-limiting factor in many organic systems[9]. The release of plant-available mineral N from organic sources such as cover crops, compost or animal manure is slow and often does not keep up with the high crop N demand during the peak growing period[9,10]. The better performance of organic legumes and perennials is not because they received more N, but rather because they seem to be more efficient at using N (Supplementary Table 7 and Supplementary Fig. 4). Legumes are not as dependent on external N sources as non-legumes, whereas perennials, owing to their longer growing period and extensive root systems, can achieve a better synchrony between nutrient demands and the slow release of N from organic matter[11].

Organic crops perform better on weak-acidic to weak-alkaline soils (that is, soils with a pH between 5.5. and 8.0; Fig. 3b). A possible explanation is the difficulty of managing phosphorus (P) in organic systems. Under strongly alkaline and acidic conditions, P is less readily available to plants as it forms insoluble phosphates, and crops depend to a stronger degree on soil amendments and fertilizers. Organic systems often do not receive adequate P inputs to replenish the P lost through harvest[12]. To test this hypothesis we need further research on the performance and nutrient dynamics of organic agriculture on soils of varying pH.

Studies that reported having applied best management practices in both systems show better organic performance (Fig. 3c). Nutrient and pest management in organic systems rely on biological processes to deliver plant nutrients and to control weed and herbivore populations. Organic yields thus depend more on knowledge and good management practices than conventional yields. However, in organic systems that are not N limited (as they grow perennial or leguminous crops, or apply large N inputs), best management practices are not required (Supplementary Table 11).

It is often reported that organic yields are low in the first years after conversion and gradually increase over time, owing to improvements in

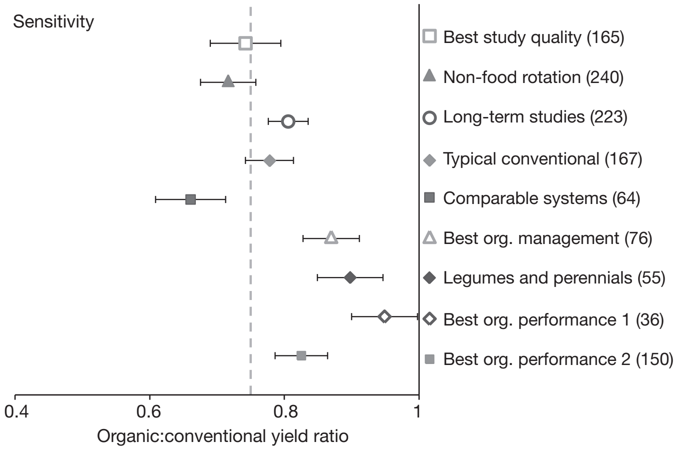

Figure 2 | Sensitivity study of organic-to-conventional yield ratios. Best study quality, peer-reviewed studies using appropriate study design and making appropriate inferences; non-food rotation, studies where both systems have a similar duration of non-food crops; long-term studies, excludes very short duration and recently converted studies; typical conventional, restricted to commercial conventional systems with yields comparable to local averages; comparable systems, studies that use appropriate study design and make appropriate inferences, where both systems have the same non-food rotation length and similar N inputs; best org. management, excludes studies without best management practices or crop rotations; legumes and perennials, restricted to leguminous and perennial crops; best org. performance 1, rain-fed legumes and perennials on weak-acidic to weak-alkaline soils; best org. performance 2, rain-fed and weak-acidic to weak-alkaline soils. Values are mean effect sizes with 95% confidence intervals. The number of observations is shown in parentheses. The dotted line indicates the effect size across all studies.

soil fertility and management skills[13]. This is supported by our analysis: organic performance improves in studies that lasted for more than two seasons or were conducted on plots that had been organic for at least 3 years (Fig. 2, Supplementary Fig. 5 and Supplementary Table 13).

Water relations also influence organic yield ratios—organic performance is −35% under irrigated conditions, but only −17% under rain-fed conditions (Fig. 3e). This could be due to a relatively better organic performance under variable moisture conditions in rain-fed systems. Soils managed with organic methods have shown better water-holding capacity and water infiltration rates and have produced higher yields than conventional systems under drought conditions and excessive rainfall[14,15] (see Supplementary Information). On the other hand, organic systems are often nutrient limited (see earlier), and thus probably do not respond as strongly to irrigation as conventional systems.

The majority of studies in our meta-analysis come from developed countries (Supplementary Fig. 1). Comparing organic agriculture across the world, we find that in developed countries organic performance is, on average, −20%, whereas in developing countries it is −43% (Fig. 3f). This poor performance of organic agriculture in developing countries may be explained by the fact that a majority of the data (58 of 67 observations) from developing countries seem to have atypical conventional yields (>50% higher than local yield averages), coming from irrigated lands (52 of 67), experimental stations (54 of 67) and from systems not using best management practices (67 of 67; Supplementary Fig. 10 and Supplementary Table 8). In the few cases from developing countries where organic yields are compared to conventional yields typical for the location or where the yield data comes from surveys, organic yields do not differ significantly from conventional yields because of a wide confidence interval resulting from the small sample size ($n = 8$ and $n = 12$ respectively, Supplementary Fig. 10a).

The results of our meta-analysis differ dramatically from previous results[6]. Although our organic performance estimate is lower than previously reported[6] in developed countries (−20% compared to −8%), our results are markedly different in developing countries

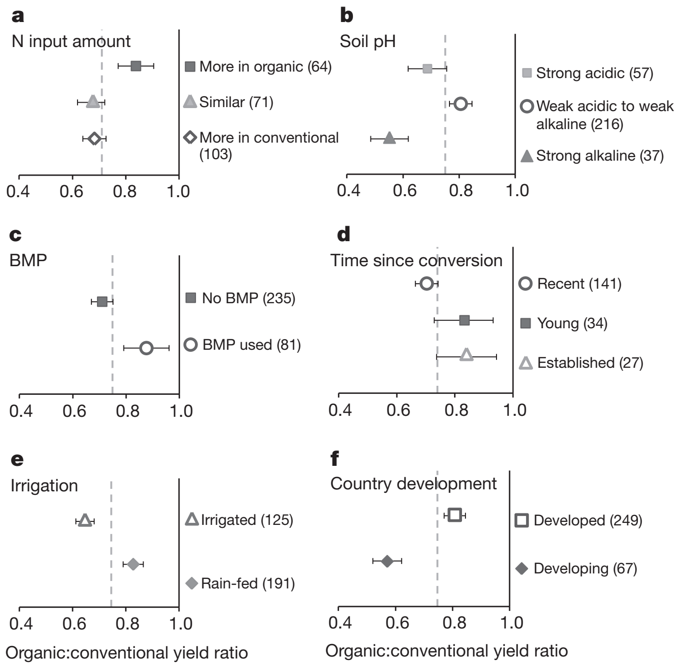

Figure 3 | Influence of N input, soil pH, best management practices, time since conversion to organic management, irrigation and country development. a–f, Influence of the amount of N input (**a**), soil pH (**b**), the use of best management practices (BMP; **c**), time since conversion to organic management (**d**), irrigation (**e**) and country development (**f**) on organic-to-conventional yield ratios. For details on the definition of categorical variables see Supplementary Tables 1–3. Values are mean effect sizes with 95% confidence intervals. The number of observations in each class is shown in parentheses. The dotted line indicates the cumulative effect size across all classes.

(−43% compared to +80%). This is because the previous analysis mainly included yield comparisons from conventional low-input subsistence systems, whereas our data set mainly includes data from high-input systems for developing countries. However, the previous study compared subsistence systems to yields that were not truly organic, and/or from surveys of projects that lacked an adequate control. Not a single study comparing organic to subsistence systems met our selection criteria and could be included in the meta-analysis. We cannot, therefore, rule out the claim[16] that organic agriculture can increase yields in smallholder agriculture in developing countries. But owing to a lack of quantitative studies with appropriate controls we do not have sufficient scientific evidence to support it either. Fortunately, the Swiss Research Institute of Organic Agriculture (FiBL) recently established the first long-term comparison of organic and different conventional systems in the tropics[17]. Such well-designed long-term field trials are urgently needed.

Our analysis shows that yield differences between organic and conventional agriculture do exist, but that they are highly contextual. When using best organic management practices yields are closer to (−13%) conventional yields (Fig. 2). Organic agriculture also performs better under certain agroecological conditions—for example, organic legumes or perennials, on weak-acidic to weak-alkaline soils, in rain-fed conditions, achieve yields that are only 5% lower than conventional yields (Fig. 2). On the other hand, when only the most comparable conventional and organic systems are considered the yield difference is as high as 34% (Fig. 2). In developed countries or in studies that use conventional yields that are representative of regional averages, the yield difference between comparable organic and conventional systems, however, goes down to 8% and 13%, respectively (see Supplementary Information).

In short, these results suggest that today's organic systems may nearly rival conventional yields in some cases—with particular crop types, growing conditions and management practices—but often they do not. Improvements in management techniques that address factors limiting yields in organic systems and/or the adoption of organic agriculture under those agroecological conditions where it performs best may be able to close the gap between organic and conventional yields.

Although we were able to identify some factors contributing to variations in organic performance, several other potentially important factors could not be tested owing to a lack of appropriate studies. For example, we were unable to analyse tillage, crop residue or pest management. Also, most studies included in our analysis experienced favourable growing conditions (Supplementary Fig. 8), and organic systems were mostly compared to commercial high-input systems (which had predominantly above-average yields in developing countries; Supplementary Figs 6b and 10a). In addition, it would be desirable to examine the total human-edible calorie or net energy yield of the entire farm system rather than the biomass yield of a single crop species. To understand better the performance of organic agriculture, we should: (1) systematically analyse the long-term performance of organic agriculture under different management regimes; (2) study organic systems under a wider range of biophysical conditions; (3) examine the relative yield performance of smallholder agricultural systems; and (4) evaluate the performance of farming systems through more holistic system metrics.

As emphasized earlier, yields are only part of a range of economic, social and environmental factors that should be considered when gauging the benefits of different farming systems. In developed countries, the central question is whether the environmental benefits of organic crop production would offset the costs of lower yields (such as increased food prices and reduced food exports). Although several studies have suggested that organic agriculture can have a reduced environmental impact compared to conventional agriculture[18,19], the environmental performance of organic agriculture per unit output or per unit input may not always be advantageous[20,21]. In developing countries, a key question is whether organic agriculture can help alleviate poverty for small farmers and increase food security. On the one hand, it has been suggested that organic agriculture may improve farmer livelihoods owing to cheaper inputs, higher and more stable prices, and risk diversification[16]. On the other hand, organic agriculture in developing countries is often an export-oriented system tied to a certification process by international bodies, and its profitability can vary between locations and years[22,23].

There are many factors to consider in balancing the benefits of organic and conventional agriculture, and there are no simple ways to determine a clear 'winner' for all possible farming situations. However, instead of continuing the ideologically charged 'organic versus conventional' debate, we should systematically evaluate the costs and benefits of different management options. In the end, to achieve sustainable food security we will probably need many different techniques—including organic, conventional, and possible 'hybrid' systems[24]—to produce more food at affordable prices, ensure livelihoods for farmers, and reduce the environmental costs of agriculture.

METHODS SUMMARY

We conducted a comprehensive literature search, compiling scientific studies comparing organic to conventional yields that met our selection criteria. We minimized the use of selection criteria based on judgments of study quality but examined its influence in the categorical analysis. We collected information on several study characteristics reported in the papers and derived characteristics of the study site from spatial global data sets (see Supplementary Tables 1–3 for a description of all categorical variables). We examined the difference between organic and conventional yields with the natural logarithm of the response ratio (the ratio between organic and conventional yields), an effect size commonly used in meta-analyses[25]. To calculate the cumulative effect size we weighted each individual observation by the inverse of the mixed-model variance. Such a categorical meta-analysis should be used when the data have some underlying structure and individual observations can be categorized into groups (for example, crop species or fertilization practices)[26]. An effect size is considered significant if its confidence interval does not overlap with 1 in the back-transformed response ratio. To test the influence of categorical variables on yield effect sizes we examined between-group heterogeneity (Q_B). A significant Q_B indicates that there are differences in effect sizes between different classes of a categorical variable[26]. All statistical analyses were carried out in MetaWin 2.0[26].

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 6 November 2011; accepted 9 March 2012.
Published online 25 April 2012.

1. Godfray, H. C. *et al.* Food security: the challenge of feeding 9 billion people. *Science* **327,** 812–818 (2010).
2. Foley, J. *et al.* Solutions for a cultivated planet. *Nature* **478,** 337–342 (2011).
3. McIntyre, B. D., Herren, H. R., Wakhungu, J. & Watson, R. T. *International Assessment of Agricultural Knowledge, Science and Technology for Development: Global Report* http://www.agassessment.org/ (Island, 2009).
4. De Schutter, O. *Report Submitted by the Special Rapporteur on the Right to Food* http://www2.ohchr.org/ english/issues/food/docs/A-HRC-16–49.pdf (United Nations, 2010).
5. Trewavas, A. Urban myths of organic farming. *Nature* **410,** 409–410 (2001).
6. Badgley, C. *et al.* Organic agriculture and the global food supply. *Renew. Agr. Food Syst.* **22,** 86–108 (2007).
7. Cassman, K. G. Editorial response by Kenneth Cassman: can organic agriculture feed the world-science to the rescue? *Renew. Agr. Food Syst.* **22,** 83–84 (2007).
8. Connor, D. J. Organic agriculture cannot feed the world. *Field Crops Res.* **106,** 187–190 (2008).
9. Berry, P. *et al.* Is the productivity of organic farms restricted by the supply of available nitrogen? *Soil Use Manage.* **18,** 248–255 (2002).
10. Pang, X. & Letey, J. Organic farming: challenge of timing nitrogen availability to crop nitrogen requirements. *Soil Sci. Soc. Am. J.* **64,** 247–253 (2000).
11. Crews, T. E. & Peoples, M. B. Can the synchrony of nitrogen supply and crop demand be improved in legume and fertilizer-based agroecosystems? A review. *Nutr. Cycl. Agroecosyst.* **72,** 101–120 (2005).
12. Oehl, F. *et al.* Phosphorus budget and phosphorus availability in soils under organic and conventional farming. *Nutr. Cycl. Agroecosyst.* **62,** 25–35 (2002).
13. Martini, E., Buyer, J. S., Bryant, D. C., Hartz, T. K. & Denison, R. F. Yield increases during the organic transition: improving soil quality or increasing experience? *Field Crops Res.* **86,** 255–266 (2004).
14. Letter, D., Seidel, R. & Liebhardt, W. The performance of organic and conventional cropping systems in an extreme climate year. *Am. J. Altern. Agric.* **18,** 146–154 (2003).
15. Colla, G. *et al.* Soil physical properties and tomato yield and quality in alternative cropping systems. *Agron. J.* **92,** 924–932 (2000).
16. Scialabba, N. & Hattam, C. *Organic Agriculture, Environment and Food Security* (Food and Agriculture Organization, 2002).
17. Research Institute of Organic Agriculture (FiBL). *Farming System Comparison in the Tropics* http://www.systems-comparison.fibl.org/ (2011).
18. Crowder, D. W., Northfield, T. D., Strand, M. R. & Snyder, W. E. Organic agriculture promotes evenness and natural pest control. *Nature* **466,** 109–112 (2010).
19. Bengtsson, J., Ahnström, J. & Weibull, A.-C. The effects of organic agriculture on biodiversity and abundance: a meta-analysis. *J. Appl. Ecol.* **42,** 261–269 (2005).
20. Kirchmann, H. & Bergström, L. Do organic farming practices reduce nitrate leaching? *Commun. Soil Sci. Plan.* **32,** 997–1028 (2001).
21. Leifeld, J. & Fuhrer, J. Organic farming and soil carbon sequestration: what do we really know about the benefits? *Ambio* **39,** 585–599 (2010).
22. Valkila, J. Fair trade organic coffee production in Nicaragua—sustainable development or a poverty trap? *Ecol. Econ.* **68,** 3018–3025 (2009).
23. Raynolds, L. T. The globalization of organic agro-food networks. *World Dev.* **32,** 725–743 (2004).
24. National Research Council. *Toward Sustainable Agricultural Systems in the 21st Century* (National Academies, 2010).
25. Hedges, L. V., Gurevitch, J. & Curtis, P. S. The meta-analysis of response ratios in experimental ecology. *Ecology* **80,** 1150–1156 (1999).
26. Rosenberg, M. S., Gurevitch, J. & Adams, D. C. *MetaWin: Statistical Software for Meta-analysis: Version 2* (Sinauer, 2000).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to the authors of the 66 studies whose extensive field work provided the data for this meta-analysis. Owing to space limitations our citations can be found in Supplementary Material. We would like to thank J. Reganold for useful comments on our manuscript. We are grateful to I. Perfecto, T. Moore, C. Halpenny, G. Seufert and S. Lehringer for valuable discussion and/or feedback on the manuscript and L. Gunst for sharing publications on the FiBL trials. D. Plouffe helped with the figures and M. Henry with compiling data. This research was supported by a Discovery Grant awarded to N.R. from the Natural Science and Engineering Research Council of Canada.

Author Contributions V.S. and N.R. designed the study. V.S. compiled the data and carried out data analysis. All authors discussed the results and contributed to writing the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to V.S. (verena.seufert@mail.mcgill.ca).

METHODS

Literature search. We searched the literature on studies reporting organic-to-conventional yield comparisons. First we used the references included in the previous study[6] and then extended the search by using online search engines (Google scholar, ISI web of knowledge) as well as reference lists of published articles. We applied several selection criteria to address the criticisms of the previous study[6] and to ensure that minimum scientific standards were met. Studies were only included if they (1) reported yield data on individual crop species in an organic treatment and a conventional treatment, (2) the organic treatment was truly organic (that is, either certified organic or following organic standards), (3) reported primary data, (4) the scale of the organic and conventional yield observations were comparable, (5) data were not already included from another paper (that is, avoid multiple counting), and (6) reported the mean (X), an error term (standard deviation (s.d.), standard error (s.e.) or confidence interval) and sample size (n) as numerical or graphical data, or if X and s.d. of yields over time could be calculated from the reported data. For organic and conventional treatments to be considered comparable, the temporal and spatial scale of the reported yields needed to be the same, that is, national averages of conventional agriculture compared to national averages of organic agriculture or yields on an organic farm compared to yields on a neighbouring conventional farm—not included were, for example, single farm yields compared to national or regional averages or before–after comparisons. Previous studies[27] have illustrated the danger of comparing yield data drawn from single plots and field trials to larger state and national averages.

The use of selection criteria is a critical step in conducting a meta-analysis. On the one hand, scientific quality and comparability of observations needs to be ensured. On the other hand, a meta-analysis should provide as complete a summary of the current research as possible. There is an ongoing debate about whether meta-analyses should adopt very specific selection criteria to prevent mixing incomparable data sets together and to minimize variation in the data set[28] or whether, instead, meta-analyses should include as wide a range of studies as possible to allow for an analysis of sources of variation[29]. We followed the generally recommended approach, trying to minimize the use of selection criteria based on judgments of study quality[30]. Instead, we examined the influence of quality criteria empirically by evaluating the differences between observations with different quality standards. We did not therefore exclude yield observations from non-peer-reviewed sources or from studies that lacked an appropriate experimental design a priori. The quality of the study and the comparability of the organic and conventional systems were assessed by evaluating the experimental design of the study as well as the form of publication. Studies that were published in peer-reviewed journals and that controlled for the possible influence of variability in space and time on experimental outcomes through an appropriate experimental design were considered to follow high quality standards.

Categorical variables. In addition to study quality criteria, information on several other study characteristics like crop species, location and timescale, and on different management practices, was collected (see Supplementary Tables 1–3). We also wanted to test the effect of study site characteristics on yield ratios and we thus collected information on biophysical characteristics of the study site. As most studies did not report climate or soil variables we derived information on several agroecological variables that capture cropland suitability[31], including the moisture index α (the ratio of actual to potential evapotranspiration) as an indicator of moisture availability to crops, growing degree days (GDD, the annual sum of daily mean temperatures over a base temperature of 5 °C) as an indicator of growing season length, as well as soil carbon density (C_{soil}, as a measure of soil organic content) and soil pH as indicators of soil quality from the latitude × longitude values of the study site and global spatial models/data sets at 5 min resolution[32,33].

We derived the thresholds for the classification of these climate and soil variables from the probability of cultivation functions previously described[31]. This probability of cultivation function is a curve fitted to the empirical relationship between cropland areas, α, GDD or C_{soil}. It describes the probability that a location with a certain climate or soil characteristic is covered by cropland. Suitable locations with favourable climate and soil characteristics have a higher probability of being cultivated. Favourable climate and soil characteristics can thus be inferred from the probability of cultivation. For α, GDD and C_{soil} a probability of cultivation under 30% was classified as 'low' suitability, between 30% and 70% as 'medium' suitability, and above 70% as 'high' suitability (Supplementary Table 3). Sites with low and medium suitable moisture indices are interpreted as having insufficient water availability, sites with low and medium GDD have short growing seasons, and sites with low and medium soil carbon densities are either unfertile because they have too small a C_{soil} and low organic matter content (and thus insufficient nutrients) or too high a C_{soil} in soils in wetlands where organic matter accumulates because they are submerged under water. For soil pH, instead, we defined thresholds based on expert judgment. Soil pH information was often given in the studies and we only derived soil pH values from the global data set if no soil pH value was indicated in the paper.

To assess whether the conventional yield values reported by studies and included in the meta-analysis were representative of regional average crop yields, we compared them to FAOSTAT yield data and a high-resolution spatial yield data set[34,35]. We used the FAO data[35], which reports national yearly crop yields from 1961 to 2009, for temporal detail and a yield data set[34], which reports sub-national crop yields for 175 crops for the year 2000 at a 5-min latitude by 5-min longitude resolution, for spatial detail. We calculated country average crop yields from FAO data for the respective study period and calculated the ratio of this average study-period yield to the year-2000 FAO national yield value. We derived the year-2000 yield value from the spatial data set through the latitude by longitude value of the study site and scaled this value to the study-period-to-year-2000 ratio from FAOSTAT. If the meta-analysis conventional yield value was more than 50% higher than the local yield average derived by this method it was classified as 'above average', when it was more than 50% lower as 'below average', and when it was within ±50% of local yield averages as 'comparable'. We choose this large yield difference as a threshold to account for uncertainties in the FAOSTAT and global yield data set[34].

Meta-analysis. The natural log of the response ratio[25] was used as an effect size metric for the meta-analysis. The response ratio is calculated as the ratio between the organic and the conventional yield. The use of the natural logarithm linearizes the metric (treating deviations in the numerator and the denominator the same) and provides more normal sampling distribution in small samples[25]. If the data set has some underlying structure and studies can be categorized into more than one group (for example, different crop species, or different fertilizer types) a categorical meta-analysis can be conducted[26]. Observations with the same or similar management or system characteristics were grouped together. We then used a mixed effects model to partition the variance of the sample, assuming that there is random variation within a group and fixed variation between groups. We calculated a cumulative effect size as weighted mean from all studies by weighting each individual observation by the reciprocal of the mixed-model variance, which is the sum of the study sampling variance and the pooled within-group variance. Weighted parametric meta-analysis should be used whenever possible to deal with heteroscedasticity in the sample and to increase the statistical power of the analysis[36]. The cumulative effect size is considered to be significantly different from zero (that is, the organic treatment shows a significant effect on crop yield) if its 95% confidence interval does not overlap zero.

To test for differences in the effect sizes between groups the total heterogeneity of the sample was partitioned into the within group (Q_W) and between group heterogeneity (Q_B) in a process similar to an analysis of variance[37]. The significance of Q_B was tested by comparing it against the critical value of the χ^2 distribution. A significant Q_B implies that there are differences among cumulative effect sizes between groups[26,38]. Only those effects that showed a significant Q_B are presented in graphs. All statistical analyses were carried out using MetaWin 2.0[26]. For representation in graphs effect sizes were back-transformed to response ratios.

Each observation in a meta-analysis is required to be independent. Repeated measurements in the same location over time are not independent. If yield values from a single experiment were reported over several years therefore the average yield over time was calculated and used in the meta-analysis. If the mean and variance of multiple years was reported, the weighted average over time was calculated by weighting each year by the inverse of its variance. Different experiments (for example, different tillage practices, crop species or fertilizer rates) from the same study are not necessarily independent. However, it is recommended to still include different experiments from the same study, as their omission would cause more distortions of the results than the lack of true independence[38]. We therefore included different experiments from a single study separately in the meta-analysis.

If data from the same experiment from the same study period were reported in several papers, the data were only included once, namely from the paper that reported the data in the highest detail (that is, reporting s.e./s.e. and n and/or reporting the longest time period). If instead data from the same experiment from different years were reported in separate papers, the data were included separately in the analysis (for example, refs 39, 40).

In addition to potential within-study dependence of effect size data, there can also be issues with between-study dependence of data[36]—data from studies conducted by the same author, in the same location or on the same crop species are also potentially non-independent. We addressed this issue by conducting a hierarchical, categorical meta-analysis (as described earlier), specifically testing for the influence of numerous moderators on the effect size. In addition, we examined the interaction between categorical variables through a combination of contingency tables and sub-categorical analysis (see Supplementary Information for the results of this analysis and for a more detailed discussion of this issue).

We performed a sensitivity analysis (see Supplementary Table 14) to compare the robustness of results under more strict quality criteria (see discussion of definition of study quality earlier) and to assess organic yield ratios under a couple of specific system comparisons.

27. Johnston, M., Foley, J. A., Holloway, T., Kucharik, C. & Monfreda, C. Resetting global expectations from agricultural biofuels. *Environ. Res. Lett.* **4,** 014004 (2009).
28. Whittaker, R. J. Meta-analyses and mega-mistakes: calling time on meta-analysis of the species richness–productivity relationship. *Ecology* **91,** 2522–2533 (2010).
29. Hillebrand, H. & Cardinale, B. J. A critique for meta-analyses and the productivity–diversity relationship. *Ecology* **91,** 2545–2549 (2010).
30. Englund, G., Sarnelle, O. & Cooper, S. D. The importance of data-selection criteria: meta-analyses of stream predation experiments. *Ecology* **80,** 1132–1141 (1999).
31. Ramankutty, N., Foley, J. A., Norman, J. & McSweeney, K. The global distribution of cultivable lands: current patterns and sensitivity to possible climate change. *Glob. Ecol. Biogeogr.* **11,** 377–392 (2002).
32. Deryng, D., Sacks, W., Barford, C. & Ramankutty, N. Simulating the effects of climate and agricultural management practices on global crop yield. *Glob. Biogeochem. Cycles* **25,** GB2006 (2011).
33. IGBP-DIS. *Soildata (V 0): A Program for Creating Global Soil-Property Databases* (IGBP Global Soils Data Task, 1998).
34. Monfreda, C., Ramankutty, N. & Foley, J. A. Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Glob. Biogeochem. Cycles* **22,** GB1022 (2008).
35. Food and Agriculture Organization of the United Nations (FAO). *FAOSTAT* http://faostat.fao.org (2011).
36. Gurevitch, J. & Hedges, L. V. Statistical issues in ecological meta-analyses. *Ecology* **80,** 1142–1149 (1999).
37. Hedges, L. V. & Olkin, I. *Statistical Methods for Meta-Analysis*. (Academic, 1985).
38. Gurevitch, J., Morrow, L. L., Wallace, A. & Walsh, J. S. A meta-analysis of competition in field experiments. *Am. Nat.* **140,** 539–572 (1992).
39. Liebhardt, W. *et al.* Crop production during conversion from conventional to low-input methods. *Agron. J.* **81,** 150–159 (1989).
40. Drinkwater, L., Janke, R. & Rossoni-Longnecker, L. Effects of tillage intensity on nitrogen dynamics and productivity in legume-based grain systems. *Plant Soil* **227,** 99–113 (2000).