ENVIRONMENTAL RESEARCH
LETTERS

Agricultural yield geographies in the United States

To cite this article before publication: Emily K. Burchfield *et al* 2021 *Environ. Res. Lett.* in press https://doi.org/10.1088/1748-9326/abe88d

Agricultural yield geographies in the United States

Emily K. Burchfield[a,*] and Katherine S. Nelson[b]
*Corresponding author

a. Assistant Professor, Department of Environmental Sciences, Emory University, 400 Dowman Drive, Atlanta GA 30322, USA. Phone: (404)727-0463, Email: emily.burchfield@emory.edu, OrcidID: 0000-0003-0459-6270

b. Assistant Professor, Department of Geography and Geospatial Sciences, Kansas State University, 920 N. 17th Street, Manhattan, KS, USA. Phone: (785)532-6727, Email: ksnelson@ksu.edu, OrcidID: 0000-0002-4240-5474

Abstract: We examine the geographies of agricultural yields in the United States, home to some of the most productive agricultural systems on the planet. We model and map yield divergence from biophysical expectations and regional norms for five major crops—corn, soy, wheat, alfalfa, and hay—and assess how this divergence interacts with farm-level resources, farm(er) characteristics, and landscape context. Our results highlight the ways in which human activity has reinforced and intensified the yield geographies defined by sun, soil, and water alone. Yield gains brought by human activity are strongly associated with increased expenditure on inputs to production and receipts from federal programs, but *not* with net revenue gains for farmers. These yield gains vary across operator race, gender, farm size, and major U.S. region. We also find that beyond a threshold, increased input expenditure is associated with marginally decreasing yields, raising important questions about the interactions between yields and farmer livelihoods. We conclude by discussing the importance of broadening the production-centric paradigm that has dominated agricultural innovation over the last century to include the well-being of the farmers and ecological systems on which agricultural production ultimately depends.

Significance Statement: For the most widely grown U.S. crops, yield divergence from biophysical expectations is strongly associated with increased expenditure on inputs to production and receipts participation in federal programs, but *not* with net revenue gains for farmers, raising important questions about the interactions between yields and farmer livelihoods. These yield divergences vary across operator race, gender, farm size, and major U.S. region.

Acknowledgements: This work is supported in part by United States Department of Agriculture (USDA) National Institute of Food and Agriculture Grant No. 2020-67019-31157. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of the U.S. Department of Agriculture. Many thanks to Britta Schumacher for invaluable help in wrangling the USDA Census data into submission. Thanks to Kaitlyn Spangler and Britta Schumacher for providing thoughtful and constructive feedback on the manuscript. Thanks to Dr. Jeremy Cowan for his valuable insights on soil dynamics.

Author Contributions: Both authors contributed equally to the conceptualization and writing of the paper. Dr. Burchfield led the data construction and analyses.

Keywords: Agricultural production, yield, United States

Agricultural yield geographies in the United States

1. Introduction

Human innovation has pushed agricultural production well beyond the bounds of what sun, soil, and water alone can support. Over the last 70 years, the U.S. yields of soy and winter wheat have more than *tripled*, while corn yields have increased over *five-fold* [1]. These yield gains have supported a more than tripling of global population and brought tremendous changes to agricultural systems worldwide. Today, the U.S. is home to some of the most productive, consolidated, and specialized agricultural systems on the planet. These systems have been strongly shaped by an economic paradigm that prioritizes production and efficiency, technological innovations that target simplified large-scale systems[2,3], federal policies that incentivize specialized commodity production[4,5], and by the increased integration of rural economies in globalized markets[6,7]. Today, over half of U.S. land is devoted to agricultural production, with over 80% of this land cultivated with corn, wheat, soy, alfalfa, or hay [8]. An increasing share of these crops are grown for export, with U.S. farmers exporting, on average, more than 20% of what they produce[1]. Moreover, the operations on which these crops are grown are increasing in size, with large-scale operations earning more than $500,000 annually comprising only 7.5% of U.S. farms but operating over 41% of agricultural land [1,9].

While there is evidence that large-scale simplified systems do, in fact, generate higher yields[10–12], there is growing concern about the implications of these trends for people and planet. U.S. farm debt has increased by over 30% in the last decade to levels not seen since the 1980s Farm Crisis [13]. Over the same period, farm income has sharply declined, with the USDA reporting a median farm income of *negative* $1,735 in 2018 [14]. Many families have left the agricultural sector, as a national trend of farm consolidation has pushed small- and medium-sized farms out of production[2,15,16]. Existing farms are increasingly exposed to fluctuations in global markets, often with devastating consequences [17–19]. At the same time, changing climate[20–24] and declining environmental quality[25–27] threaten agricultural collapse in many of the same regions.

With these trends in mind, scientists, practitioners, and policy makers have started to question the production-centric paradigm that has dominated agricultural innovation over the last half-century, questioning its viability for the farmers and ecological systems on which agricultural production ultimately depends [28–30]. If one of the most productive agricultural systems on the planet is also associated with farmer debt, farm consolidation, declining incomes, and a slew of negative environmental externalities, is production really the best metric of agricultural "success"?

We begin to explore this question by identifying the farm-level resources, farm(er) characteristics, and landscape context of the most productive agricultural systems in the U.S. We identify unexpectedly high yielding agricultural regions by estimating yield divergence from what would be expected given only biophysical conditions and regional norms for the five major crops that comprise nearly 80% of cultivated acreage in the U.S.: corn, soy, wheat, hay, and alfalfa. By removing yield variance explained by biophysical conditions and regional norms, we isolate the *human yield signal,* or the human contribution to yield geographies. We then assess the farm-level resources, farm(er) characteristics, and landscape context of counties the for which the human yield signal is particularly strong.

We find that human activity has intensified and amplified the yield geographies explained by sun, soil, and water alone. In areas with unexpectedly high yields for these five crops, farmers spend more on major inputs to production (fertilizer, chemicals, machinery, and labor) and report higher rates of federal support, both in terms of government receipts and crop insurance. Surprisingly, we find that, on average, farmers in these unexpectedly high yielding systems *do not* earn higher net revenues. We also find that, beyond a threshold, increased input expenditure is associated with marginally decreasing yields. Taken together, this work suggests that while U.S. agriculture excels at meeting human demand, it falls short in satisfying other important objectives. This raises serious questions about the future viability of U.S. agriculture and highlights the importance of expanding the production-centric paradigm by reconceptualizing agriculture as a major force for positive, regenerative change for both people and planet.

2. Methods

We constructed panel datasets merging county-level yield estimates from the USDA NASS Survey with seasonal sun, soil, and water characteristics for U.S. counties cultivating alfalfa, corn, hay, soy, and (winter) wheat from 2008 to 2018, excluding counties with fewer than three yield observations over this period. Together, these crops dominate U.S. agriculture, covering 78% of agricultural land in the U.S. (Figure S10) [8] and representing nearly 60% of crop value produced in the U.S.[32] We constructed two county-season weather indicators—growing degree days and total precipitation—from gridded daily four-kilometer temperature and precipitation data provided by the PRISM Climate Group (2004) [33]. For each indicator, we defined season duration using the spatially explicit, crop-specific planting and harvesting dates provided by [34]. Growing degree days—an indicator of cumulative temperature exposure—are the sum of mean daily temperatures above a crop-specific threshold over the crop's growing season,[1] while total precipitation is the sum of precipitation (in millimeters) throughout the growing season. We also collected data describing the percent of a county's agricultural land irrigated from the USDA NASS [1]. When this data was unavailable, we replaced missing values with linearly interpolated estimates from the MiRAD project[37] standardized by agricultural extent estimates derived from the USDA's Cropland Data Layer. We constructed county-level indicators of soil characteristics by averaging at the county scale four variables in the Harmonized World Soil Database [38]: topsoil pH, topsoil organic carbon, topsoil cation exchange capacity, and topsoil exchangeable sodium percentage. These predictors were selected from the full HWSD dataset based on their known importance to agricultural production and their contributions to the prediction of yield in a random forests model (Figure S9).

To capture well-established non-linearities in the effects of seasonal precipitation and temperature on yields [39,40], we modeled the interactions between yields and county-year seasonal weather predictors using a second-order random walk function $f(X_{tc})$. This structure allows the effect of these predictors to vary non-linearly while also accounting for temporal autocorrelation in predictors effects. To avoid overfitting, we modeled soil parameters and percent irrigated land as county-level linear controls. We also included a dummy indicator for year ($TIME_t$) to capture any dynamics that affect all counties across the country in a particular year such as major market or policy changes. Any spatial variability in the impact of these shifts is captured in the county-level random effects described below.

One of the challenges of modeling county-level agricultural production is strong spatial dependency in the data, i.e. neighboring counties tend to have similar yields. We address this issue by modeling county-level random effects using a Besag-York-Mollié model [41] that includes both county-specific spatially *unstructured* random effects (modeled as exchangeable) and county-specific spatially *structured* effects modeled using a spatial weights matrix linking counties that share borders. The *unstructured* random effects (v_{0cr}) capture time-invariant factors associated with a county that influence yield (individual spice), while the *spatially-structured* effects (u_{0cr}) account for the fact that observations from neighboring counties exhibit higher correlation than more distant counties. Following Cinner et al. (2016), we also include regional random effects for each of the USDA's nine Farm Resource Regions (v_{00r}). These regions capture important variability in farm size, farmer

---

[1] GDD baseline temperature of 0°C for winter wheat and hay, 10°C for corn and soy and 15°C for alfalfa.[35,36]

demographics, cropping patterns, sociocultural context, land management norms, and market access across the U.S.[42] (Figure S1). By examining county effects in the context of the region in which they are nested, we can assess yields *given the norms* of a region. The final model is specified as:

$$y_{tcr} \sim N(\mu_{tcr}, \sigma^2)$$
$$\mu_{tcr} = b_{000} + v_{00r} + u_{0cr} + v_{0cr} + f(\boldsymbol{X}_{tc}) + \beta_1 \boldsymbol{Z}_{tc} + \beta_2 TIME_t$$

where t indexes time, c indexes counties, and r indexes regions. It is the *difference* between the county-level random effects ($u_{0cr} + v_{0cr}$) and the expected regional yields (v_{00r}) that serves as the basis of the remaining analyses. This difference can be thought of as the extent to which, given biophysical conditions, regional norms, and national temporal trends in yield, a county's yields have diverged from national averages over the last decade. This approach removes the share of variance explained by biophysical factors alone, allowing us to assess the characteristics of yield geographies *unexplained* by the biophysical. These residual geographies represent the influence of human activities on yield, which we refer to as the *human yield signal*.

We estimated models using the `R-INLA` package [43] which uses Integrated Nested Laplace Approximations to increase the computational efficiency of model estimation. The model uses uninformative (reduced precision) prior distributions for linear effects (Z_{tc}) and penalized complexity priors for non-linear seasonal weather predictors ($\boldsymbol{X}_{tc}$). We employed default and recommended settings for penalized complexity priors as provided by Simpson et al. (2014) [44]. Regional effects (v_{00r}) were modeled using a logGamma prior with weakly informative priors. County-level random effects were modeled with a Besag-Mollie-York model which includes both iid random effects and Besag spatial effects using R-INLA default priors. Model fit was evaluated using the deviance information criterion (DIC), the conditional predictive ordinate (CPO), mean squared error (MSE) and Bayesian R-squared (R2) (Table S1). Model scripts and additional information on model diagnostics and robustness checks are available at https://github.com/####/US_production_geographies.

We collected available data describing farm resources and farm(er) characteristics from the USDA Agricultural Census. We included only variables with less than 10% missing data for the period of interest (Census years 2007, 2012, and 2017, Table S3), with the exception of operator race. 74% and 41% of counties were missing data describing operated acreage by Black and Hispanic farmers respectively; data were most likely withheld by the USDA NASS to preserve confidentiality in counties where very few minority operators cultivate. Despite this limitation, we decided to include these variables in our descriptive assessments due to their growing importance in our national conversation about diversity and sustainability. We standardized predictors (where applicable) to facilitate comparison across counties using "total operated acres" which includes agricultural land used for crops, pasture, or grazing, as well as woodlands, farm roads, and farm buildings. To model the effect of variability in land use patterns on agricultural production, we also constructed additional indicators of landscape composition and configuration from the USDA Cropland Data Layer–a 30-meter annual land use dataset based on satellite imagery and extensive ground truth data. We include two indicators of *landscape composition*—an indicator of crop diversity (Shannon Diversity Index) and a measure of the predominance of undeveloped and uncultivated landcover classes (Percent Natural Cover)—and two indicators of *landscape configuration*—an indicator of largest patch size (Mean Patch Size) and an indicator of the overall patchiness of a landscape (Edge Density). We compare values of each variable across a gradient of low (-1 sd) to high (+1 sd) unexpected yields using box and whisker plots. We also assess the contribution of each variable to the prediction of the human yield signal using a random forests analysis implemented using the `randomForests` package in `R` [45]. Random forests is well-suited to high dimensional, multicollinear data, and makes no distributional assumptions about the data on which it is trained, allowing it to handle complex nonlinear interactions among predictors. Though the variable importance estimates generated by the random forest algorithm *do not* have causal interpretations, by ranking the relative contribution of farm-level resources, farm(er) characteristics, and landscape context to the prediction of yield divergence from expectations, they provide initial insights into the human factors that may be associated with surprising yields.

3. Results

3.1. Biophysical correlates of yield

Variations in the agricultural yields of corn, soy, wheat, alfalfa, and hay are largely explained by differences in access to sun, soil, and water (Figure 1, Table S1). Seasonal temperature and precipitation have strong effects on the yields of all crops, with stronger effects for temperature than for precipitation (Figure 1A). Beyond a threshold, increasing temperature exposure is associated with significant yield declines, particularly for corn and soy, corroborating prior research highlighting the threat posed by global heating [40,46–48]. The yields of corn, wheat, and soy also decline above a seasonal precipitation threshold (Figure 1B), though hay and alfalfa show sustained yield gains across the range of seasonal precipitation. Though effect strength varies across crops, irrigation and topsoil pH have consistent positive effects on crop yields (Figure 2A). Finally, the significant effects of year on yields (Figure 2B) highlights the importance of major shifts in management, technology, markets, and policies that affect production nationally. These effects are comparable in size to seasonal weather effects.

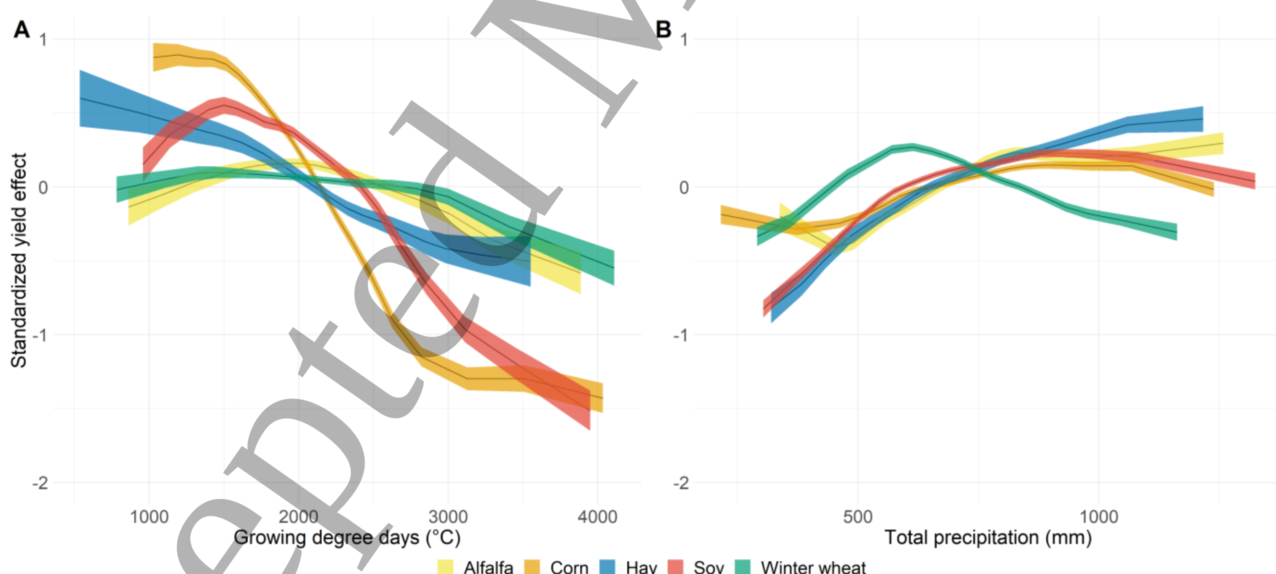

Figure 1: The crop-specific yield response to (a) seasonal growing degree days and (b) total seasonal precipitation estimated using a second-order random walk function. Solid lines show the median effect and shaded bands the 95% credibility limits. Model R^2 for each crop above 0.7 with the exception of hay (Table S1).

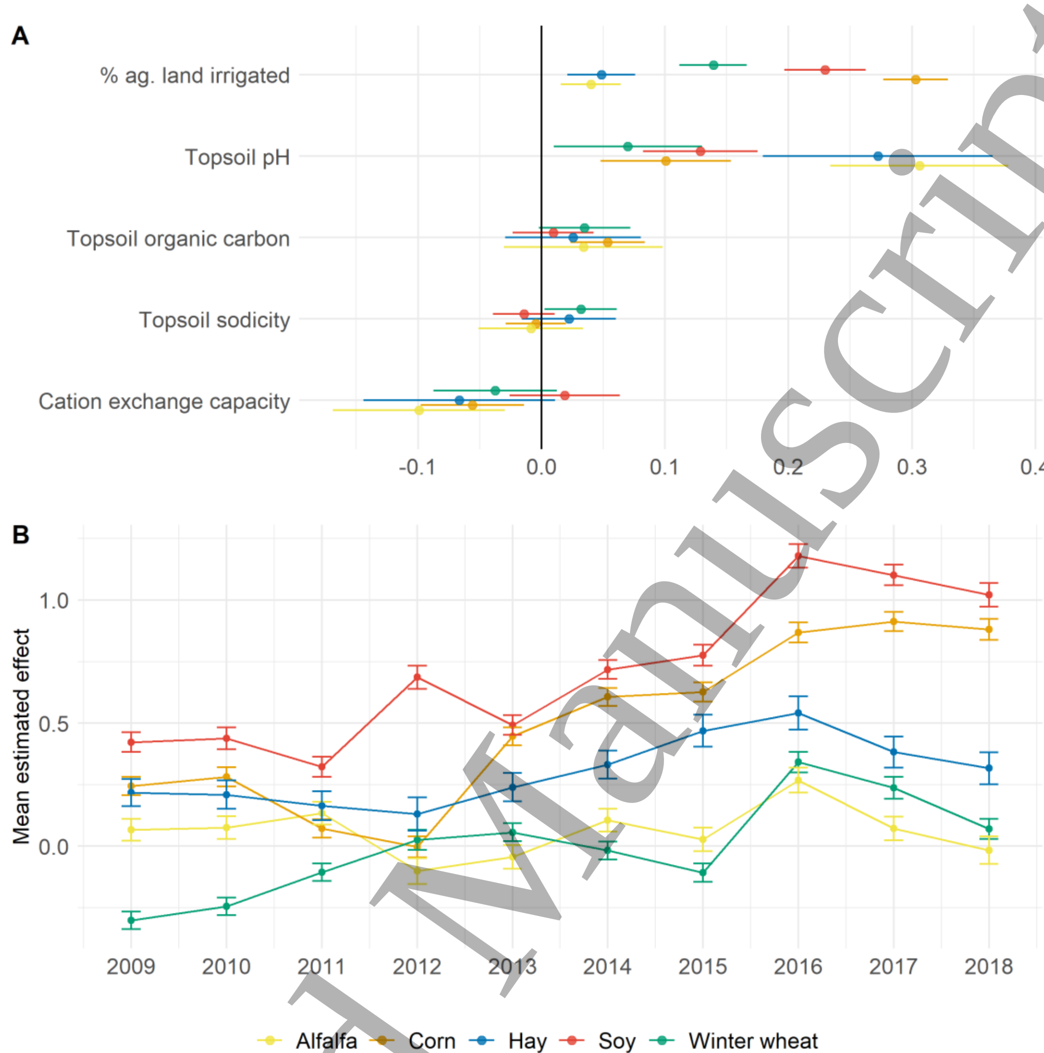

Figure 2: Regression coefficients for (a) the average effect of linear soil and irrigation predictors on yields and (b) the average effect of each year on yields. Year effects capture major shifts in management, technology, markets, and policies that affect production nationally. The points represent the mean estimate, while the error bars represent the 95% confidence interval. All continuous variables were standardized prior to running the models, so these effects represent the standard deviation change in yield associated with a one standard deviation increase in each predictor. For temporal fixed effects, the baseline year was set to 2008.

In addition to controlling for county-level variability in biophysical factors, we also include regional random effects which place county-level yield dynamics in the context of the larger region in which they are located. This allows us to compute county-level yield expectations *given* regional norms, which we define using the USDA Farm Resource Regions.[42] We observe higher regional yield norms in the Midwestern U.S. (Heartland, Mississippi Portal, Southeastern Seaboard, and Eastern Uplands) and lower yield norms in the Northern Great Plains and Prairie Gateway regions (Figure S2).

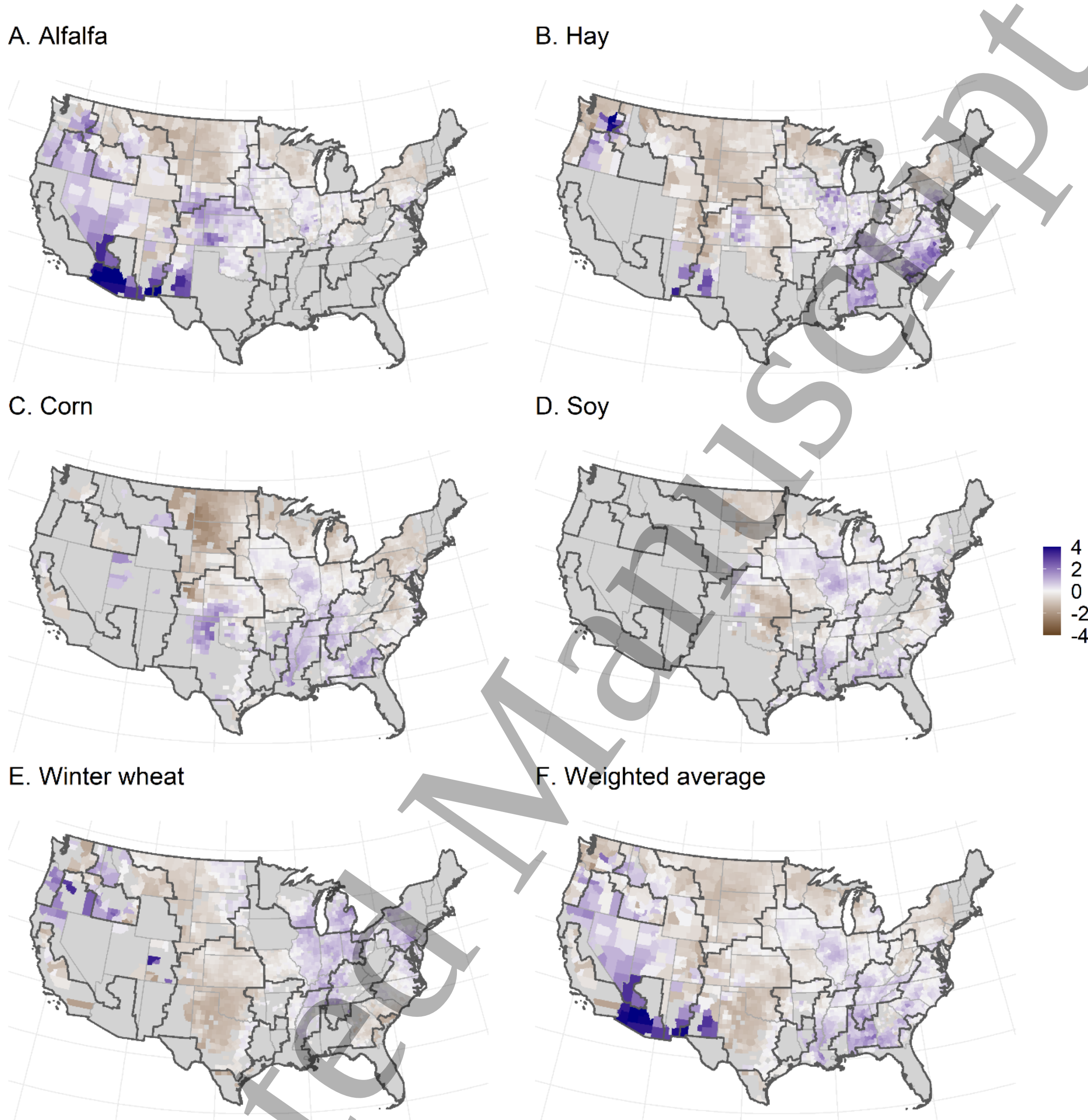

Figure 3: The human yield signal, or yield geographies *unexplained* by biophysical conditions (sun, soil, and climate), national temporal trends, and regional norms. For example, a county with a value of two has observed yields that—given county-level biophysical conditions, national yield growth, and regional norms—are two standard deviations above the national average yield. Counties for which insufficient yield data was available (fewer than three observations through time) are shaded gray. Panel F shows the average of unexpected yields across crops, weighted by the proportion of a county's agricultural land cultivated with each crop and divided by the total proportion of agricultural land cultivated with the five crops we assess. Farm resource region boundaries are shown in dark gray.

County-level variability in yields unexplained by biophysical conditions, national temporal trends, and regional norms exhibits a strong spatial structure (Figure 3), confirmed by the dominance of the spatially-structured component of the county-scale random effects (u_{0cr}) as compared to the unstructured component (v_{0cr}) and the presence of strong spatial autocorrelation in yield divergence from expectations (Table S2). Even after controlling for the biophysical factors and regional norms that make the Midwestern U.S. particularly well-suited to corn and soy production, the yields of these crops are significantly higher in the Midwest than expected. Corn, soy, alfalfa, and hay yields are also unexpectedly high in counties with access to the High Plains Aquifer. Alfalfa, hay, and winter wheat are particularly productive in the Pacific Northwest, though with significant intra-regional variability. In Arizona and New Mexico, alfalfa yields are up to five standard deviations above national averages, even after accounting for regional norms. The Southeastern U.S. and the region along the Mississippi River show unexpectedly high yields across all crops, while the northern Great Plains show unexpectedly low yields across crops. We also compute average unexpected yields across crops, weighted by the proportion of a county's agricultural land cultivated with each crop, which deemphasizes high yield estimates in counties where little of a crop is cultivated (Figure 3F). Once again, the Midwestern and Southeastern U.S. emerge as unexpectedly productive regions, as well as the Southwestern U.S., though this signal is largely dominated by highly productive and large-scale alfalfa cultivation.

3.2. The human yield signal

What factors, beyond the biophysical controls and regional norms included in our models, might help to explain unexpected yields? To begin to answer this question, we collected and constructed data describing three categories of human activity known to directly or indirectly influence yields: farm-level resources, farm and farmer characteristics, and the land use decisions that shape landscape context (Table S3). Farm-level resources include per acre expenditure on major inputs to production (machinery, labor, fertilizer, and chemicals) and per acre receipts from federal programs. These factors bound the seasonal on-farm decisions that have a strong, direct influence on production. Farm(er) characteristics such as farmer race, gender, experience, tenure, and farm size can influence access to farm-level resources as well as the adoption of new technologies and practices that influence production [49,50]. Finally, the landscape context in which a farm operates can strongly influence the provisioning of ecosystem services essential to agricultural production [51–56]. Agricultural activity has dramatically transformed landscapes in many part of the U.S.—today, over half of U.S. land is devoted to agricultural production [8]. Recent research suggests that differences in the composition—the quantity of land cover categories on a landscape—and configuration—how these land cover categories are arranged on a landscape—of the landscape in which a farm operates can have significant impacts on production outcomes [57].

To understand the major components of the human yield signal, we examine how farm-level resources, farm(er) characteristics, and landscape context vary across counties with unexpectedly high yields (counties with yields more than one standard deviation above expectations) and unexpectedly low yields (counties with yields less than one standard deviation below expectations).[2] Because these high and low yielding categories are defined by the yields of only five crops—albeit five crops that comprise nearly 80% of what's grown in the U.S. and which make up over 80% of cultivated acreage in over two-thirds of the counties included in this analysis (Figure S10)—we note that the patterns in these human factors may be influenced by other crops in a county's cropping mix. These other crops, however, reported insufficient yields to include in this national analysis. Though this data availability limits our investigation, several compelling patterns emerge.

First, in areas of unexpectedly high production for major U.S. crops, farmers spend more, on average, on major inputs to production (fertilizer, labor, machinery, chemicals) than in low

[2] We use one standard deviation to have a reasonable sample size, but note that when using two standard deviations, this strengthens the differences we describe. We include summaries of *average systems* (between -1 and +1 SD) in the SI.

production areas (Figure 4A). Median fertilizer expense in unexpectedly productive regions was $27.90 per acre as compared to $7.81 in less productive regions. Median per acre expenditure on chemicals and machinery in high production areas was $14.86 and $ 389.07 respectively, as compared to $4.31 and $165.48 per acre in low production areas. This expenditure pattern is consistent across all crops included in the analysis (Figure S3). Farmers in highly productive regions also receive over double the receipts from government programs[3] than farmers in low production areas (median values of $10.61 and $5.12 per acre respectively) and, on a per crop basis, have slightly higher rates of participation in federal crop insurance programs, particularly those cultivating soy (Figure S3F).

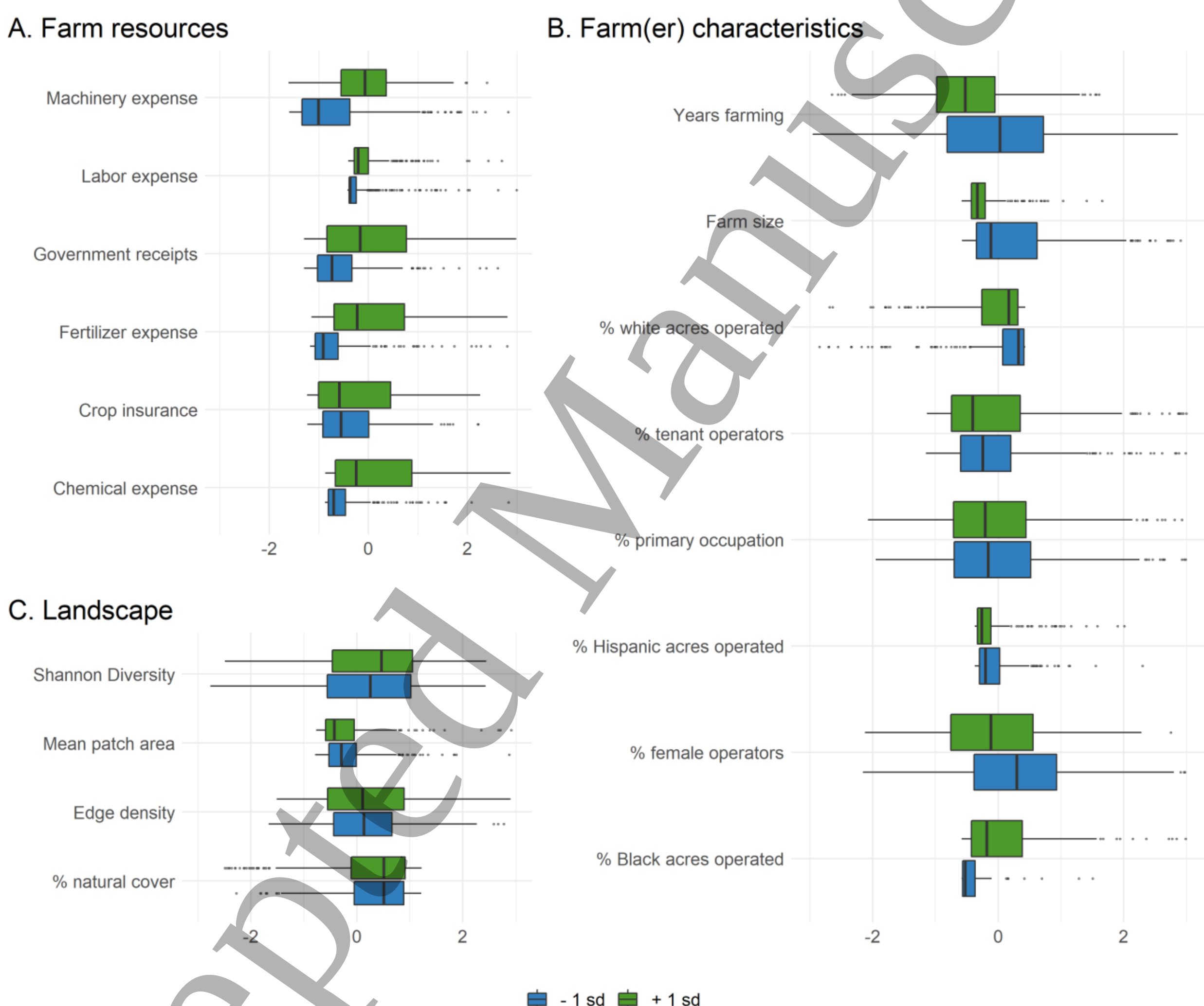

Figure 4: Box and whisker plot showing differences in (a) farm-level resources, (b) farm(er) characteristics, and (c) landscape context in counties where the across-crop weighted average of yield divergence from expectations is greater than one standard deviation above the mean (+ 1 sd) or one

[3] This category consists of direct payments from the government and includes: payments from Conservation Reserve Program, Wetlands Reserve Program, Farmable Wetlands Program, and Conservation Reserve Enhancement Program; loan deficiency payments; disaster payments; other conservation programs; and all other federal farm programs under which payments were made directly to farm operators. Commodity Credit Corporation (CCC) proceeds, local and state government agricultural program payments, and federal crop insurance payments are not tabulated in this category (USDA NASS, 2019, p. 759).

standard deviation below the mean (- 1 sd). Bold vertical lines represent the median. Values are standardized for comparison. For crop-specific results, see Figure S3.

Across the available farm and farmer characteristics we examine, the strongest consistent association we observe is for *farm size* (Figure 4B). Surprisingly, larger farms are associated with *lower* than expected yields across the five crops included in our analysis (Figure S4D). Median farm size for farmers in unexpectedly productive counties is 85 acres as compared to 172 acres in less productive counties. Farmer knowledge—measured as years of farming experience and whether the principal operator's primary occupation is farming—as well as the nature of land ownership do not have strong associations with unexplained yields, with the exception of tenancy, for which a higher percent of agricultural acreage operated by tenants is associated with noticeably higher yields for soy, alfalfa, and corn (Figure S4C).

Another finding that merits further discussion is the interaction between farmer race and yields. In most regions of the U.S., white farmers operate well over 90% of cultivated land (Figure S4G and S7C). Despite the systematic underrepresentation of Black farmers in U.S. agriculture, increased acreage operated by Black farmers is associated with higher than expected yields, particularly for soy and hay (Figure S4E). Hispanic farmers are slightly more prevalent in areas with higher alfalfa yields (Figure S4F), while higher rates of female operated acreage are weakly associated with lower yields (Figure S4H). We note, however, that these findings are tentative given the low level of data availability for farmer race.

We find no clear association between indicators of landscape composition (diversity and percent natural cover), configuration (edge density and mean patch size) and yields (Figure 4C). Across crops, lower percent natural cover and higher mean patch area (larger contiguous areas of a single land cover type) are weakly associated with higher production levels (Figure S5).

We use random forests regression to determine which factors most strongly predict divergence from yield expectations (Figure 5). While we were unable to investigate racial associations with yields due to high levels of missing data, we included other available predictors of farm resources, farm(er) characteristics, and landscape context. We find that fertilizer and chemical expenses, which have strong positive associations with yield divergence (Figures S3A and S3B, Figure 4A), are the strongest predictors of yield divergence from expectations. Government receipts are also strongly predictive of divergence from expectations, also with a strongly positive association (Figure S3E, Figure 4A). Median farm size is also predictive of unexpected yields, with smaller farm sizes associated with increased yields (Figure S4D, Figure 4B). Finally, edge density and crop diversity emerge as landscape characteristics most predictive of unexpected yields, with predictive strengths similar to those for machinery and labor inputs and stronger than those of farm(er) characteristics, though the direction of these effects are unclear (Figure S5, Figure 4C).

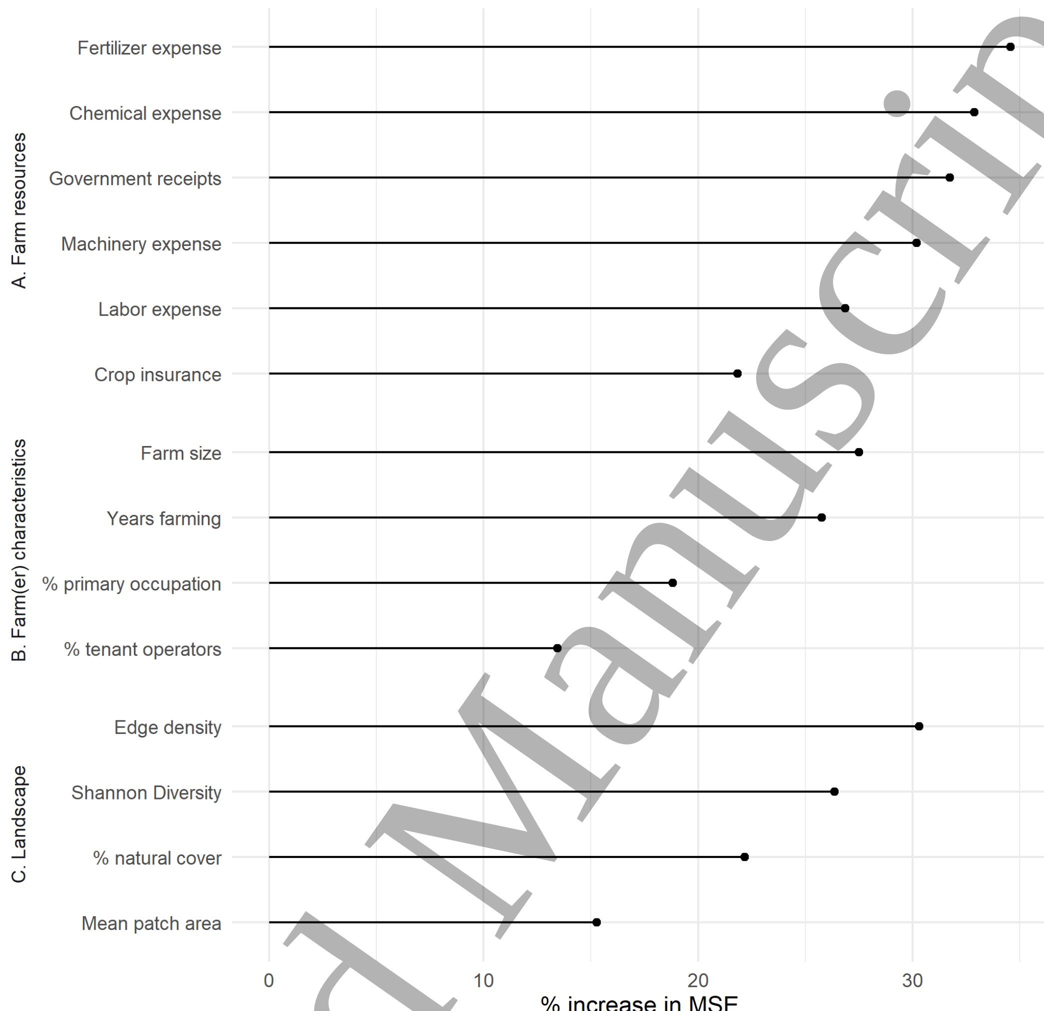

Figure 5: Variable importance plot showing the percent increase in mean squared error (MSE) associated with removing each predictor from the model. Model MSE on 25% held-out data was 0.287. Note that we had to exclude race from this analysis due to data availability. Results here predicting the weighted average of unexpected yields across crops.

4. Discussion

Our results reveal several interesting and important patterns in the yield geographies of major U.S. crops. First, for these five crops, yields that are not explained by differences in sun, soil, water, and regional norms exhibit a strong spatial structure, highlighting both the importance of explicitly accounting for spatial structure in yield models and the contribution of the broader context in which a county operates to agricultural yields. This finding also implies that major shifts in agricultural yields occur at a scale between U.S. county—the scale at which yields are systematically reported by the USDA—and USDA Farm Resource Regions. This is supported by the relatively small effects of Farm Resource Regions on yields (Figure S2), which suggests that yields vary nearly as much *within* regions as they do *across* regions.

Despite controlling for county-level differences in access to sun, soil, and water, the geographies presented in Figure 3 reveal yield gradients that strongly follow known ecological and biophysical boundaries. For example, despite controlling for irrigation in our models, corn, soy, and hay show abrupt yield shifts that align strongly with the boundaries of the High Plains Aquifer (Figure 6). Areas that are characterized by an abundance of surface water resources and highly suitable soil, such as the Mississippi River Valley, show unexpectedly high yields across all crops, while areas with more constrained water resources such as the Western U.S. show high yields for only a limited number of crops. We interpret the persistence of these biophysical signals as the human amplification of natural patterns of biophysical suitability. Agricultural production expands and intensifies in regions particularly well-suited to the cultivation of a specific crop [58]; increased agricultural activity further amplifies yield gains, increasing farm-level resources, and reinforcing and intensifying the suitability geographies defined by sun, soil, and water access.

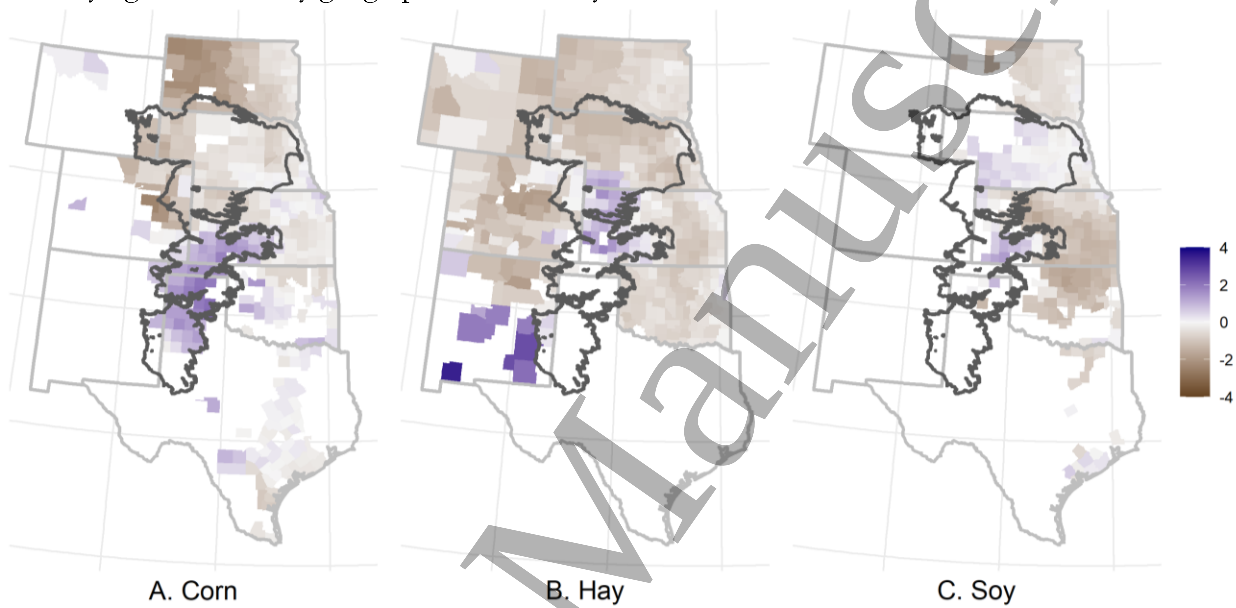

Figure 6: Yield divergence from expectations in the High Plains Aquifer.

This persistent spatial structure also highlights the *limits* to human amplification of biophysical suitability. Consider, for example, the regions at the periphery of the Midwestern U.S. that exhibit lower than expected yields for the two commodities that dominate this region: corn and soy. Over the last fifty years, the intersecting forces of consumer demand, trade policy, and market prices have pushed the cultivation of corn and soy beyond the regions that are most biophysically suited to their cultivation [5,59,60]. Since the 1960s, harvested corn and soy acreage has increased by 76%, with corn and soy alone comprising over 56% of harvested cropland in the U.S. [9].

The second major finding is that farm-level expenditures on fertilizer, chemicals, machinery, and labor are consistently higher in unexpectedly productive regions of the U.S. This finding is consistent across all crops and inputs (Figure S3). We estimate that the median total per acre expense on these inputs is $458 per acre in unexpectedly productive counties, over *two times* higher than the $185 per acre spent in low yielding counties. Farmers may benefit from increased input use *if* increased input expenditures result in higher yields; however, our comparison of livelihood indicators available from the USDA suggests that though crop sales are higher in unexpectedly productive regions, there is no observable difference in net farm income across this yield gradient (Figure 7A). This finding is consistent across crops—with the exception of soy and, to a lesser extent, alfalfa (Figure S6B). While this finding may be influenced by other farm activities beyond the cultivation of the five major crops examined here, such as high value livestock production or cultivation of high value specialty crops in areas with low crop yields, it nevertheless points to a disconnect between increased efforts and expenditures to boost production of the most widely grown crops in the U.S. and compensation for those efforts. The partial dependence plots from the random forest regression also suggest that increased input expenditure is associated with diminishing yield returns (Figure 7B). Only at exceptionally high expenditure rates of around $260 and $217 dollars per acre for fertilizer and chemical expense respectively are increased inputs associated with higher yields. This merits further exploration, especially given that per farm average expenditures—of which labor, machinery, fertilizer, and chemicals comprise over 30%—reached over $176,000 in 2017 [1]. These increased expenditures are strongly associated with unprecedented levels of farmer debt [13] and a nearly 40% drop in net farm income since 2013. [14,59,61]

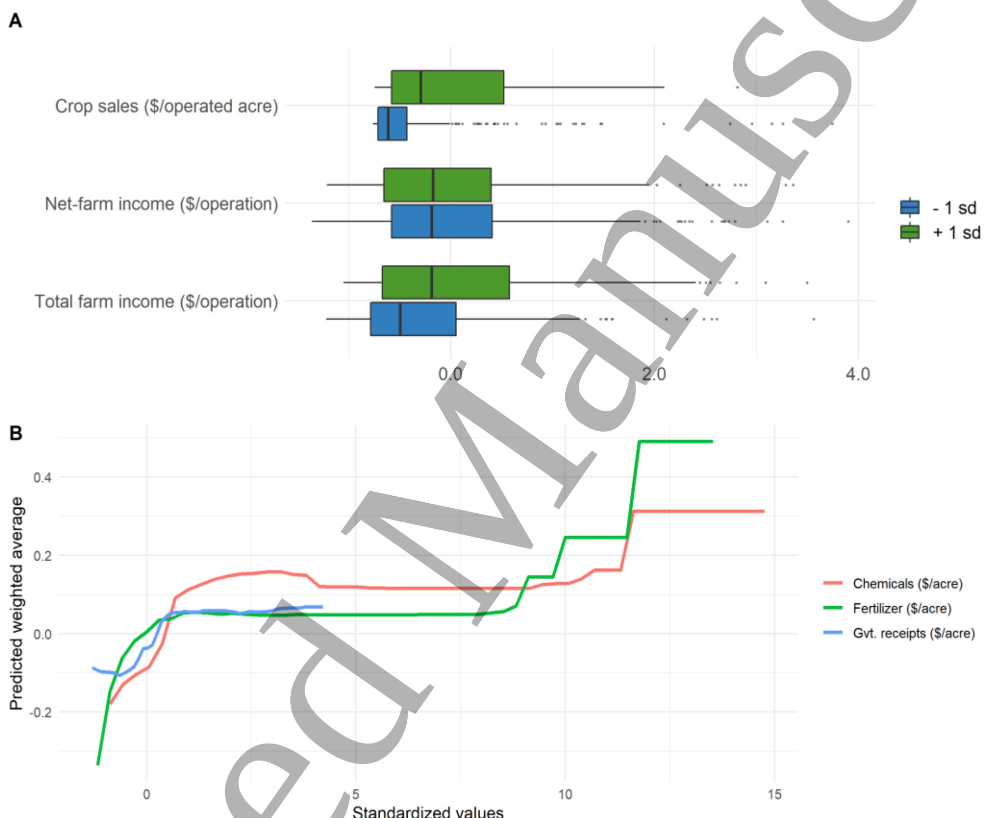

Figure 7: (A) Box and whisker plot showing differences in standardized county-level livelihood indicators across weighted average yield divergence from expectations. Bold vertical lines represent the median. Values are standardized for comparison. Here, net farm income is derived by subtracting total farm expenses from total sales, while total farm income is the average operation-scale income, before taxes and expenses. (B) Partial dependence plots for inputs and government receipts. Partial dependence is the dependence of the outcome on one predictor after averaging out the effects of the other predictors in the model, graphically characterizing the relationship between an individual predictor and the predicted values of yield divergence.

Farm-level decisions are strongly bounded by the larger structures, institutions, and policies that shape U.S. agriculture [5,9]. Our results highlight the strong influence of these factors on yield outcomes across the nation. First, the temporal effects we estimate are comparable in magnitude to the effects of seasonal weather on yields, highlighting the strong contribution of interannual variability in markets and policies to yield outcomes. Note, for example, the diverging effects for corn and soy over the period from 2010 to 2013—a period of tremendous price volatility for both crops. Note also, the steady decline in year effects from 2016 to 2018. This period is marked by geopolitical unpredictability as well as trade wars between the U.S. and China that significantly strained U.S. commodity exports. We also observe that farmers in unexpectedly productive regions receive over two times more support from government programs than those cultivating in unexpectedly low yielding regions. Government receipts are also highly predictive of unexplained yields (Figure 5), though like physical inputs to production, they show diminishing returns in their contribution to the prediction of unexpected yields (Figure 7B). The implications of the strong association between government receipts and high yields are mixed. Federal programs are an important source of income stabilization for U.S. farmers; however, research suggests that participation in these programs may have negative implications for farm-level adaptive capacity and resource use [62–65]. Large farms tend to have greater access to these programs[2,4], and while the link between farm size and yields is mixed[11], this finding may highlight that higher-yielding systems are those that are more likely to have access to and benefit from government support.

The third major finding is that while there are no clear overall associations between farmer characteristics and unexplained yields, increased acreage operated by Black farmers is associated with higher than expected yields for several crops. This is important given the relative challenges Black, Indigenous, and other farmers of color face in accessing land and capital [66,67] and the ways in which government programs (strongly associated with higher yields) have been shown to systematically privilege white farmers [68]. Our exploration of available data also highlights the strong spatial concentration of minority farmers, with Black operators cultivating primarily in the southeastern, and Hispanic operators in the southwestern U.S. (Figure S7).

Fourth, and rather surprisingly, larger farms are associated with *lower* than expected yields. Over the last decades, the number of small- and mid-sized farms in the U.S. has declined. In 2018, large-scale farms (sales greater than 500K) covered over 40% of agricultural land but comprised only 7.5% of the total number of farms [1,9]. The shift towards large-scale industrial farming is strongly driven by the assumption that larger farms are more productive [69,70]. Though this assumption has been widely debated[69], recent evidence suggests that in developed economies large farms tend to be more productive[3,10,11]. This evidence is particularly strong in grain production, where many recent technological advances, such as large combine harvesters and precision agricultural technologies, most strongly benefit large-scale farms [3,11]. While we note that while yields (outputs) are distinct from productivity (the ratio of inputs to outputs), this finding provides some preliminary evidence that, after controlling for the natural assets of a farm, size may not necessarily bring high yields.

Finally, though no clear associations between landscape context and yields emerge from these analyses, for specific crops, counties with larger contiguous areas of the same crop see higher than expected yields. While this appears to contradict the finding that larger farm operations are associated with lower than expected yields, it may suggest that it is not the *size* of the operation but the extent of *specialization* within the operation that is associated with yield gains. We also observe lower than expected yields in areas with greater percent natural cover, though this likely reflects the fact that areas with more natural cover may also be those that are less well-suited to agricultural production. Though no major differences emerge across yield divergence from expectations (Figure 4C), edge density and agricultural diversity are strongly predictive of yields. This suggests that landscape characteristics, though not significantly different across yield divergence from expectations, still show interesting nonlinear impacts on the prediction of yields after accounting for other factors. The difference here also highlights how the effects of these landscape characteristics on yields are likely nonlinear and not fully captured by high to low yield gradient we explore. As edge density increases, predicted production decreases, suggesting that increased configurational complexity of a landscape may not necessarily be good for production. Interestingly, the effect of crop diversity on production corroborates published research suggesting that diversity, particularly high levels of diversity, is associated with higher yields [57,71] (Figure S8B).

5. Conclusions

Taken together, these results highlight the ways in which human activity has amplified biophysical suitability signals in the geography of U.S. agricultural production. While unexpectedly high yields are strongly associated with increased expenditure on major inputs to production and higher federal support, they are *not* strongly associated with net revenue gains for farmers. We also find that beyond a threshold, increased input expenditure is associated with marginally decreasing returns to production. This provides preliminary evidence that the *intensification* of agricultural production in highly suitable areas is not necessarily associated with improvements in net farm revenues. This raises important questions about the tradeoffs between yields and farm livelihoods. What are the ultimate goals of agricultural production? Meeting human demand for food, fuel, and fiber is a must, but what of the livelihoods of the individuals who grow our food? And what of the environmental base that supports all food production? Our analyses suggest that some of the most unexpectedly productive regions of the U.S. fail to meet the equally important objectives of economic viability and environmental sustainability.

The geographies revealed by these analyses also highlight regions of concern. First, the strong yield signal in the High Plains Aquifer highlights production intensification brought by heavy groundwater irrigation. This aquifer is expected to lose an estimated 24% of irrigated area in this century due to agricultural withdrawals [72], and whether these high yields can be sustained into the future is currently under debate [20,73,74]. Second, federal support of corn and soy production has pushed corn and soy production beyond the periphery of the Corn Belt, into regions we find have lower than expected yield outcomes for these crops (29, 42). This expansion has displaced other crops that may be better suited to this region and exerted a number of environmental externalities [76–78]. While changing climate that will likely *increase* the biophysical suitability of these peripheral regions to the cultivation of corn and soy [23,60], the strong market and policy signals driving the expansion of corn and soy may limit farmer capacity to explore more sustainable alternatives [79].

Many have argued that meeting projected food demand while also preserving rural livelihoods and ecological integrity will require a significant reconceptualization of the ultimate goals of agriculture [80–82]. One such reconceptualization proposed by the USDA defines *sustainable* agriculture as an integrated system that satisfies human demand while *also* enhancing environmental quality, efficiently using natural resources, supporting farmer livelihoods, and enhancing quality of life for farmers and society as a whole [83]. Our research shows that while U.S. agriculture excels at meeting human demand, we fall short in satisfying other important objectives, raising serious concerns about the economic and ecological viability of U.S. agriculture over time. We applaud and encourage those in the agricultural community who are critically expanding the production-centric paradigm to reposition agriculture as a force for positive, regenerative change for both people and planet. We also urge members of the scientific community to continue to reimagine and redefine agriculture as tool to cultivate socioecological well-being.

References

1. USDA National Agricultural Statistics Service. NASS - QuickStats. *2020* (2020).
2. Bruckner, T. Agricultural Subsidies and Farm Consolidation. *Am. J. Econ. Sociol.* **75**, 623–648 (2016).
3. Foster, A. D. & Rosenzweig, M. R. *Are There Too Many Farms in the World? Labor-Market Transaction Costs, Machine Capacities and Optimal Farm Size*. https://www.nber.org/papers/w23909 (2017) doi:10.3386/w23909.
4. Graddy-Lovelace, G. & Diamond, A. From supply management to agricultural subsidies—and back again? The U.S. Farm Bill & agrarian (in)viability. *J. Rural Stud.* **50**, 70–83 (2017).
5. McGranahan, D. A., Brown, P. W., Schulte, L. A. & Tyndall, J. C. A historical primer on the US farm bill: Supply management and conservation policy. *J. Soil Water Conserv.* **68**, 67A-73A (2013).
6. Alston, J. M. & Pardey, P. G. Agriculture in the Global Economy. *J. Econ. Perspect.* **28**, 121–146 (2014).
7. Woods, M. Engaging the global countryside: globalization, hybridity and the reconstitution of rural place. *Prog. Hum. Geogr.* **31**, 485–507 (2007).
8. USDA-NASS. USDA National Agricultural Statistics Service Cropland Data Layer. (2020).
9. Spangler, K., Burchfield, E. K. & Schumacher, B. Past and Current Dynamics of U.S. Agricultural Land Use and Policy. *Front. Sustain. Food Syst.* **4**, 98 (2020).
10. Sheng, Y. & Chancellor, W. Exploring the relationship between farm size and productivity: Evidence from the Australian grains industry. *Food Policy* **84**, 196–204 (2019).
11. Key, N. Farm size and productivity growth in the United States Corn Belt. *Food Policy* 10 (2019).
12. Rada, N. E. & Fuglie, K. O. New perspectives on farm size and productivity. *Food Policy* **84**, 147–152 (2019).
13. Key, N., Burns, C. & Lyons, G. Financial Conditions in the U.S. Agricultural Sector: Historical Comparisons. 36.
14. USDA ERS - Farm Household Income and Characteristics. https://www.ers.usda.gov/data-products/farm-household-income-and-characteristics/farm-household-income-and-characteristics/#Farm%20Household%20Income%20Forecast.
15. Manjunatha, A. V., Anik, A. R., Speelman, S. & Nuppenau, E. A. Impact of land fragmentation, farm size, land ownership and crop diversity on profit and efficiency of irrigated farms in India. *Land Use Policy* **31**, 397–405 (2013).
16. Lowder, S. K., Skoet, J. & Raney, T. The Number, Size, and Distribution of Farms, Smallholder Farms, and Family Farms Worldwide. *World Dev.* **87**, 16–29 (2016).
17. Conger, R. *Families in Troubled Times: Adapting to Change in Rural America*. (Routledge, 2020).
18. Hendricks, N. P. *et al.* The environmental effects of crop price increases: Nitrogen losses in the U.S. Corn Belt. *J. Environ. Econ. Manag.* **68**, 507–526 (2014).
19. Bouët, A. & Laborde, D. *US Trade Wars with Emerging Countries in the 21st Century: Make America and Its Partners Lose Again*. https://papers.ssrn.com/abstract=3029615 (2017).
20. Cotterman, K. A., Kendall, A. D., Basso, B. & Hyndman, D. W. Groundwater depletion and climate change: future prospects of crop production in the Central High Plains Aquifer. *Clim. Change* **146**, 187–200 (2018).
21. Schlenker, W., Hanemann, W. M. & Fisher, A. C. Water availability, degree days, and the potential impact of climate change on irrigated agriculture in California. *Clim. Change* **81**, 19–38 (2007).
22. Wang, S. L., Ball, V. E., Nehring, R., Williams, R. & Chau, T. Impacts of Climate Change and Extreme Weather on US Agricultural Productivity: Evidence and Projection. in *Understanding Productivity Growth in Agriculture* (University of Chicago Press, 2017).

23. Burchfield, E., Matthews-Pennanen, N., Schoof, J. & Lant, C. Changing yields in the Central United States under climate and technological change. *Clim. Change* **159**, 329–346 (2020).
24. Nelson, G. C. *et al.* Climate change effects on agriculture: Economic responses to biophysical shocks. *Proc. Natl. Acad. Sci.* **111**, 3274–3279 (2014).
25. Bommarco, R., Vico, G. & Hallin, S. Exploiting ecosystem services in agriculture for increased food security. *Glob. Food Secur.* **17**, 57–63 (2018).
26. Duarte, G. T., Santos, P. M., Cornelissen, T. G., Ribeiro, M. C. & Paglia, A. P. The effects of landscape patterns on ecosystem services: meta-analyses of landscape services. *Landsc. Ecol.* **33**, 1247–1257 (2018).
27. Lawler, J. J. *et al.* Projected land-use change impacts on ecosystem services in the United States. *Proc. Natl. Acad. Sci.* **111**, 7492–7497 (2014).
28. German, R. N., Thompson, C. E. & Benton, T. G. Relationships among multiple aspects of agriculture's environmental impact and productivity: a meta-analysis to guide sustainable agriculture. *Biol. Rev.* **92**, 716–738 (2017).
29. Prokopy, L. S. *et al.* The urgency of transforming the Midwestern U.S. landscape into more than corn and soybean. *Agric. Hum. Values* 1–3 (2020) doi:10.1007/s10460-020-10077-x.
30. Altieri, M. A. *Agroecology: The Science Of Sustainable Agriculture, Second Edition*. (CRC Press, 2018).
31. 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**, (2008).
32. USDA National Agricultural Statistics Service. Crop values: 2019 summary. (2019).
33. PRISM Climate Group, Oregon State University, http://prism.oregonstate.edu, created 4 Feb 2004.
34. Ramankutty, N., Evan, A., Monfreda, C. & A. Foley, J. *Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000*. vol. 22 (2008).
35. *Agweather connection*. https://www.mesonet.org/mesonet_connection/V2_No8.pdf.
36. *Corn growing degree days*. https://ndawn.ndsu.nodak.edu/help-corn-growing-degree-days.html.
37. Pervez, M. S. & Brown, J. F. Mapping Irrigated Lands at 250-m Scale by Merging MODIS Data and National Agricultural Statistics. *Remote Sens.* **2**, 2388–2412 (2010).
38. Nachtergaele, F. O. *et al.* Harmonized World Soil Database (version 1.0). (2008).
39. Lobell, D. B. *et al.* Greater sensitivity to drought accompanies maize yield increase in the US Midwest. *Science* **344**, 516–519 (2014).
40. Schlenker, W. & Roberts, M. J. Nonlinear temperature effects indicate severe damages to US crop yields under climate change. *Proc. Natl. Acad. Sci.* **106**, 15594–15598 (2009).
41. Besag, J., York, J. & Mollié, A. Bayesian image restoration, with two applications in spatial statistics. *Ann. Inst. Stat. Math.* **43**, 1–20 (1991).
42. Heimlich, R. E. *Farm Resource Regions*. https://ideas.repec.org/p/ags/uersab/33625.html (2000).
43. Rue, H., Martino, S. & Chopin, N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **71**, 319–392 (2009).
44. Simpson, D. P., Rue, H., Martins, T. G., Riebler, A. & Sørbye, S. H. Penalising model component complexity: A principled, practical approach to constructing priors. *ArXiv14034630 Stat* (2014).
45. Liaw, A. & Wiener, M. Classification and Regression by randomForest. **2**, 5 (2002).
46. Lobell, D. B. *et al.* The critical role of extreme heat for maize production in the United States. *Nat. Clim. Change* **3**, 497–501 (2013).
47. Rosenzweig, C. *et al.* Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *Proc. Natl. Acad. Sci.* **111**, 3268–3273 (2014).

48. Schauberger, B. *et al.* Consistent negative response of US crops to high temperatures in observations and crop models. *Nat. Commun.* **8**, 13931 (2017).
49. Baumgart-Getz, A., Prokopy, L. S. & Floress, K. Why farmers adopt best management practice in the United States: A meta-analysis of the adoption literature. *J. Environ. Manage.* **96**, 17–25 (2012).
50. Knowler, D. & Bradshaw, B. Farmers' adoption of conservation agriculture: A review and synthesis of recent research. *Food Policy* **32**, 25–48 (2007).
51. Dainese, M. *et al.* A global synthesis reveals biodiversity-mediated benefits for crop production. *bioRxiv* (2019) doi:10.1101/554170.
52. Garibaldi, L. A. *et al.* Mutually beneficial pollinator diversity and crop yield outcomes in small and large farms. *Science* **351**, 388–391 (2016).
53. Hass, A. L. *et al.* Maize-dominated landscapes reduce bumblebee colony growth through pollen diversity loss. *J. Appl. Ecol.* **0**,.
54. Fahrig, L. *et al.* Farmlands with smaller crop fields have higher within-field biodiversity. *Agric. Ecosyst. Environ.* **200**, 219–234 (2015).
55. Fahrig, L. *et al.* Functional landscape heterogeneity and animal biodiversity in agricultural landscapes: Heterogeneity and biodiversity. *Ecol. Lett.* **14**, 101–112 (2011).
56. Sirami, C. *et al.* Increasing crop heterogeneity enhances multitrophic diversity across agricultural regions. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 16442–16447 (2019).
57. Burchfield, E. K., Nelson, K. S. & Spangler, K. The impact of agricultural landscape diversification on U.S. crop production. *Agric. Ecosyst. Environ.* **285**, 106615 (2019).
58. Lark, T. J., Salmon, J. M. & Gibbs, H. K. Cropland expansion outpaces agricultural and biofuel policies in the United States. *Environ. Res. Lett.* **10**, 044003 (2015).
59. Parton, W. J., Gutmann, M. P. & Ojima, D. Long-term Trends in Population, Farm Income, and Crop Production in the Great Plains. *BioScience* **57**, 737–747 (2007).
60. Lant, C., Stoebner, T. J., Schoof, J. T. & Crabb, B. The effect of climate change on rural land cover patterns in the Central United States. *Clim. Change* **138**, 585–602 (2016).
61. Mishra, A. K. & Sandretto, C. L. Stability of Farm Income and the Role of Nonfarm Income in U.S. Agriculture. *Rev. Agric. Econ.* **24**, 208–221 (2002).
62. Annan, F. & Schlenker, W. Federal Crop Insurance and the Disincentives to Adapt to Extreme Heat. 23.
63. Di Falco, S. & Perrings, C. Crop biodiversity, risk management and the implications of agricultural assistance. *Ecol. Econ.* **55**, 459–466 (2005).
64. O'Donoghue, E. J., Roberts, M. J. & Key, N. Did the Federal Crop Insurance Reform Act Alter Farm Enterprise Diversification? *J. Agric. Econ.* **60**, 80–104 (2009).
65. Bruckner, T. Agricultural Subsidies and Farm Consolidation. *Am. J. Econ. Sociol.* **75**, 623–648 (2016).
66. Calo, A. & De Master, K. After the Incubator: Factors Impeding Land Access Along the Path from Farmworker to Proprietor. *J. Agric. Food Syst. Community Dev.* 1–17 (2016) doi:10.5304/jafscd.2016.062.018.
67. Horst, M. & Marion, A. Racial, ethnic and gender inequities in farmland ownership and farming in the U.S. *Agric. Hum. Values* **36**, 1–16 (2019).
68. Minkoff-Zern, L.-A. & Sloat, S. A new era of civil rights? Latino immigrant farmers and exclusion at the United States Department of Agriculture. *Agric. Hum. Values* **34**, 631–643 (2017).
69. Rada, N. E. & Fuglie, K. O. New perspectives on farm size and productivity. *Food Policy* **84**, 147–152 (2019).

70. Duffy, M. Economies of Size in Production Agriculture. *J. Hunger Environ. Nutr.* **4**, 375–392 (2009).
71. Renard, D. & Tilman, D. National food production stabilized by crop diversity. *Nature* **571**, 257–260 (2019).
72. Deines, J. M. *et al.* Mapping three decades of annual irrigation across the US High Plains Aquifer using Landsat and Google Earth Engine. *Remote Sens. Environ.* **233**, 111400 (2019).
73. Rudnick, D. R. *et al.* Deficit Irrigation Management of Maize in the High Plains Aquifer Region: A Review. *JAWRA J. Am. Water Resour. Assoc.* **55**, 38–55 (2019).
74. Ajaz, A., Datta, S. & Stoodley, S. High Plains Aquifer–State of Affairs of Irrigated Agriculture and Role of Irrigation in the Sustainability Paradigm. *Sustainability* **12**, 3714 (2020).
75. Green, T. R., Kipka, H., David, O. & McMaster, G. S. Where is the USA Corn Belt, and how is it changing? *Sci. Total Environ.* **618**, 1613–1618 (2018).
76. Wright, C. K. & Wimberly, M. C. Recent land use change in the Western Corn Belt threatens grasslands and wetlands. *Proc. Natl. Acad. Sci.* **110**, 4134–4139 (2013).
77. Lu, C. *et al.* Increasing carbon footprint of grain crop production in the US Western Corn Belt. *Environ. Res. Lett.* **13**, 124007 (2018).
78. Berti, M., Johnson, B., Ripplinger, D., Gesch, R. & Aponte, A. Environmental impact assessment of double- and relay-cropping with winter camelina in the northern Great Plains, USA. *Agric. Syst.* **156**, 1–12 (2017).
79. Roesch-McNally, G. E., Arbuckle, J. G. & Tyndall, J. C. Barriers to implementing climate resilient agricultural strategies: The case of crop diversification in the U.S. Corn Belt. *Glob. Environ. Change* **48**, 206–215 (2018).
80. Perfecto, I. & Vandermeer, J. The agroecological matrix as alternative to the land-sparing/agriculture intensification model. *Proc. Natl. Acad. Sci.* **107**, 5786–5791 (2010).
81. Bommarco, R., Kleijn, D. & Potts, S. G. Ecological intensification: harnessing ecosystem services for food security. *Trends Ecol. Evol.* **28**, 230–238 (2013).
82. Coomes, O. T., Barham, B. L., MacDonald, G. K., Ramankutty, N. & Chavas, J.-P. Leveraging total factor productivity growth for sustainable and resilient farming. *Nat. Sustain.* **2**, 22–28 (2019).
83. Sustainable Agriculture Program | National Institute of Food and Agriculture. https://nifa.usda.gov/program/sustainable-agriculture-program.