Diversification practices reduce organic to conventional yield gap

Lauren C. Ponisio1, Leithen K. M’Gonigle1,2, Kevi C. Mace1, Jenny Palomino1, Perry de Valpine1 and Claire Kremen1

1Department of Environmental Science, Policy, and Management, University of California, Berkeley, 130 Mulford Hall, Berkeley, CA 94720, USA
2Department of Biological Science, Florida State University, Tallahassee, FL 32306, USA

Agriculture today places great strains on biodiversity, soils, water and the atmosphere, and these strains will be exacerbated if current trends in population growth, meat and energy consumption, and food waste continue. Thus, farming systems that are both highly productive and minimize environmental harms are critically needed. How organic agriculture may contribute to world food production has been subject to vigorous debate over the past decade. Here, we revisit this topic comparing organic and conventional yields with a new meta-dataset three times larger than previously used (115 studies containing more than 1000 observations) and a new hierarchical analytical framework that can better account for the heterogeneity and structure in the data. We find organic yields are only 19.2% (+3.7%) lower than conventional yields, a smaller yield gap than previous estimates. More importantly, we find entirely different effects of crop types and management practices on the yield gap compared with previous studies. For example, we found no significant differences in yields for leguminous versus non-leguminous crops, perennials versus annuals or developed versus developing countries. Instead, we found the novel result that two agricultural diversification practices, multi-cropping and crop rotations, substantially reduce the yield gap (to 9 ± 4% and 8 ± 5%, respectively) when the methods were applied in only organic systems. These promising results, based on robust analysis of a larger meta-dataset, suggest that appropriate investment in agroecological research to improve organic management systems could greatly reduce or eliminate the yield gap for some crops or regions.

1. Introduction

While tremendously productive, our current agricultural food system causes many environmental problems, often trading off long-term maintenance of ecosystem services for short-term agricultural production [1,2]. Resultant problems include biodiversity loss, massive soil erosion and degradation, eutrophication and oceanic dead zones, pesticide effects on humans and wildlife, greenhouse gas emissions, and regime shifts in hydrological cycling [3–12]. Furthermore, although agriculture produces a food surplus at the global scale, over 1 billion people are chronically hungry. These problems of hunger, food insecurity and environmental harms will only be exacerbated if current trends in population growth, food and energy consumption, and food waste continue [13–15]. To maintain the Earth’s capacity to produce food, it is imperative that we adopt sustainable and resilient agricultural practices as soon as possible [16,17]. Yet it is also broadly perceived that such practices will produce lower yields [17–19], leading to a conundrum—how do we maintain or increase food production without sacrificing sustainability and resilience? Previous analyses have concluded that improving the distribution of food while also reducing waste and meat consumption will greatly contribute to sustainably meeting future global demands [9,15], although how these goals are to be achieved is not yet clear.
Advocates of ‘sustainable intensification’ have focused on increasing production efficiency while minimizing economic and environmental costs. An emphasis on efficiency, however, may not necessarily lead to development of sustainable, resilient production systems that can buffer unexpected changes resulting from the complex socio-ecological interactions that influence agriculture [17,19,20]. To achieve environmental sustainability, we must grow food in a manner that protects, uses and regenerates ecosystem services (e.g. favours natural pest control over the use of synthetic pesticides), rather than replacing them [19,21,22]. Replacing ecosystem services often has unintended, negative consequences (e.g. lethal or sub-lethal effects of pesticides on humans, beneficial insects and wildlife [7,10,12,23]). Broad adoption of sustainable agricultural methods is unlikely, however, unless such methods are similarly productive and/or cost-effective, such that they improve livelihoods. Hence, there is much incentive to determine whether a yield gap exists between ‘conventional’ agriculture (i.e. chemically intensive and biologically simplified) and alternative, more sustainable forms of agriculture, and if so, how it can be reduced or eliminated.

Such systems (e.g. agroecological, ecologically intensive, biologically diversified or regenerative farming systems) use cultivation techniques that, through plot- to landscape-scale diversification, specifically encourage ecological interactions that generate soil fertility, nutrient cycling and retention, water storage, pest/disease control, pollination, and other essential agricultural inputs/ecosystem services [22]. The most widely practised and studied alternative to conventional agriculture is organic, which now takes place on 0.9% of agricultural lands [24]. Organic agriculture is defined as having no synthetic inputs, but organic farms may or may not practise full suite of cultivation techniques characterizing sustainable agriculture [21,25]. Although the terms ‘organic’ and ‘sustainable’ agriculture are not equivalent, studies of organic agriculture have revealed better performance than conventional systems on some (but not all) sustainability metrics, including species richness and abundance, soil fertility, nitrogen uptake by crops, water infiltration and holding capacity, and energy use and efficiency [26–32]. Here, we provide the most comprehensive calculation of the yield gap between organic and conventional agriculture, building on the work of others [33–36].

Early reviews comparing organic to conventional agriculture found yield gaps of 8–9% in developed countries [33,34], but yield gains of as much as 180% in developing countries. Two recent meta-analyses, however, found organic yields to be 20–25% lower than conventional yields [35,36]. That these studies differed so much in their conclusions can probably be attributed to two factors. First, each study used different criteria for selecting the data to be included in its review or meta-analysis. For example, for developing countries, Badgley et al. [34] focused primarily on comparing sites using techniques of sustainable agriculture with ‘resource-poor’ sites, rather than strict organic versus conventional comparisons, accounting for the yield gains they found for ‘organic management’ in developing countries [18,21,37]. Second, each of the above studies used different analytical methods to combine the data across the different sub-studies. For example, the reviews of de Ponti et al. [35], Stanhill [33] and Badgley et al. [34] did not account for the sampling variance within studies, which is the recommended practice to deal with heteroscedasticity in the sample of studies [38]. Seufert et al.’s [36] meta-analysis, while accounting for sampling variance, combined nested data (e.g. several experiments reported within the same study) without accounting for the hierarchy (electronic supplementary material, §S1). This introduced pseudo-replication that effectively understated the Type I error rate of their analysis by an order of magnitude (electronic supplementary material, figure S1), and biased estimates of the yield gap and its statistical uncertainty (electronic supplementary material, figure S2). Given these methodological and data-related critiques, a new study is needed to produce a more robust estimate of the gap between organic and conventional yields.

Here, we develop a hierarchical meta-analytic framework that overcomes the methodological pitfalls of previous studies by accounting for both the multi-level nature of the data and the yield variation within studies. Furthermore, via a literature search we compiled a more extensive and up-to-date meta-dataset, comprising 1071 organic versus conventional yield comparisons from 115 studies—over three times the number of observations of any of the previous analyses. Our meta-dataset includes studies from 38 countries and 52 crop species over a span of 35 years.

2. Material and methods

(a) Search details

In our search, we used similar terms to those employed by Seufert et al. [36] and de Ponti et al. [35]. The search term used was a complex Boolean search containing (i) the term ‘organic’ or ‘ecological’ and (ii) the term ‘agriculture’, ‘farming’, ‘production’ or ‘cropping’ in combination with (iii) terms equal or similar to the terms ‘yield’ and ‘compare’. We used the search engines Academic Compete Search, Google Scholar and Web of Science. The last search was conducted in January 2013. The complete list of studies and yield data are provided in the electronic supplementary material, table S5.

(b) Inclusion criteria

We adopted Seufert et al.’s [36] rigorous inclusion criteria, except we excluded (i) comparisons of subsistence yields (unimproved agriculture) against improved agricultural methods (e.g. [34,39]) and (ii) comparisons of yields taken from different years. Additionally, in cases where the means of organic and conventional yields were reported but the variance of those means were not (a necessary component for inclusion in meta-analysis), we obtained an estimate of the variance directly from the original authors, whenever possible. Of the 99 studies lacking variance estimates, we obtained variance estimates or original data from the authors of 28 of them. In cases where the authors did not reply and there were multiple years of data reported, we took the mean and variance across years (59 studies, 232 organic versus conventional comparisons), as did Seufert et al. [36]. Because the variance across years is not a perfect estimate of the within-year variation, however, we also conducted analyses excluding these studies (electronic supplementary material, §S2.5 and figure S2). Together, the search and data request yielded 115 studies from which we extracted 1071 organic versus conventional comparisons.

(c) Meta-analytic model

We built a hierarchical meta-analytic model to generate an estimate of the yield gap (see the electronic supplementary material, §S2, for details). Following standard practice, we compared the natural log of the ratios between organic and conventional yields (the
‘response ratio’) across studies [36,40]. The response ratio is more normally distributed than the raw ratio and independent of the units of measurement used within a study, and thus is comparable across studies [40].

To analyse the yield data, which contained several levels of hierarchy, we employed two methods. First, for studies that compare multiple treatments (usually organic) with one control treatment (usually conventional), and are thus not independent [41,42], we calculated a combined response ratio and corresponding standard error for the entire study using the method presented in eqn 3 and 8 in [41], and then used these combined response ratios in the nested analysis.

Next, we constructed a hierarchical regression model in a Bayesian framework to account for the dependencies in the yield data. We expanded on the traditional random effects model [43] by considering three additional sources of random variation (i.e. random effects): (1) between studies, (2) within a study between years and (3) within a year between response ratios (e.g. across replicated trials of a crop planted at different times in the season). We also considered whether the variances of the random effect distributions for (2) and (3) were shared across studies, or study-specific.

The traditional random effects meta-analytic model includes a random effect of study (electronic supplementary material, figure S1). We then added the additional random effects sequentially and determined whether the posterior distribution of the added parameter was clearly differentiated from zero (electronic supplementary material, Section S2). If it was, we concluded that the data supported adding that layer of hierarchy. We confirmed our selected hierarchical model, prior to model selection, with all sources of random variation supported adding that layer of hierarchy. We confirmed our model selection using the deviance information criterion, which agreed in this case [44,45]. The full possible model, prior to model selection, with all sources of random variation is

\[
\begin{align*}
    y_{ijk} & = \mu + \alpha_i + \beta_j + \eta_{ih} + \epsilon_{ijk}, \\
    \alpha_i & \sim N(0, \sigma^2_{\alpha}), \\
    \beta_j & \sim N(0, \sigma^2_{\beta}), \\
    \eta_{ih} & \sim N(0, \sigma^2_{\eta}), \\
    \epsilon_{ijk} & \sim N(0, \sigma^2_{\epsilon}), \\
    \sigma^2_{\alpha} & \sim \Gamma(CV_{\alpha}, \text{scale}) \\
    \text{and } \sigma^2_{\beta} & \sim \Gamma(CV_{\beta}, \text{scale}),
\end{align*}
\]

where \( y_{ijk} \) is the observed magnitude of the \( k \)th response ratio from the \( j \)th year of the \( i \)th study, \( \mu \) is the mean response ratio across studies, \( \alpha_i \) is the effect of \( i \)th study, \( \beta_j \) is the effect of \( j \)th year of the \( i \)th study, \( \eta_{ih} \) is the effect of the \( h \)th category of a specific explanatory variable (e.g. crop species), \( \epsilon_{ijk} \) is the residual. \( \sigma^2_{\alpha} \) is the between-study variance, \( \sigma^2_{\beta} \) is the within-year, between-response ratio variance of study \( i \), \( \sigma^2_{\eta} \) is the within-year, between-response ratio variance of study \( i \), and \( \sigma^2_{\epsilon} \) is the variance of response ratio \( y_{ijk} \) as reported by its study. \( CV_{\alpha} \) and \( CV_{\beta} \) and scale\( \alpha \) and scale\( \beta \) are the coefficient of variation and scale parameters of the gamma distributions of the study-specific between- and within-year variances. When response ratios that shared a common control were combined, \( y_{ijk} \) corresponds to the aggregate within-study response ratio (eqn 3 in [41]) and \( \sigma^2_{\eta} \) is its pooled variance (eqn 8 in [41]).

We also extended this model in order to accommodate analyses of study characteristics such as crop type and management practices. We analyse these additional explanatory variables one at a time because not all studies reported all explanatory variables. In these analyses, for cases where multiple organic treatments represented different categories for a specific explanatory variable, they could not be combined using Lajeunesse’s method [41]. The potential bias resulting from non-independence of the response ratios in these cases, however, would be minimized by the fact that they are not pooled together in the analysis [41].

Letting \( h \) index the categories for a particular explanatory variable (e.g. crop species), we then have

\[
y_{ijk} = \mu + \gamma_h + \alpha_i + \beta_j + \eta_{ih} + \epsilon_{ijk},
\]

where \( \gamma_h \) is the effect of the \( h \)th category, and the rest of the model parallels that given in equation (2.1).

In order to facilitate comparison between our results and those of previous analyses, we used the same categories as those defined by Seufert et al. [36]. We also examined the sensitivity of our results to explanatory variables related to study quality, again using the study quality categories defined by Seufert et al. [36]. The coefficients of explanatory variables were considered to be different from each other if the posterior of the difference between the 95% credibility intervals around the group means did not overlap zero.

We used JAGS through the R package rjags interface [46,47] to implement Markov chain Monte Carlo (MCMC) sampling. Inference was made from three chains each with 10³ samples of the posterior distribution after a burn-in of 10² and with a thin rate of 10¹. We used Gaussians with large variances to define priors, except for variance terms, where we used a uniform (0, 100) prior on the standard deviation. Initial values were chosen randomly. Convergence was assessed by visual assessment of MCMC chains and using the Gelman–Rubin statistic (‘Rhat’ in R package R2jags, with values less than 1.1 indicating convergence [48]). Credible intervals around parameter estimates were calculated as the 2.5% and 97.5% quantiles of the posterior. We also checked for bias in our meta-dataset using a funnel plot and QQ-plot (electronic supplementary material, figure S3).

3. Results

We found the data supported including random variation between studies and study-specific variation with a year, but not random variation between years of a study (electronic supplementary material, table S1 and figure S6). Using the selected hierarchical model, we found a smaller yield gap between conventional and organic cropping systems than that reported in recent meta-analyses [35,36]. We found that organic yields were 19.2% lower than conventional yields, with a 95% credible interval ranging from 15.5% to 22.9% (figure 1). Conventional yields were significantly higher than organic for all crop types and the yield ratios of most crop types did not vary significantly from one another (figure 1). At the finer scale of crop species, however, yield ratios differed significantly between some pairs of species (electronic supplementary material, figure S3).

The most dramatic difference between our findings and earlier work is an almost complete lack of significant differences between groups for all of the explanatory variables investigated. Unlike Seufert et al. [36], we did not find significant differences in yields for leguminous and non-leguminous crops nor for perennial and annuals (figure 1). Nor did we find a difference between the yield gaps for studies conducted in developed versus developing countries (electronic supplementary material, table S2; see also [34]). Our results were robust to including a between-year random effect (though including this parameter was not supported by the data; electronic supplementary material, table S1 and figure S6).
et al. is that Seufert and the majority of the explanatory variables between our study and that of Seufert for the crop rotations. The yield gap between organic polycultures and conventional monocultures (9 ± 4%) was significantly smaller than when both treatments were monocultures (17 ± 3%) or both polycultures (21 ± 6%). We found a similar result with crop rotations. The yield gap was smaller when the organic system had more rotations (8 ± 5%) compared with when both treatments had a similar number of rotations (20 ± 2%) or did not have crop rotations at all (16 ± 5%). These results also suggest that polyculture and crop rotations increase yields in both organic and conventional cropping systems (figure 2). Seufert et al. [36] found no such differences between cropping or rotation systems. There is some overlap between studies that reported the yields of organic polycultures with more rotations, so these practices could work synergistically to close the yield gap, or one of the practices could be producing the majority of the effect.

We found evidence of bias in the meta-dataset towards studies reporting higher conventional yields relative to organic (electronic supplementary material, Section S3). We also detected a trend towards larger yield gaps in more recent studies, though it is difficult to determine the causal mechanism for this trend (electronic supplementary material, Section S3). Our results should therefore be interpreted as presenting an estimate of the yield gap from the available literature that is likely favouring studies reporting higher conventional yields than organic.

4. Discussion

Our extensive dataset including over three times more yield comparisons than previous studies [35,36] and our hierarchical analytical framework, provides the most up-to-date estimate of the yield gap between organic and conventional agriculture, and how this yield gap is influenced, or not, by management practices and crop types. The lower bound of our credible interval around the yield ratio overlaps the upper bounds of the two previous meta-analyses [35,36], but because these analyses did not account for the hierarchy of their data and/or the sampling variance within studies, these prior estimates are subject to high levels of Type I error (underestimated uncertainty), which likely results in inaccuracy in estimating the yield gap and its statistical uncertainty. Further, we found entirely different effects of crop types and management practices on the yield gap than previous studies [36].

The results of our analysis are limited by modelling considerations and the studies available for inclusion. We modelled...
as many layers of non-independence in our meta-dataset as the data supported, but others may exist. In addition, we found a bias towards reporting of higher conventional to organic yield ratios in the literature; therefore, even though our estimate of the yield gap is more robust and smaller than previous analyses, it may still be an overestimate.

The estimate of the organic-to-conventional yield ratio is an average over many disparate systems and crop types. The over-representation of specific practices or crops in the dataset may therefore excessively influence the estimate of the yield gap. For example, cereal crops, which exhibit the greatest difference in yield of the crop types between organic and conventional systems, were greatly over-represented (53% of comparisons). The finding that cereal productivity (including wheat, barley, rice and maize) is lower in organic systems is of interest because of its central importance in the human diet and predominance in cultivated land area. This larger difference, however, is not surprising, given the extensive efforts since the Green Revolution to increase cereal yields by breeding high-yielding cereal varieties adapted to work well with conventional inputs [50,51].

Given that there is such a diversity of management practices used in both organic and conventional farming, a broad-scale comparison of organic and conventional production may not provide the most useful insights for improving management of organic systems. Instead, it might be more productive to investigate explicitly and systematically how specific management practices (e.g. intercrop combinations, crop rotation sequences, composting, biological control, etc.) could be altered in different cropping systems to mitigate yield gaps between organic and conventional production. Historically, research and development of organic cropping systems has been extensively underfunded relative to conventional systems [16,52,53]; thus, research priorities would need to shift to provide for this needed work. Our meta-analyses found relatively small, and potentially overestimated, differences in yield between organic and conventional agriculture (i.e. between 15.5 and 22.9%), despite historically low rates of investment in organic cropping systems. These yield differences dropped to 9 ± 4% and 8 ± 5% when diversification techniques (multi-cropping and crop rotations, respectively) were used. We therefore suggest that further investment in agroecological research has the potential to improve productivity of sustainable agricultural methods to equal or better conventional yields in various cropping systems, as has indeed been demonstrated through long-term studies (e.g. [54,55]).

Further, many comparisons between organic and conventional agriculture use modern crop varieties selected for their ability to produce under high-input (conventional) systems. Such varieties are known to lack important traits needed for productivity in low-input systems, potentially biasing towards finding lower yields in organic versus conventional comparisons. By contrast, few modern varieties have yet been developed to produce high yields under organic conditions [50,56]; generating such breeds would be an important first step towards reducing yield gaps when they occur. Finally, reducing the yield gap between organic and conventional agriculture (or, more accurately, between biologically diversified versus chemically intensive farming systems) has the potential benefit of reducing the loss of biodiversity and ecosystem services degrade? BioScience 60, 576 – 589. (doi:10.1525/bio.2010.60.8.4)


forests were the primary sources of new agricultural land in the 1980s and 1990s. Proc. Natl. Acad. Sci. USA 107, 16 732 – 16 737. (doi:10.1073/pnas.091027/5107)


14. FAO. 2009 How to feed the world in 2050. Rome, Italy: Food and Agriculture Organization.


55. Rodale Institute. 2012 Farming systems trial, 30-year report. Technical report, Rodale Institute, Kutztown, PA, USA.


