LETTER

Combining randomized field experiments with observational satellite data to assess the benefits of crop rotations on yields

Dan M Kluger1∗, Art B Owen1 and David B Lobell2

1 Department of Statistics, Stanford University, Stanford, CA, 94305, United States of America
2 Department of Earth System Science and Center on Food Security and the Environment, Stanford University, Stanford, CA, 94305, United States of America

∗ Author to whom any correspondence should be addressed.

E-mail: kluger@stanford.edu

Keywords: causal forest, data fusion, observational data, crop rotation, US Corn Belt, causal inference, crop yields

Supplementary material for this article is available online

Abstract

With climate change threatening agricultural productivity and global food demand increasing, it is important to better understand which farm management practices will maximize crop yields in various climatic conditions. To assess the effectiveness of agricultural practices, researchers often turn to randomized field experiments, which are reliable for identifying causal effects but are often limited in scope and therefore lack external validity. Recently, researchers have also leveraged large observational datasets from satellites and other sources, which can lead to conclusions biased by confounding variables or systematic measurement errors. Because experimental and observational datasets have complementary strengths, in this paper we propose a method that uses a combination of experimental and observational data in the same analysis. As a case study, we focus on the causal effect of crop rotation on corn (maize) and soybean yields in the Midwestern United States. We find that, in terms of root mean squared error, our hybrid method performs 13% better than using experimental data alone and 26% better than using the observational data alone in the task of predicting the effect of rotation on corn yield at held-out experimental sites. Further, the causal estimates based on our method suggest that benefits of crop rotations on corn yield are lower in years and locations with high temperatures whereas the benefits of crop rotations on soybean yield are higher in years and locations with high temperatures. In particular, we estimated that the benefit of rotation on corn yields (and soybean yields) was 0.85 t ha−1 (0.24 t ha−1) on average for the top quintile of temperatures, 1.03 t ha−1 (0.21 t ha−1) on average for the whole dataset, and 1.19 t ha−1 (0.16 t ha−1) on average for the bottom quintile of temperatures. This association between temperatures and rotation benefits is consistent with the hypothesis that the benefit of the corn-soybean rotation on soybean yield is largely driven by pest pressure reductions while the benefit of the corn-soybean rotation on corn yields is largely driven by nitrogen availability.

1. Introduction

The task of understanding which farm management practices lead to increased crop yields has been important for millennia. This task is especially critical in a time with increasing rates of food insecurity [1] and with climate change threatening agricultural productivity [2]. Increasing crop yields with better farm management practices can mitigate food insecurity issues and can substantially decrease the amount of land used for agriculture—a sector which currently uses about 37% of the land on Earth [3]. Yet identifying which practices are truly ‘better’ for a given location and climatic regime can be very difficult given the typically strong interactions of management with soil and weather conditions.

Historically, researchers have turned to randomized experiments on designated research croplands to answer questions about whether a particular agricultural practice leads to higher crop yields. Such randomized experiments are the gold standard in causal inference because they can be used to
get an unbiased estimate of the causal effect of a farm management practice on yield. On the other hand, randomized field trials often suffer from small sample sizes leading to wider confidence intervals than desired for these causal effects. In addition, of even greater concern, is that randomized experiments can suffer from external validity issues [4], meaning that their conclusions may not apply to farms with different soil, climate, or management conditions than those of the experimental site. A related limitation is that randomized experiments can only investigate a few management factors at once (if the goal is to estimate the effect of a particular agricultural practice on yield, this is an external validity issue). Recent open science initiatives that aggregate the data from many different field experiments [5–7] mitigate this limitation, and moreover, they can reduce small sample size and external validity issues. However, even with aggregated datasets that include a variety of field experiments, small sample size and external validity issues could remain, as the aggregated dataset may only have a few experiments where the management practice of interest is randomized.

One approach to studying the effectiveness of farm management practices that does not suffer from external validity issues, but could suffer from accuracy issues, is to use cropping systems simulations. Crop system simulators, such as the Agricultural Productions Systems Simulator [8] and the Decision Support System for Agrotechnology Transfer [9] are well tested [10, 11] and have been used to study management practices such as tillage [12], irrigation [13], optimal fertilizer application [14], and the benefits of crop rotation [14, 15]. Studies using crop simulations to investigate management practices typically leverage experimental data to either validate or calibrate the crop simulations, and they subsequently use calibrated simulations to draw conclusions about the management practice of interest in settings beyond the experimental sites. While crop simulators provide a helpful approach to studying farm management practices, they inevitably rely on simplifying assumptions and typically only consider a subset of the physical, chemical, and biological processes involved (e.g. they typically omit pests and weeds). The focus in this manuscript is on data-centric approaches.

Thanks to the recent big data revolution coupled with advances in satellite-based remote sensing technologies, the effects of agricultural practices on yields are now being investigated with much larger datasets that span a wide array of growing conditions [16–18]. Despite the appeals of using large satellite-based datasets for drawing causal inferences, they suffer from two main drawbacks. First, satellites can give inaccurate estimates of treatment variables such as crop rotation or tillage as well as inaccurate estimates of the yield, although this issue will be of decreasing importance as the technology and algorithms continue to improve. Second, such datasets are observational and lack randomized treatments, so causal inferences drawn from such datasets would rely heavily on the assumption that there are no unmeasured founders. If strong assumptions fail to hold, an observational study could fail to satisfy internal validity, meaning that the estimates for the causal effect based on the observational study are biased and do not properly estimate the causal effect in the study sample.

The advantages and drawbacks of experimental and observational datasets are complimentary, suggesting benefits for combining the two types of datasets in the same analysis. In particular, it is plausible to leverage the high internal validity of experimental data to improve upon causal estimates from observational studies which could have large biases due to systematic measurement errors or unmeasured founders. Conversely, it is plausible to leverage the large sample size and high external validity of observational data to improve upon causal estimates from experimental data, which could have high variance due to small sample sizes and may not be representative of the fields and growing conditions of interest. Statistical methods for fusing both experimental and observational data is a growing area of research [19], and a number of recent theoretical results from the field of statistics suggest that when both experimental and observational data are available, it is strictly better to use methods that leverage both the experimental and observational data in tandem rather than to only use the experimental data [20, 21] or to only use the observational data [22]. To our knowledge, these hybrid methods have not previously been used in agronomy or the earth sciences.

The goal in this paper is to develop an approach for combining data from randomized field experiments with observational satellite-based datasets to estimate the causal effect of agricultural practices on yields. As a case study, our focus is on estimating the causal effect of crop rotation on corn and soybean yields in the Midwestern United States.

Crop rotation has long been known to be an important component of agricultural systems and its benefits on crop yields are well established via randomized experiments. For example, for the predominant 2 year corn-soybean rotation in the United States, the median benefit of rotation on corn yields was 0.87 tons per hectare (t ha$^{-1}$) and ranged from −0.32 to 1.68 t ha$^{-1}$ (5%–95% quantiles) across 28 field experiments conducted prior to 2008 [23]. Meanwhile, the estimated benefit of the corn-soybean rotation on soybean yields ranged from 0.15 to 0.48 t ha$^{-1}$ (median = 0.26 t ha$^{-1}$) across seven different experiments in the United States conducted prior to 2014 [24–27].

The yield benefit of rotation is commonly explained by reductions in pest or disease pressure, as well as by increased nutrient availability and soil quality. Less is understood about how the
effectiveness of crop rotation varies across geographical regions and weather conditions and about how the effectiveness of crop rotation might change under the extreme temperature and precipitation conditions that are forecasted in a changing climate. To this end, interactions between crop rotation benefits and weather or soil covariates have been explored using solely experimental data [28] or solely observational data [17, 18]. The results in [28] suffer from small sample sizes and suggest that the rotation benefit on corn yields is higher at low temperatures whereas the results from [17, 18] do not use data from randomized experiments and suggest the opposite relationship between rotation benefit on corn yields and temperatures.

In this paper, we introduce a calibration-based approach for combining both experimental and observational data in the same analysis. Our approach is similar to that in [29], although their approach is used for calibration of causal estimates within large experimental datasets. We use leave-one-out cross validation to check that our approach is helpful for predicting the treatment effects in new locations. Finally, we explore interactions between our estimated crop rotation benefits and weather covariates and discuss our findings.

2. Methods
In this section, we summarize the dataset and methods used. We refer readers to the Materials and Methods section in the supplementary file (available online at stacks.iop.org/ERL/17/044066/mmedia) for a more detailed description of the data and methods used and for an explanation of some concepts and choices in the analysis.

2.1. Datasets
Our study region was the Corn Belt of the United States, which is characterized by high-yielding commercial agriculture for corn and soybeans and contributes nearly 30% of the global production for these crops [30]. We focus on a nine-state region within the Corn Belt (figure 1) spanning the 19 year period from 2000 to 2018 due to availability of yield data.

Our observational satellite-dataset consisted of 180 000 randomly sampled 30 \times 30 m agricultural pixels from the nine-state study region. For each pixel and for each year in the study span, the dataset contained information on yield, crop rotation, as well as weather and soil covariates. Our yield data was taken from previously published yield maps for corn [31] and soybean [32] that were produced using the Scalable Crop Yield Mapper [31–34], a method that estimates yields using a combination of satellite data and crop modelling simulations. For each year and pixel in our dataset, crop type was determined using satellite-based crop maps [35, 36]. Crop rotation within each 2 year window was inferred from this data and was categorized into five types of observed rotations: corn planted before corn (CC), soybean planted before corn (SC), soybean planted before soybean (SS), corn planted before soybean (CS), and rotations involving other crops. Our analysis ignores rotations involving other crops, because the specific crop type among the other crops could not be determined with accuracy that was close to that of the corn and soybean classifications and because yield maps were not available for the other crops. Weather and soil covariates were extracted from a combination of sources [37–40], and the pre-selected covariates are summarized in table 1.

Our experimental dataset came from two open-source datasets [5, 6], each containing the yield data from multiple different randomized field experiments in North America. For assessing the benefit of rotation on corn yields we selected experiments that contained instances of both SC and CC, and for assessing the benefit of rotation on soybean yields we selected experiments that contained instances of both CS and SS (see section M.2.2 for our other inclusion criteria). Our final dataset consisted of 11 experiments, each lasting at least three years, and combined they included 89 site-years of yield data and spanned the years 2000–2016. All experiments in our final dataset were completely randomized block designs with 2–5 replications per site [42–46]. In figure 1, we see that the 11 experiments in our final dataset were spread out throughout the Corn Belt and were conducted in a variety of climatic conditions. Despite this variety, our dataset contained no experiments from certain states, only one experiment in a climate with low early season precipitation, and only one experiment in a climate with a large number of extreme degree days.

2.2. Estimation of the rotation benefits (the causal effects of rotation on yield)
Throughout the text, we will use the term ‘rotation benefit’ to denote the causal effect of crop rotation on yield (or equivalently, the term denotes the treatment effect for a setting where the treatment variable is crop rotation and the outcome variable is yield). In this subsection, we describe three approaches for estimating the rotation benefits. One uses only the satellite-based dataset, the second uses only the experimental dataset, and the third is our proposed hybrid approach which uses both the experimental and satellite-based dataset.

To estimate the benefit of crop rotation on yield using only satellite-based observations, we fit causal forests [47, 48] using the grf R package [49]. The causal forest method is a recent adaptation of the classical random forest algorithm [50] designed for estimation and inference of causal effects in settings where it is of interest either to understand how the
Figure 1. Maps of average early season precipitation and extreme temperatures in the Corn Belt. In these maps the colored regions indicate the geographical span of the satellite-based dataset, and the text indicates the location of the 11 experiments. The map on the left shows the variation in geography of early season precipitation (1 January–30 April) averaged across the years 2000–2018, while the map on the right shows the geographical variation in extreme heat, measured in terms of extreme degree days averaged across the years 2000–2018.

Table 1. Pre-selected weather and soil covariates and their descriptions. Each row corresponds to an element of the covariate vector $x$.

| Covariate name       | Description                                                                 | Measurement frequency | Source          |
|----------------------|-----------------------------------------------------------------------------|-----------------------|-----------------|
| Latitude             | Latitude of $30 \times 30$ m pixel                                          | Once                 |                 |
| Longitude            | Longitude of $30 \times 30$ m pixel                                         | Once                 |                 |
| Year                 | Year in which measurements were taken                                       | Annually             | PRISM [37,38]   |
| Growing degree days  | Aggregated temperature exceeding $8 \degree C$ (April–September)            | Annually             | PRISM [37,38]   |
| Extreme degree days  | Aggregated temperature exceeding $30 \degree C$ (April–September)          | Annually             | PRISM [37,38]   |
| Early season precipitation | Rainfall between 1 January and 30 April                                          | Annually             | GRIDMET [40]    |
| Growing season         | Rainfall between 1 May and 15 September                                      | Annually             | GRIDMET [40]    |
| Previous year early season precipitation | Early season precipitation from the previous year                                           | Annually             | GRIDMET [40]    |
| Previous year growing season precipitation | Growing season precipitation from the previous year                                           | Annually             | GRIDMET [40]    |
| rootznaws             | Root zone available water storage                                            | Once                 | SSURGO [39]     |
| aws0_100              | Top 1 meter available water storage                                          | Once                 | SSURGO [39]     |
| Corn productivity index (NCCPI [41]) | Score of how favorable the soil is for growing corn                                           | Once                 | SSURGO [39]     |
| Soybean productivity index (NCCPI [41]) | Score of how favorable the soil is for growing soybean                                        | Once                 | SSURGO [39]     |

Causal effects vary as a function of the covariates in the model or to predict the causal effect for specific samples. For example, in our setting, a causal forest can be used to investigate how the causal effect of crop rotation on yield varies as a function of the weather and soil covariates, and it can also be used to predict the effect of crop rotation on yield for a field with known weather and soil covariates. The causal forest method has previously been used to study the effect of tillage on crop yields [16], the impact of weather on agricultural productivity [51], and the effectiveness of forest management policies [52], fishery policies [53], and growth mindset interventions [54].

Letting $x$ denote the vector of our covariates (see table 1), our first causal forest learned a function which we call $ECRB_{Sat}(x)$, denoting the satellite-based estimated corn rotation benefit. In particular, $ECRB_{Sat}(x)$ is an estimate of the average second-year corn yield in a soybean-corn rotation minus that in a corn-corn rotation, stratified by the specific year, geolocation, soil covariates and weather covariates encoded in $x$. Analogously, we fit a separate causal forest to learn a function $ESRB_{Sat}(x)$, the satellite-based estimated soybean rotation benefit. It should be emphasized that $ECRB_{Sat}(x)$ and $ESRB_{Sat}(x)$ may not give reliable estimates of the causal effect of rotation on corn and soy, because they can be biased due to unmeasured confounders (such as fertilizer) and because of errors in the treatment and outcome variables (the rotation and yield variables in this observational dataset were estimated using satellite imagery). See section M.3.1 for more details and discussion about the quantities $ECRB_{Sat}(x)$ and $ESRB_{Sat}(x)$.

In the second approach, rotation benefits were estimated using only the experiments. In each site $s$, for each year $t$, we estimated the effect of rotation on corn yield by taking the difference of means between SC subplots and CC subplots:
ECRB_{Exp}(s, t) = \text{average yield across SC subplots} - \text{average yields across CC subplots.}

Similarly, for each site \( s \) and year \( t \), we estimated the experimental benefit of rotation on soybean yields with

ESRB_{Exp}(s, t) = \text{average yield across CS subplots} - \text{average yield across SS subplots.}

Because most sites did not have yield data for a continuous soybean rotation, we were only able to calculate ECRB_{Exp}(s, t) at three out of the 11 experimental sites in our study. For details on inclusion criterion for the subplots, see section M.3.2.

For our hybrid approach, we leverage both the experimental dataset and the satellite-based observational dataset by fitting a linear calibration of the ECRB_{Sat} values at the experimental sites towards the ECRB_{Exp} values. In particular, at each experimental site \( s \) and year \( t \), we used the precise latitudes and longitudes of the experimental site to determine the covariate vector \( x(s, t) \) and compute ECRB_{Sat}(x(s, t)) using the previously fitted causal forest. Using all the site and year combinations that had both ECRB_{Sat}(x(s, t)) and ECRB_{Exp}(s, t) estimates, we fit the following mixed effects regression model

\[
ECRB_{Exp}(s, t) = a + b \times ECRB_{Sat}(x(s, t)) + \alpha_s + \beta_t + \varepsilon_{s,t},
\]

where \( \alpha_s \) is a random effect for site, \( \beta_t \) is a random effect for year, and \( \varepsilon_{s,t} \) is mean zero, independent Gaussian noise with variance that depends on the number of SC and CC subplots at each site. After fitting the mixed effects model, we used the estimated intercept \( \hat{a} \) and slope \( \hat{b} \) to construct a calibrated estimate of the effect of crop rotation on corn yield for any choice of covariate vector \( x \):

\[
ECRB_{Calib}(x) = \hat{a} + \hat{b} \times ECRB_{Sat}(x).
\]

Because we only had three distinct experimental sites with continuous soybean yield data, to calculate calibrated estimates of the soybean rotation benefit, ESRB_{Calib}(x), we fit a more parsimoniuous calibration model, producing calibrations of the form ESRB_{Calib}(x) = \hat{c} + ESRB_{Sat}(x). See section M.3.3 for more details on the calibration and discussion about the modelling choices. Variants of the above approaches for estimating the rotation benefit that were attempted but not ultimately included in this manuscript are acknowledged in section M.5.

2.3. Leave-one-out cross validation of rotation benefit estimates

To estimate the errors associated with our calibration approach in predicting the experimental rotation effect on corn yield at an unobserved experimental site, we used a variant of leave-one-out cross validation (see section M.4 for a description of the leave-one-out cross validation approach used and the alternative prediction approaches that were used for a comparison). The leave-one-out cross validation was only performed for corn rotation benefits and not for soybean rotation benefits because only three experimental sites had available yield data from both CS and SS subplots.

3. Results and discussion

3.1. Satellite-based estimates, experiment-based estimates and their association

The satellite-based and experimental estimates for the benefit of crop rotation are visualized in Figures S2 and S3 for corn and soy, respectively. As expected, we generally see a positive effect of rotation on both corn and soybean yield, although the observational corn rotation benefit, ECRB_{Sat}, is negative in some regions. We suspect that in many of the regions where ECRB_{Sat} is negative, rotation still has a positive effect on corn yields, and the apparent harm of rotation is due to bias in the uncalibrated satellite-based estimates (although, there could be some cases where the causal effect of rotation on corn yield is truly negative).

To directly visualize how correlated the satellite-based estimates are with the experiment-based estimates, we create a scatterplot of ECRB_{Sat}(x(s, t)) versus ECRB_{Exp}(s, t) (figure 2, top) and ESRB_{Sat}(x(s, t)) versus ESRB_{Exp}(s, t) (figure 2, bottom). Each point in the scatter plot corresponds to one year of data at one site.

We observe a significant positive correlation between ECRB_{Exp} and ECRB_{Sat}, but with clear differences between the two estimates. In particular, the satellite-based estimates exhibit a much smaller range than the experimental estimates, which is not surprising for two reasons. First, due to inaccurate classifications of the rotation in the satellite-based dataset, we expect the satellite-based estimates to be biased towards zero by a multiplicative factor smaller than 1 (this well-established phenomena \([55–57]\) is known as attenuation bias, and a mathematical justification for this phenomena in the setting of a causal forest is presented in appendix B). Second, satellite-based estimates are estimated using a large sample and therefore are not very noisy, whereas each experimental estimate is based on yields from only a few subplots. The experimental estimates are therefore much noisier (see 95% confidence intervals on figure 2), and hence the extreme positive and negative experimental benefits can be partially explained by noise. The noise in the experimental estimates can also partially explain some unexpected observations where ECRB_{Exp} is negative and ECRB_{Sat} is positive. Despite the inconsistencies between satellite-based estimates and experimental estimates, the differing ranges of the two types of estimates and the apparent
Figure 2. Comparison between satellite-based estimates and experimental estimates of rotation benefits on corn yields (top) and soybean yields (bottom). Top: each point corresponds to one year and one experimental site. The $x$ coordinate gives the $\text{ECRB}_{\text{Sat}}$ value. The $y$ coordinate gives the value of $\text{ECRB}_{\text{Exp}}$ at a particular experimental site and year, with 95% confidence intervals for the year and site-specific averages given by the grey lines. The 95% confidence interval pooled standard error estimates across years but not across sites. Triangular points represent sites from [6] while circular points represent sites from [5]. Bottom: this is similar to the top panel except here we look at estimates for the effect of rotation on soybean yields rather than those for corn yields.

A positive association between $\text{ECRB}_{\text{Exp}}$ and $\text{ECRB}_{\text{Sat}}$ suggests the benefits of using our linear calibration approach over a satellite-only approach.

We fit the model described in section M.3.3 and estimated $\hat{a} = 0.58$ (95% CI = [0.00, 1.15]) and $\hat{b} = 1.75$ (95% CI = [0.37, 3.12]). These confidence intervals are based on the estimated standard errors in the mixed effects model, and the positive confidence interval for $b$ gives evidence that there is a positive association between $\text{ECRB}_{\text{Sat}}$ and $\text{ECRB}_{\text{Exp}}$. To confirm the statistical significance of the positive association between $\text{ECRB}_{\text{Sat}}$ and $\text{ECRB}_{\text{Exp}}$, we also ran 10,000 iterations of the cluster bootstrap, where the experiments were sampled with replacement, but the data within each experiment was not (see for example, ‘Strategy 1’ in section 3.8 of [58]). The 95% bootstrap confidence interval for the Pearson correlation between $\text{ECRB}_{\text{Sat}}$ and $\text{ECRB}_{\text{Exp}}$ was [0.03, 0.36] and the 95% bootstrap confidence interval for the linear regression slope of $\text{ECRB}_{\text{Exp}}$ on $\text{ECRB}_{\text{Sat}}$ was [0.33, 4.11].

For soybean calibration, the additive bias correction term $\hat{c}$ was equal to 0.003 t ha$^{-1}$ (95% CI = [−0.10, 0.11]). Therefore, our calibrated estimates, $\text{ESRB}_{\text{Calc}}$, and our satellite-based estimates, $\text{ESRB}_{\text{Sat}}$, are nearly identical, and some readers may prefer to interpret our plots of $\text{ESRB}_{\text{Calc}}$ as simply plots of $\text{ESRB}_{\text{Sat}}$ values.
3.2. Leave-one-out cross validation

The results for the leave-one-out cross validation indicate that the hybrid calibration approach is better at predicting the rotation benefits in an unobserved experimental site than various reasonable approaches to using only the experimental data (table 2). We also found that the hybrid calibration is better at predicting the treatment in an unobserved experimental site than using only the observational data. Overall, the hybrid approach resulted in errors that were 13% lower than using experimental data alone, and 26% lower than using the satellite data alone. The second and third rows of table 2 indicate that these findings are not sensitive to our calibration model choices of including a random effect for site and year and our choice of not including tillage in the model. In table S2, we consider an additional sensitivity check to see if our conclusions would still hold had we taken a weighted average rather than a standard average of the squared prediction errors when performing leave-one-out cross validation.

These tables suggest that the hybrid calibration approach of section M.3.3 performs the best overall in predicting the experimental rotation benefit at a held-out site (see section M.4, for explanation of the term ‘held-out site’), but it does not perform substantially better than fitting a model that uses just the experimental data and the early season precipitation weather covariate. However, whereas the hybrid calibration approach was a principled choice in the absence of a priori knowledge about which weather covariates are associated with the actual rotation benefits, the model with early season precipitation was merely found to perform well empirically, among four possible choices of weather covariates, on a dataset with a small number of different experiments.

3.3. Positivity and heterogeneity of estimated rotation benefits

Having established that the calibration leads to improved estimates of the effect of rotation on corn yields, we visualize a map of these calibrated effects in figure 3 (top). We also visualize ESRB_{Calib} in figure 3 (bottom), but we emphasize that because we have only three experiments with continuous soybean data, our calibration merely accounted for additive bias rather than both additive bias and multiplicative bias, and the approach was not validated on held-out data. The causal effects of rotation on yield were found to be geographically heterogeneous, with corn yields benefitting most from rotation in the northwest parts of the Corn Belt and soybean yields benefitting most from rotation in the southern and central parts.

We can also see from figure 3 that the calibrated rotation benefits for corn yield are generally positive. While only 91.5% of ECRB_{Sat} values are positive, 99.8% of ECRB_{Calib} values are positive (one sided 95% CI = [91%, 100%], using the same cluster bootstrap as in section 3.1), implying that the causal effect of rotation on corn yields is negative less often than observational data alone would suggest.

A key benefit of using observational data in addition to experimental data is the ability to study the rotation benefits under weather conditions that were not observed in the experimental data. In figure 4, we plot a heatmap of our calibrated estimates of the rotation benefit for soybean and corn yields as a function of temperature and precipitation, with points indicating the weather covariates observed at experimental sites.

We observe that the corn rotation benefit is smaller at high temperatures and low growing season precipitation values, but this likely could not be inferred from experimental data alone as only one site-year was observed to have temperature greater than 2400 growing degree days and growing season precipitation less than 500 mm. The lower absolute effects of rotation on corn yield in the high temperature and low precipitation regime are perhaps partly explained by these conditions leading to lower corn yields in general [59–61]. Heatmaps of ECRB_{Calib} as a function of other temperature and precipitation variables are presented in figure S4. While it appears that the benefit of rotation on corn yield is particularly high when at the lowest observed early season precipitation values, this phenomenon only occurs in a narrow window of temperature values and is potentially an artifact of most of such points being in South Dakota, where the rotation benefit is seen to be particularly high. This phenomenon may also be explained by the greater uncertainty and inaccuracy of our estimates near the boundary of covariate space, which we discuss below.

We caution readers that near the boundaries of the heatmaps (e.g. outside 1st–99th percentile range for the covariates) in figures 4, S4 and S5, the estimated rotation benefits have higher uncertainty and are likely less accurate than those well within the interior of the heatmaps for two reasons. First, at each point near the boundary of covariate space, there are fewer samples within a neighborhood of that point, so the causal forest model will return ECRB_{Sat} or ESRB_{Sat} values that have higher variance. Second, the points near the boundary of covariate space are further away from the experimental sites, so if the true relationship between the satellite-based rotation benefits and the experimental-based rotation benefits is not the linear model described in section M.3.3, the extrapolation errors of using this calibration model will be largest near the boundary.

For soybean we observe a different and clearer pattern (figure 4, bottom). When the number of growing degree days is high, the rotation benefit for soybean yields is high and the opposite is true when the number of growing degree days is low. Intriguingly, this pattern is not observed when we
Table 2. Comparison of errors associated with various methods of predicting the average experimental effects. Each entry is reported in terms of root mean squared error (t ha\(^{-1}\)) when using leave-one-out cross validation. In columns 3–8, values were calculated by fitting the model in section M.3.3 when the ECRB\(_{sat}(x(s,t))\) term is dropped and is either not replaced (column 3), or replaced by one weather covariate (columns 4–7), or replaced by all four weather covariates in a multivariate regression (column 8). In the first row, the preferred mixed effects model is used. The second row includes a sensitivity check where we drop the random effects for site and year from the calibration model and fit a linear model instead. The third row includes another sensitivity check where we add an indicator of whether each experimental subplot had tillage as a variable in the mixed effects calibration model. The entries with the lowest root mean square error in each row are in bold.

| Calibration approach (section M.3.3) | Preferred model | Linear model | Tillage model | Just predict with ECRB\(_{sat}\) | Nearest experiment | All other experiments | Early season precipitation | Growing season precipitation | GDD | EDD | All four weather covariates | Hybrid |
|-------------------------------------|-----------------|--------------|---------------|---------------------------------|-------------------|------------------------|----------------------------|--------------------------------|-----|-----|-----------------------------|--------|
| Preferred model                     | 0.99            | 0.99         | 0.99          | 0.97                            | 0.77              | 0.83                   | 0.83                       | 0.88                           |     |     | 0.83                        | 0.73   |
| Linear model                        | 0.97            | 0.75         | 0.99          | 0.97                            | 0.83              | 0.92                   | 0.87                       | 0.92                           |     |     | 0.82                        | 0.73   |
| Tillage model                       | 1.24            | 0.88         | 0.88          | 0.92                            | 0.88              | 0.96                   | 0.92                       | 0.93                           |     |     | 0.93                        | 0.84   |
Figure 3. Maps of estimated rotation benefits for corn (top) and soybean (bottom), based on calibrated estimates. The text on these maps indicates the locations of experimental sites. In the top panel, \(E_{\text{calib}}\) is averaged across the years 2000–2018 and across square latitude and longitude bins (\(\sim 10 \times 10\) km in size). In the bottom panel, \(E_{\text{calib}}\) is averaged across the years 2000–2018 and across square latitude and longitude bins (\(\sim 10 \times 10\) km in size).

Environ. Res. Lett. 17 (2022) 044066
D M Kluger et al

3.4. Discussion

We recognize several limitations of this study. First, there were only three experimental sites for assessing soybean rotation benefit in our dataset. These three

measure temperature with extreme degree days rather than growing degree days, and in fact, for a fixed value of growing degree days, the benefit of rotation on soybean yields does not increase with extreme degree days (figure S5).

As a complimentary analysis to the heatmaps in figure 4, we compute the average calibrated rotation benefits in the top and bottom quintile for temperature (GDD). The average calibrated corn rotation benefit was 0.85 versus 1.19 t ha\(^{-1}\) across observations in the top and bottom quintile of GDD values, respectively. For soybean yields, we observe the opposite pattern where the rotation benefit is larger with higher temperatures. In particular, the average calibrated soybean rotation benefit was 0.24 versus 0.16 t ha\(^{-1}\) across observations in the top and bottom quintile of GDD values, respectively. For reference, across the whole dataset, the average calibrated corn rotation benefit across the Corn Belt was 1.03 t ha\(^{-1}\) and the average calibrated soybean rotation benefit was 0.21 t ha\(^{-1}\).

We also check whether there is a clear temporal trend in rotation benefits during the 19 year study period. In figure S6, we see that there is no clear or substantial temporal trend in the rotation benefits for either corn or soybean. It appears that \(E_{\text{calib}}\) has a slight positive trend while the \(E_{\text{calib}}\) values may have a slightly negative trend throughout the duration of the study period, a phenomenon which would be explainable by decreasing temperatures in the Corn Belt between 2000 and 2019.
sites were not representative of the entire Corn Belt, as they were geographically clustered (figure 3, bottom) and did not contain any observations at lower temperatures (figure 4, bottom). Therefore, the calibration fitted to the three experiments may not extrapolate as well to western, northern, and eastern parts of the Corn Belt and to settings with lower temperatures, and future studies on soybean rotation benefit should collate a larger and more diverse set of experiments if possible. While calibration suffers in such settings with a limited collection of experimental sites, our example highlights that in such settings, observational data is especially critical for addressing external validity issues.

Second, our analysis implicitly assumed that the experimental data gave unbiased estimates of the rotation benefits, as our proposed approach involved transforming the satellite-based estimates to better match the experimental estimates. In reality, the experimental estimates could be based on a biased sample, as experimental sites that observed an unexpected negative rotation benefit may have been more likely to be excluded from the two experimental databases we used [5, 6] (e.g. [6], identified experimental sites to include in their database via a literature search, so their collection of sites could be subject to publication bias).

Third, by only looking at the crop type in 2 year windows, our analysis does not directly address the relevant question of quantifying the long-term rotation benefits in the corn-soybean rotation. Further complicating the matter, a sizable percentage of both the experimental dataset and the satellite-based dataset that we used involved farms with a corn-soybean
rotation over a long period of time, meaning that our approach is likely estimating some weighted average of the short-term and long-term rotation benefits. Future work can analyze crop rotations using more sophisticated and longer-term rotational diversity scores such as the one proposed in [62].

Fourth, we merely provide point estimates and do not provide standard errors for our calibrated estimates. The standard errors for the calibrated estimates might be quite large due to the small number of experiments, large noise in the experimental estimates, and propagation of the noise in the satellite-based estimates.

Despite these limitations, our findings about the relationship between temperature corn rotation benefits are consistent with recent results in the literature (see appendix D for a more detailed comparison with results from [46], a study which uses 11 long-term experimental sites). In addition, as we discuss next, our findings about the corn and soybean rotation benefits are consistent with current understanding of why rotation is beneficial for each crop.

Pest pressure could largely explain our findings that the benefit of crop rotation on soybean yields is larger in observations with higher temperatures (in GDD, but not in EDD). It is well established that the benefits of crop rotation for soybean yields result largely because of reductions in pest pressure. For example, [63] found that the benefit of rotation on soybean yields in two Louisiana experiments was largely explained by a reduction in soybean cyst nematode and [64] found crop rotation to be effective in reducing soybean cyst nematode. Meanwhile, using pesticide application rate as a proxy for pest pressure, [65] showed that increased yearly minimum temperatures are associated with increased pest pressure, as many agricultural pests are not able to survive low winter temperatures. Therefore, in warmer conditions, we would expect there to be higher pest pressure on soy, rendering crop rotation a more beneficial management practice in increasing soybean yields because of its effectiveness in reducing pest pressure.

To explore the hypothesis that ESRB is larger in warmer conditions due to increased pest pressure, we plot pesticide use per acre versus ESRB values in figure 5. Because pesticide use on soybean crops is thought to be a proxy for pest pressure on soybeans (pesticides tend to be applied after pests are reported in the region), the figure indicates that soybean rotation benefit tends to be higher under conditions with greater pest pressure.

The mechanisms driving the rotation benefit for corn are different than those for soybean. While crop rotation also alleviates pest pressure on corn, pest pressure reduction is unlikely to be the main reason that corn benefits from crop rotation because of the prevalence of effective pest-resistant corn hybrids (for example, in 2020 Bt corn represented 82% of US corn acreage, while Bt soybeans were not yet commercially available [68]). Instead, the rotation benefit for corn is largely driven by effects on soil nitrogen. Fixation of atmospheric nitrogen in soybean benefits the subsequent corn yields, whereas corn residues can compete for nitrogen with the

Figure 5. Plot of estimated pesticide use for soybean versus average ESRBCalib values. Each point corresponds to one state and one year (only years between 2008 and 2017 are plotted). To generate this plot, soybean acreage estimates were taken from NASS quickstats [66], and the high estimates of pesticide use were taken from the USGS [67]. We did not include pesticide use data from 2018 because this data was preliminary. We also did not include pesticide use data from prior to 2008 in figure 5 because pesticide use did not appear to be associated with increased temperatures prior to 2008 (figure S7) and because no direct measurements of pest pressure could be leveraged instead.
subsequent crop. Soybean nitrogen fixation is less efficient at higher temperatures [69], while mineralization of nitrogen from corn residues occurs at a faster rate at higher temperatures [70, 71], which would tend to reduce nutrient limitations. These mechanisms would both serve to reduce the benefit of rotation for subsequent corn crops at higher temperatures.

Beyond nutrient dynamics, crop residues can also affect corn by lowering early season soil temperatures, which can inhibit early season corn growth when temperatures are low [72]. This effect is diminished in rotation because soybean produces less residues than corn. Thus, the penalty from growing two consecutive corn crops (i.e. the benefit of rotating) would be smaller at higher temperatures. We investigated the influence of crop residues on rotation benefits using previous year county-level crop yields [66] as a proxy for residues; however, the results of this additional analysis suggested that either the impacts of crop residues on rotation benefits are relatively minor compared to the impacts of temperature or that previous year county-level yields are a poor proxy for crop residues (see appendix D).

4. Conclusion

We developed an approach to combine experimental and observational datasets to assess the effect of agricultural practices on yields. The approach can be used to simultaneously mitigate external validity issues of randomized field experiments and internal validity issues of observational studies. Open-source initiatives which collated the results of many field experiments were essential to our approach. While we used a linear calibration approach, future work could explore more sophisticated approaches in settings with many experimental sites.

Focusing on the causal effect of crop rotation on yields in the Midwestern United States as a case study, our hybrid approach led to better predictions of the causal effect at held-out experimental sites than did the standard approaches of using only experimental data or using only observational data (tables 2 and S2). Furthermore, the case study highlighted some of the issues with observational data and experimental data. We found that using only the observational data led to a biased underestimate of the benefit of crop rotation on corn yields. This bias is partially explained by fertilizer being an unmeasured confounder because rotating fields tend to use less fertilizer. We also see that the experimental data alone do not adequately capture the geographical and weather scenarios of interest (figures 1, 3, and 4, S4 and S5).

Our results suggest that the benefit of rotation on corn yield is smaller in warmer conditions whereas the benefit of rotation on soybean yield is larger in warmer conditions. The former observation could be partly explained by the relationship between temperature and nitrogen availability, while the latter observation could be partly explained by increased pest pressure in warmer conditions. As we anticipate a warming climate, these findings have implications for how strongly crop rotations should be encouraged in the future. Specifically, our findings suggest that rotation may become a less beneficial management practice for corn yields in a future with higher temperatures, though rotation will remain a net beneficial practice. Meanwhile, our findings suggest that rotation will become an increasingly important management practice for maximizing soybean yields, and therefore, soybean farmers should be increasingly encouraged to rotate their crops.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

D M K was supported by the James and Nancy Kelso Stanford Interdisciplinary Graduate Fellowship and by a Stanford Graduate Fellowship. A B O was supported by NSF Grant IIS-1837931. D B L was supported by the NASA Harvest Consortium (NASA Applied Sciences Grant No. 80NSSC17K0652, Sub-Award 54308-Z6059203).

The authors gratefully acknowledge Jill Deines, Dominik Rothenhäusler, Michael Sklar, Matthieu Stigler, Paul Switzer, Sherrie Wang and three anonymous reviewers for helpful discussions or comment. The authors would also like to thank Lori Abendroth and Timothy Bowles, for providing relevant information about the experimental datasets that were used. Finally, the authors thank Jill Deines for assistance with extracting the observational data from Google Earth Engine.

ORCID iDs

Dan M Kluger © https://orcid.org/0000-0002-7865-259X
Art B Owen © https://orcid.org/0000-0001-5860-3945
David B Lobell © https://orcid.org/0000-0002-5969-3476

References

[1] FAO, IFAD, UNICEF, WFP and WHO 2021 The state of food security and nutrition in the world 2021. Transforming food systems for food security, improved nutrition and affordable healthy diets for all (Rome: FAO)
[2] Porter J R 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 II to the Fifth Assessment
