nature sustainability

Article

https://doi.org/10.1038/s41893-023-01152-2

# Key role of planted and harvested area fluctuations in US crop production shocks

Received: 21 November 2022

Accepted: 22 May 2023

Published online: 15 June 2023

Check for updates

Dongyang Wei[1], Jessica A. Gephart[2], Toshichika Iizumi[3], Navin Ramankutty[4] & Kyle Frankel Davis[1,5]

Food production stability against climate variability and extremes is crucial for food security and is influenced by variations in planted area, harvested area and yield. Yet research has focused on yield responses to climate fluctuations, ignoring how planted area and harvestable fraction (that is, the ratio of planted area to harvested area) affect production stability. Here we apply a time series shock detection approach to county-level data (1978–2020) on seven crops in the United States, finding that shocks (that is, sudden statistically significant declines) in planted area and harvestable fraction co-occur with 51–81% of production shocks, depending on the crop. Decomposing production shock magnitudes, we find that yield fluctuations contribute more for corn (59%), cotton (49%), soybean (64%) and winter wheat (40%), whereas planted area and harvestable fraction have a greater role for others. Additionally, climatic variables explain considerable portions of the variance in planted area (22–30%), harvestable fraction (15–28%) and yield (32–50%). These findings demonstrate that crop production shocks are often associated with fluctuations in planted area and harvestable fraction. This highlights the (largely ignored) importance of producer decision-making about cropping patterns in stabilizing food production against climate variability and emphasizes the need to consider all three production components to improve food system stability.

Many countries are facing growing levels of food insecurity, with 193 million people acutely food insecure worldwide[1]. To achieve the United Nations' Sustainable Development Goal of Zero Hunger (SDG2) by 2030, urgent action is needed to ensure food security in all aspects. Least studied among the four food security pillars (availability, access, use and stability) is food stability, which refers to the ability of an individual, household or population to have reliable access to adequate, safe and nutritious food[2–4]. While stability can be affected at any step in the food supply chain, the largest number of disruption entry points are found at the production stage[5]. Food production shocks (that is, sudden and unexpected losses in production) can be caused by a wide variety of factors, including climate variability, extreme weather events and economic and political disruptions[6–8]. With increasing climate variability[9] and climate extremes expected to become more frequent, intense and prolonged[10–12], it is critical to understand the pathways through which environmental shocks impact production to develop more effective strategies for stabilizing crop production.

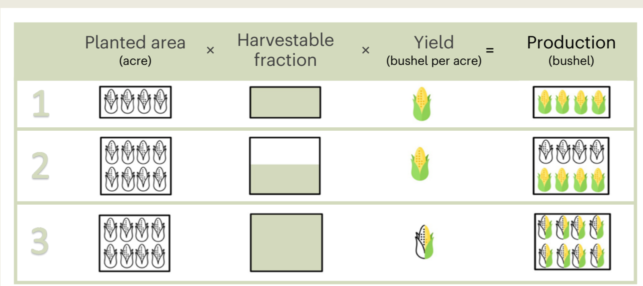
Food production instability (that is, the occurrence and magnitude of year-to-year variability for a certain period) is determined by variability and shocks in planted area, harvestable fraction (that is, the ratio between planted and harvested area) and yield (Box 1), each involving varying degrees of human decision-making. Changes in planted area are determined mainly by farmer decisions before the growing season based on economic, policy and climatic conditions[13,14].

[1]Department of Geography and Spatial Sciences, University of Delaware, Newark, DE, USA. [2]Department of Environmental Science, American University, Washington, DC, USA. [3]Institute for Agro-Environmental Sciences, National Agriculture and Food Research Organization, Tsukuba, Japan. [4]School of Public Policy and Global Affairs, Institute for Resources, Environment and Sustainability, University of British Columbia, Vancouver, British Columbia, Canada. [5]Department of Plant and Soil Sciences, University of Delaware, Newark, DE, USA. e-mail: dywei@udel.edu

**BOX 1**

## Production outcomes of component shocks

Crop production is calculated as the product of yield and harvested area. Harvested area can be further separated into planted area×harvestable fraction (calculated as the ratio of harvested to planted area). Each of these components can suffer a sudden loss independent of the other but still have the same consequences for production outcomes, as shown in the illustration below. In the first scenario, a smaller amount of area is planted compared to the other two scenarios, but there are no shocks to harvestable fraction or yield. In the second scenario, only half of the planted area is harvestable but yield is unaffected. In the third scenario, there is no difference between planted area and harvested area but yield is reduced by half. Across all of these scenarios, shocks in different components contribute to the same amount of production loss. In addition, should more than one component experience a shock at the same time, this would collectively amplify the resultant production shock. Note that these scenarios presume a single season per year and the sum of seasonal production is additionally required for multi-crop systems.




Conversely, harvestable fraction (that is, the portion of planted area that is harvestable rather than lost within-season) is influenced by exogenous natural forces—such as climate extremes—and to some extent farmer decisions. Flooding, for example, can cause a portion of a field to be washed away, thereby reducing the harvested area but leaving yield unaffected[15]. Farmers may also decide not to harvest their crops—perhaps due to low yields, inferior quality or low market prices—because the expected low revenue would not justify their time and effort[16,17]. For yield, changes are jointly dictated by within-season management decisions (for example, irrigation, varietal choice) and environmental conditions (for example, heatwave, drought and pests). Yet while all three of these components (planted area, harvestable fraction and yield) can influence production outcomes, most research on production instability to date has focused on the role of yield variations[7,18–23]. However, there are emerging efforts to understand how the different components of production contribute to its stability. For instance, some studies showed that production losses were associated with both harvested area and yield[24,25]. Other work in Brazil showed that harvested area and cropping frequency were more sensitive to climate variability than yield[26]. While these few studies suggest the importance of other non-yield components for determining production outcomes, the extent to which all three components of production (planted area, harvestable fraction and yield) influence stability across different crops and regions is unknown[14]. Improving our understanding beyond yield variations can provide a more complete picture of the vulnerabilities of current crop production practices and can better inform strategies for adapting crop production to increasingly variable and extreme climate and other natural and human-made disruptions.

In this study, we focused on crop production shocks in the United States, the world's largest producer and exporter of cereal grains[27]. Because of its important role in the global food system, understanding the components that most contribute to US production shocks can improve strategies to ensure stable and reliable crop production and better protect national and global food supplies. To this end, we investigated how variations in planted area, harvestable fraction and yield contribute to US crop production shocks and the extent to which these three components are affected by climate variability and extremes. We first assembled 43 years (1978–2020) of county-level agricultural data for seven major crops (barley, corn, cotton, sorghum, soybean, spring wheat and winter wheat), which account for 70% of US cropland[28]. We then detected shocks (that is, sudden and statistically significant decreases) in production and its components—planted area, harvestable fraction and yield—using an automated quantitative statistical method that captures sudden changes in time series while ignoring long-term gradual fluctuations[29]. Through this approach, we quantified the number of years with production shocks (frequency) and estimated their co-occurrence with shocks in each of the three components. We then used a decomposition approach to investigate to what extent each of the three components contributes to the magnitude of production shocks[30]. Finally, we built random forest regression models to determine to what extent interannual variations in production and its three components are explained by climate variability and extremes. Together, these lines of investigation can provide valuable insights beyond the role of yield in influencing production stability and can serve as a basis for expanding the option space for interventions to address climate-related crop production losses.

## Production shock frequency

Using an automated quantitative statistical shock detection method[29], we detected instances of negative deviations of production (hereafter, production shocks) ranging from 449 total negative shocks (for spring wheat) to 2,532 shocks (for corn) in all counties between 1978 and 2020 (shock can only be detected from the second year), the years for which data were available for all study crops. Production shocks varied both spatially and temporally between crops (Fig. 1). Production shock frequency has increased significantly for corn, cotton and soybean ($P < 0.05$, two-tailed Mann–Kendall test), with no significant trends in shock frequency being observed for all other study crops. In terms of geographical heterogeneity, we observed higher shock frequencies in Iowa, Illinois and Missouri for corn, soybean and winter wheat, and in North Dakota for barley and spring wheat (Fig. 1).

We then compared the co-occurrence of production shocks with shocks in each of the three components of production. We found that more than half of the production shocks for six of the seven study crops (barley, corn, cotton, sorghum, winter wheat and spring wheat) co-occurred with shocks of area-related components (Fig. 2 and Supplementary Fig. 1). Conversely, for soybean, the association with yield shocks dominates, co-occurring with 65% of production shocks. Across all seven crops, shocks related to planted area co-occurred with between 33% and 53% of production shocks, whereas shocks associated with harvestable fraction co-occurred with between 19% and 43% of production shocks. We found that between 17% and 31% of production shocks were associated with a combination of yield and area-related shocks, highlighting the fact that these components of production are not entirely independent of one another, depending on the nature of the disruption. We also compared shock outcomes between rain-fed and irrigated conditions across three crops—corn, winter wheat and soybean—for which there were sufficient data from 1978 to 2018 (Supplementary Fig. 2). Not surprisingly, we found that shocks of area-related components co-occur more often with production shocks

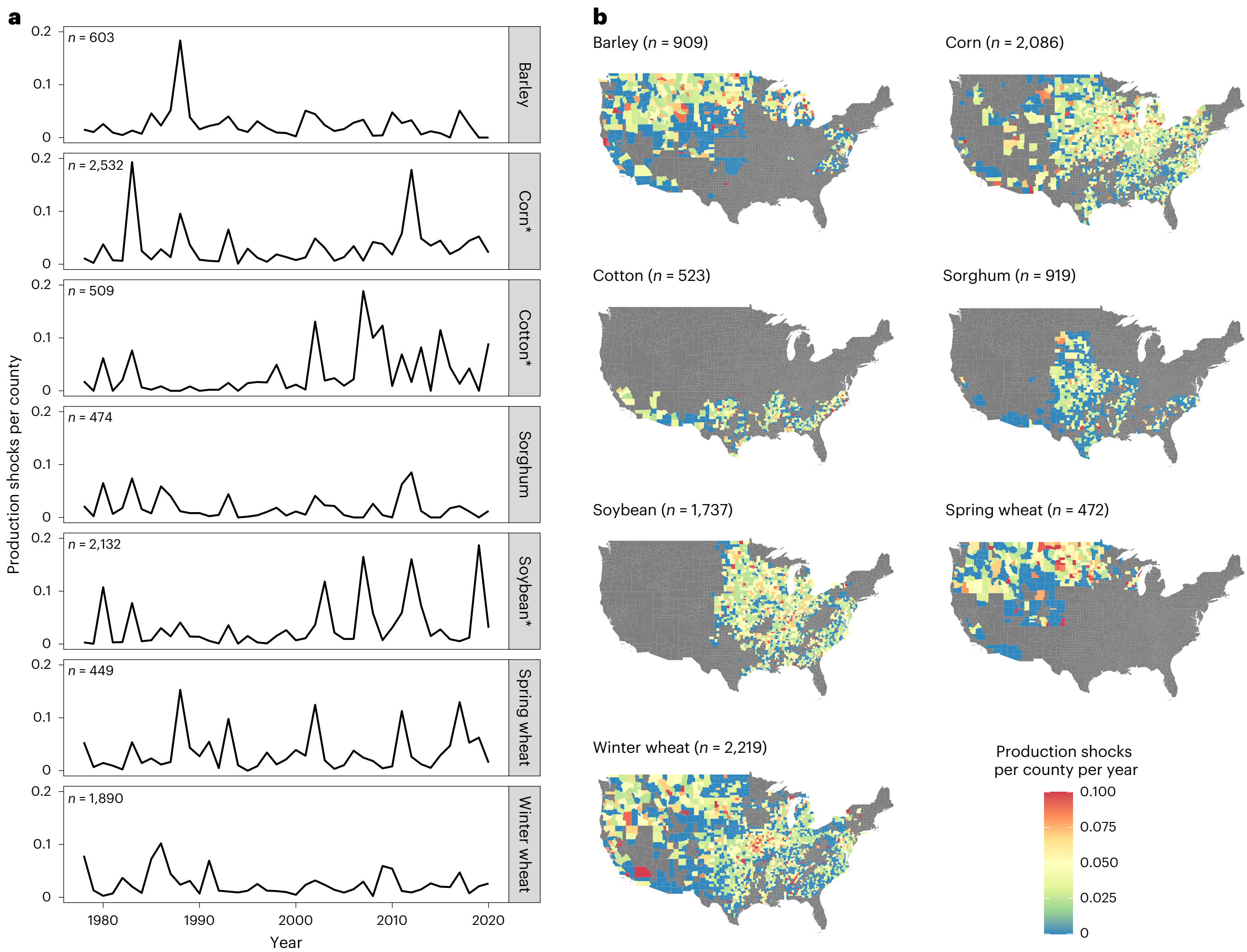



**Fig. 1 | Production shock frequency and hotspot maps for study crops.** **a**, Production shock per county from 1978 to 2020. The asterisk indicates a statistically significant trend (two-sided $P < 0.05$; corn, $P = 0.015$; cotton, $P = 0.009$; soybean, $P = 0.017$). *n* represents the total number of shocks for the whole period. **b**, Maps of shock frequency in each county over the study period. The total number of counties (*n*) examined is included in parentheses. Values higher than 0.1 are displayed in red.

under irrigated conditions (potentially due to the buffering effects of irrigation on yield), whereas under rain-fed conditions, shocks related to yield consist of most production shock co-occurrences. Looking across all component shocks (that is, not limited to those co-occurring with production shocks), we found that shocks in any individual component are unlikely to result in production shocks (Supplementary Table 1). Further, we found that the presence of co-occurring shocks in harvestable fraction and yield is still unlikely (except for corn) to result in a production shock in the same year, suggesting that in many cases planted area may have had an important compensatory role in mitigating production shocks (Supplementary Table 1).

## Production shock magnitude

We next quantified the magnitude of each of the detected production shocks and decomposed the contributions of each of the three components. On average, yield accounted for the largest portion of the production shock magnitude (31–64%) across the study crops, followed closely by planted area (22–53%) and then harvestable fraction (5–29%) (Fig. 3 and Supplementary Figs 3 and 4). Yield was dominant in explaining the magnitude of production shocks for corn (average of 59% across available years), cotton (49%), soybean (64%) and winter wheat (40%). Planted area was more important for barley (50% on average), sorghum (48%) and spring wheat (53%). As expected, the contribution of harvestable fraction shocks was larger for crops with longer growing periods (GPs), while it was smaller for crops with shorter GPs. Over time, we also observed changing influences of the three components on production shock magnitude. For instance, while harvestable fraction represents a relatively small contribution to production shock magnitude for most crops, we saw an overall statistically significant increasing trend ($P < 0.01$) of its contribution in corn (Supplementary Table 2), and a fluctuating large proportion in cotton and winter wheat (Fig. 3). We also found a significant decreasing trend ($P < 0.01$) of the contribution of yield in cotton (Supplementary Table 2).

## Links between climate indicators and agricultural factors

Lastly, we used random forest models to examine the associations between climate variables (that is, climate variability and extremes) (Supplementary Table 3) and anomalies in planted area, harvestable fraction and yield. While we expected that anomalies in these components would largely be affected by climate variables in the same GP, we also considered a 1-GP lag between climate variables and planted area

https://doi.org/10.1038/s41893-023-01152-2

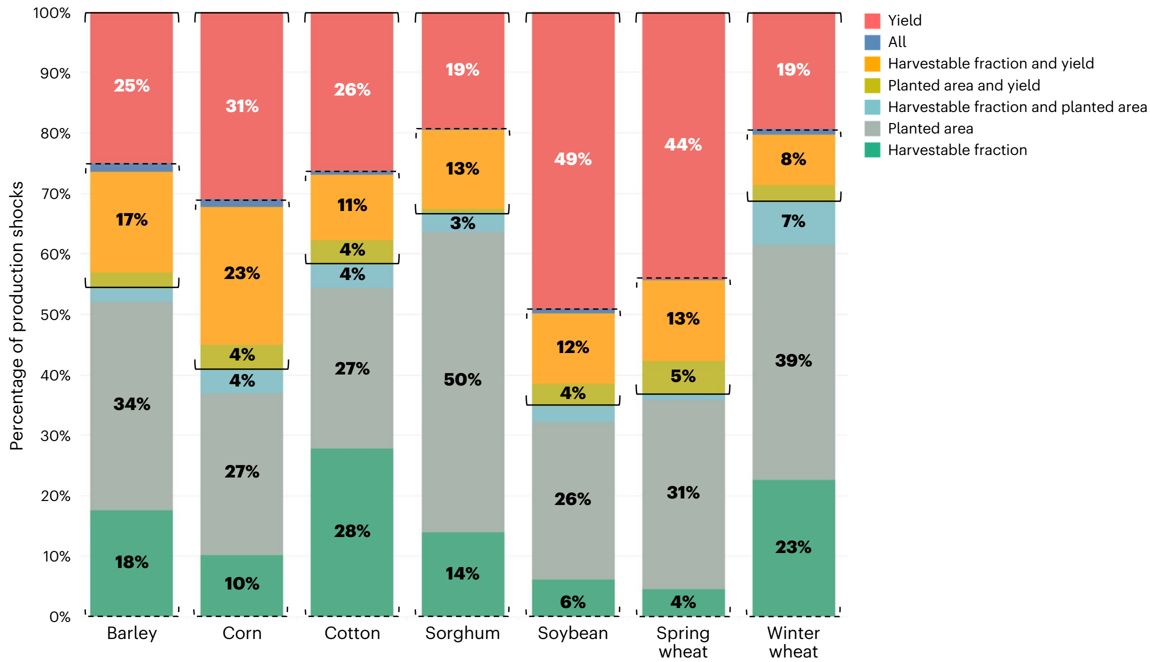



**Fig. 2 | Proportion of production shocks co-occurring with different component shocks.** The solid brackets indicate yield-related shocks and the dashed brackets include area-related shocks. Note that a small fraction of production shocks co-occurred with both yield- and area-related shocks. The fraction of total detected production shocks that did not have co-occurring shocks with any of the three components are not shown. The two main reasons for no co-occurrence were: (1) the changes in the three components were minor but amplified one another; or (2) there was high variability in the time series of one or more of the components and a shock was not statistically detectable.

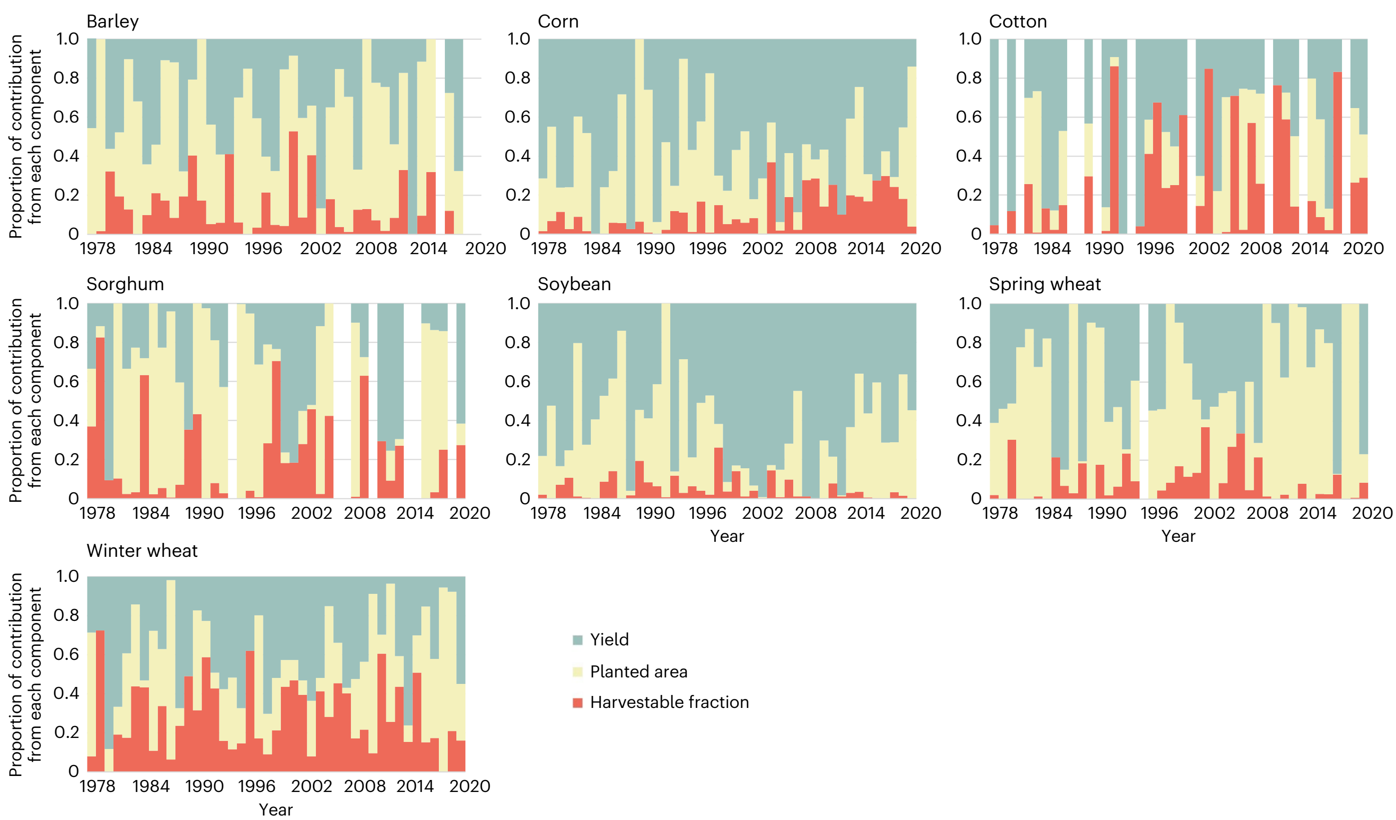



**Fig. 3 | Annual contribution from planted area, harvestable fraction and yield to shock-related production losses.** For each year for a specific crop, counties with production shocks were summed to represent the national production loss due to production shocks. Each component's contribution to the total loss was then calculated using shock decomposition[30]. The gap years (for example, 1979 for cotton) mean that no production shock was detected across all counties for that crop.

https://doi.org/10.1038/s41893-023-01152-2

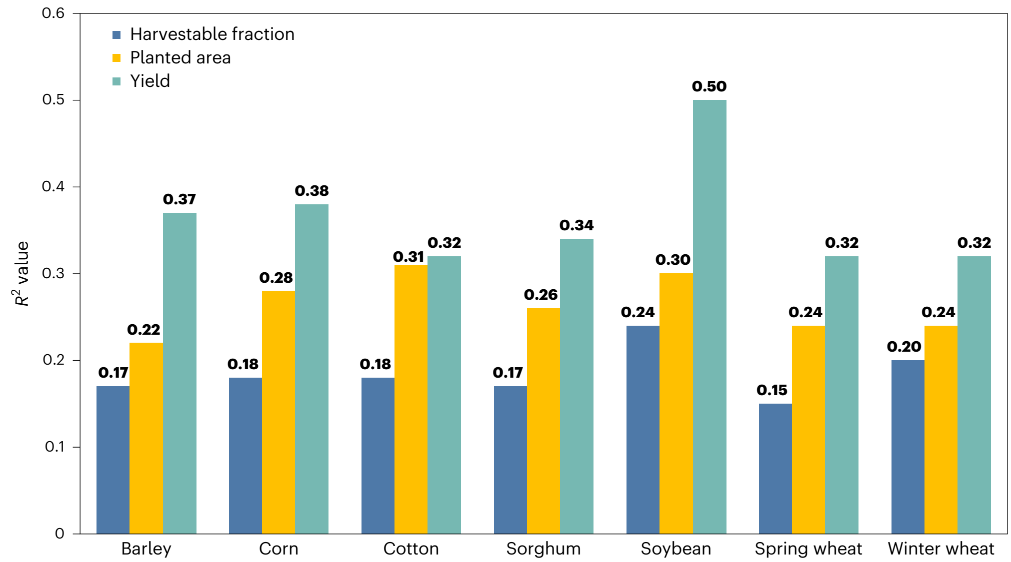



**Fig. 4 | Explained variance from random forest regressions.** Predictor variables are the climate variables listed in Supplementary Table 3. Response variables are the anomalies of the three agricultural components.

because farmers may base their planting decisions partially on the (un) favourability of climate conditions in the previous GP (Supplementary Fig. 5). We found that climate variability and extremes explain between 32% and 50% of yield anomalies, with yields for soybean (50%), corn (38%) and barley (37%) having the highest associations (Fig. 4). For planted area anomalies, we found that climate variables explained between 22% (for barley) and 31% (for cotton) of their variance. We obtained similar results when using climate variables from the previous GP (Supplementary Fig. 5). Finally, for the harvestable fraction, we found that climate variables explained between 15% (spring wheat) and 24% (soybean) of anomalies. The explanatory power of climate variables was highest for yield anomalies, then planted area and lastly harvestable fraction. The relatively low associations with harvestable fraction may be because the indicators for climate extremes considered in this study could not (due to data limitations) include variables of natural disasters (for example, flooding, landslides), which are presumably the most influential factors on harvestable fraction[25] and replanting in the same growing season after disaster. Across all three components and all crops, we found that temperature-related variables ranked highest in importance (Supplementary Table 4). Thus, these findings suggest that all components of production merit consideration as avenues for climate adaptation, particularly with regard to temperature. We note that our approach, which focuses on sudden, short-term reductions probably does not capture the effects of longer-term, persistent climate extremes (for example, multi-year droughts), which can also exercise influence on the levels of crop production.

## Discussion

The rise in climate variability, climate extremes and other disruptions poses a growing threat to the stability of food supply chains. This is especially true for production, which has many entry points for environmental, economic and political disruptions[5,31], now and in the future[32–34]. Responding to these growing disruptions requires a comprehensive view of production and the components that dictate its outcomes. To this end, this study provides new insights into the extent to which three factors—planted area, harvestable fraction and yield—affect production stability and the degree to which they are affected by climate variability and extremes. We found that planted area, harvestable fraction and yield all substantially influence both the frequency and magnitude of production shocks to varying degrees across crops. Considering shock frequency, shocks of area-related components co-occur with at least 50% of production shocks across all crops while yield-related shocks account for more than 31%. Although the effect of area-related components on production shock magnitude is generally lower than yield, we found large effects for certain crops (for example, spring wheat) that deserve particular attention. Further, we found that climate variability and extremes can explain substantial fractions of the observed variations in each of the three components, indicating that there is a combination of complex factors (both climate-related and otherwise) that can contribute to instability in production. Together our results underline the importance of considering all three components to develop holistic approaches to improve production stability and the ability to withstand and recover from disruptions under ongoing climate change. Understanding the reasons behind the crop-to-crop differences in the relative importance of these three components will be an important next step of inquiry towards the development of adaptation strategies.

Addressing food production shocks has direct implications for the entire food supply chain, negatively affecting food supply stability and posing a threat to food security. Limited availability of food can be a direct consequence of a production shock. At the same time, food production instability can dramatically increase food prices when stock is limited, lowering consumer purchasing power and potentially compromising human nutritional status, particularly among lower-income groups[35,36]. Sudden declines in production may also result in a decrease in food stocks; for instance, global grain reserves in 2008 fell to 18% of annual demand[37], which aggravated food system vulnerability. Because countries are becoming more reliant on global food trade, production shocks are affecting not only local markets and consumers but also global and distant markets when shocks cascade through the food trade network[5,38]. Despite international trade increasing the availability and diversity of food[39,40], it also exposes people to external disruptions in food production, particularly in regions that rely heavily on imports[41,42]. For instance, drought and extreme heat in 2012 caused a decline in US agricultural production that subsequently led to increases in global grain prices and compromised food access worldwide, especially for the world's poorest people[43]. This growing interconnectivity of nations means that increasing the stability of major grain producing nations' food production is a promising strategy for protecting global food security.

Our findings demonstrate that efforts are required in all components to stabilize production (with an exclusive focus on yield stability severely constraining the solution space) and that the stability of production is influenced by a variety of factors, including climate variability and extremes to a considerable degree. As such, holistic approaches that account for a variety of potential economic, political and environmental disruptions—and their collective influences on all three components of production—are necessary to truly enhance the stability of crop production. Yield has received the bulk of research and policy attention over the past few decades, with governments, international organizations and other agencies developing cultivars with climate-resilient traits (for example, heat tolerance) as well as practices to reduce the effects of environmental fluctuations on crop yields (for example, agricultural inputs such as irrigation and soil organic matter)[23,44]. But such interventions provide little opportunity for improving the stability of planted areas and harvestable fractions, which are influenced through entirely different mechanisms. For instance, planted area is determined by farmer decisions that are influenced by a host of factors including environmental policies (for example, the US Conservation Reserve Program), market demand and food prices (which enable farmers to select the most profitable crops year after year), and farmer experience in accordance with weather forecasts over time. Economic incentives that account for these various influences can help to avoid sudden shifts in planted areas from year to year. In addition, harvestable fraction is influenced by both extreme events and farmer decisions. While extreme events are difficult to predict, a suite of proactive actions can ameliorate their effects on harvestable fraction, including shifting cropping patterns, adjusting planting times to prevent loss of harvested area caused by environmental disruptions and zoning within cropland to avoid using land with a high probability of experiencing localized extreme events (for example, floods). Meanwhile, strategies that improve crop quality, price and market access could encourage harvesting and thus reduce harvestable fraction losses at the harvest stage.

Stabilizing food production is a growing challenge for agricultural development. Although governments and researchers have worked to increase yield stability[19,45,46], focusing only on yield may miss a variety of important opportunities to stabilize production in the face of disruptions. This is well aligned with recent calls in the sustainability science community to actively design and manage response diversity to a growing suite of disruptions[47]. Our findings reveal that the relative importance of the different components of production varies according to crop. Some crops are grown in a variety of locations throughout the United States (for example, corn, winter wheat), allowing our approach to be applied at regional scales to tailor strategies to local circumstances. As such, developing strategies that employ a suite of interventions targeted at planted area, harvestable fraction and yield offers the greatest flexibility for responding to local vulnerabilities and a variety of potential climatic and non-climatic disruptions.

# Methods

## Data

We relied on United States Department of Agriculture (USDA) survey data for US county-level harvestable fraction, planted area, yield and production for seven field crops, covering 70% of the planted area in the United States[28]. Harvested area is the product of harvestable fraction and planted area. We separated the two components to better disentangle the influence of human and environmental influences on area-related shocks to production, with harvestable fraction more affected by within-season environmental factors, and planted area largely influenced by farmer decisions. Our analysis was limited to crops with available county data that represent 60% or more of national production for 20 consecutive years. The study covered the years 1978 to 2020, which are the years for which data were available across all study crops. It is worth noting that the data for barley did not fully meet our criteria for inclusion after 2014; however, we examined them in the interest of completeness. Data for climatic variables were derived from the PRISM database (https://prism.oregonstate.edu/), which provides high-resolution (4 km) daily and monthly mean, maximum and minimum temperature, and precipitation data for the whole United States from 1981 to 2020. All spatial data were re-gridded to county level by taking an area-weighted average of the grid cells within each county. GP data were derived from the latest USDA survey on usual planting and harvesting dates in 2010[48]. Although climate change has altered sowing dates and crop phenology, we used fixed crop calendars to calculate growing season climate indices[49] because recent observed shifts in planting and harvesting dates have been less than 5 d per 1 °C warming[50]. Using the example of corn, we found that our results were not sensitive to this choice of crop calendar (Supplementary Table 5). GPs were then converted from dates to months to calculate climate variables over all months of each crop's GP. Following Vogel et al.[20], the climate variables calculated in our study included mean monthly temperature, mean monthly precipitation, maximum temperature, minimum temperature, warm day frequency, cold night frequency, maximum 5-d rainfall, diurnal temperature range, frost day frequency, mean 6-month Standardized Precipitation Index (SPI-6) and mean 6-month Standardized Precipitation and Evapotranspiration Index (SPEI-6) (Supplementary Table 3). Climate variability is represented by the first two variables (temperature and precipitation); climate extremes are represented by the others. All county-level agricultural and climate variables were detrended using the singular spectrum analysis method in R to remove temporal trends due to technological progress, management changes and long-term climatic changes. Because climate variables in particular can exhibit distinct temporal trends, detrending prevents the explained variance from being inflated as a result of the regression of two strongly trending variables. Except for SPEI-6 and SPI-6, which are already standardized, all variables were then standardized by dividing by their s.d. to enable the comparison of values across different locations.

## Shock detection

To identify and match the shock occurrence among planted area, harvestable fraction, yield and production in all counties and all crops, we adopted an automated quantitative statistical shock detection method after Gephart et al.[29]. It is a method to capture sudden drops in a time series, with less sensitivity to high variable data and long-term, gradual fluctuations. The process of shock detection is mainly divided into four steps (Supplementary Fig. 6): (1) fitting the time series data by locally weighted scatterplot smoothing regression (red line in Supplementary Fig. 6a) with a span of 2/3; (2) calculating the residuals (that is, the difference between the fitted and actual values; Supplementary Fig. 6b); (3) plotting residuals against the time-lagged residuals (that is, residuals of its previous year; Supplementary Fig. 6c); and (4) using Cook's distance ($D$) to identify extreme points in the regression of residuals versus time-lagged residuals (Supplementary Fig. 6d). Counties with fewer than 20 data points were excluded because of their poor performance in shock detection. Points with Cook's $D$ greater than the $4/n$ ($n$ is the number of data points in a time series) were identified as shocks. While this shock detection method can identify both positive and negative deviations, we only considered production losses (that is, negative production anomalies). Shocks for which the corresponding production data either did not have a value in the previous year or had a value identical to the previous year were not considered because of data irregularities. Using this method, we identified shocks in all counties and all crops for each of the four agricultural variables (for example, planted area, harvestable fraction, yield and production).

To compare the frequency of production shocks to those in the three component factors, we examined whether each of the three components also experienced shocks when a production shock occurred. For example, corn production in Iowa County in Wisconsin had five production shocks over 43 years, and three of them happened in the same year as harvestable fraction shocks. Then, we specified that three harvestable fraction shocks coincide with production shocks (Supplementary Fig. 7). The same approach was used for shocks in yield and planted area. Thus, we determined the number of production shocks that co-occurred with planted area, harvestable fraction and yield shocks. It is worth noting that a production shock can occur in conjunction with shocks in several components, or it may not co-occur at all. Because the shock detection method only captures relatively large drops and ignores gradual fluctuations, and deals with each component of production independently, the two main reasons for no co-occurrences are (1) because the changes in the three components are minor but amplify one another or (2) there is high variability in the time series of one or more of the components and a shock is not statistically detectable. Note that the shock frequency evaluation of our study does not account for differences in the area of each county. Because shocks were counted by county, this may potentially mute the average effect of the component shocks in the counties with larger areas and higher production.

### Shock decomposition

Based on the detected production shocks, we used a decomposition method[30] to measure the contribution of each component to the magnitude of the production shocks. Decomposition follows the index decomposition analysis (IDA)[51] to express the overall change in an aggregate quantity as a sum of contributions from each of its components. The production of each county $i$ is the product of planted area ($A$), harvestable fraction ($F$) and yield ($Y$). We used additive decompositions in the IDA that converted the difference of national production between two consecutive years (equation (1), difference of all counties between year $t$ and $t-1$) into the sum of contributions from each component (equation (2)), by calculating the logarithmic mean Divisia Index (equation (3), example for yield). This approach was applied to every two consecutive years to estimate the contributions of each component to annual national production loss caused by production shocks as:

$$\triangle P_t = P^t - P^{t-1} = \sum_i Y_i^t A_i^t F_i^t - \sum_i Y_i^{t-1} A_i^{t-1} F_i^{t-1} \quad (1)$$

$$\triangle P_t = \triangle Y + \triangle A + \triangle F \quad (2)$$

$$\triangle Y = \sum_i (P_i^t - P_i^{t-1}) / \left(\ln P_i^t - \ln P_i^{t-1}\right) \times \ln\left(\frac{Y_i^t}{Y_i^{t-1}}\right) \quad (3)$$

### Random forest and cross-validation

We applied a random forest machine learning algorithm to evaluate the correlation between each agricultural variable (for example, planted area, harvestable fraction, yield and production) and a suite of climate variables (Supplementary Table 3). Random forests is a non-parametric statistical method that uses decision trees to make regression or classification and is robust to overfitting[52]. This method has been previously applied in the analysis of yield or production anomalies in association with climate variables[20,53].

The random forest model was built to examine the relationship between the anomalies (that is, deviations from an overall trend) of each agricultural variable and all climate indicators for each crop. All data were randomly partitioned into an 80% and 20% split for training and validation. Hyperparameters (that is, the number of trees to build, maximum depth of the tree, minimum leaf node size and number of features to use for splitting) used in each model were tuned based on a grid search approach[54]. To estimate and compare the variance explained by climate for each agricultural variable, we calculated $R^2$ values from cross-validated predictions. We further generated 'variable importance ranks' to assess the relative effect of the climate indicators on each agricultural variable for each crop.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

All raw data are publicly available online. County-level harvested fraction, planted area, yield and production data for the US are available at https://www.nass.usda.gov/Quick_Stats/Lite/index.php. Climatic data from the PRISM database are available at https://prism.oregonstate.edu/. GP data are available at https://www.nass.usda.gov/Publications/Todays_Reports/reports/fcdate10.pdf.

## Code availability

The R code for shock detection was derived from Gephart et al.[29] (https://github.com/jagephart/Shock_Detection). The R code for the decomposition and the Python code for the random forest analysis are available at https://github.com/Dongyang2020/US_Shock.

## References


1. World Food Programme. *Global Report on Food Crises 2022* (GRFC, 2022); https://www.wfp.org/publications/global-report-food-crises-2022
2. Firdaus, R. R., Senevi Gunaratne, M., Rahmat, S. R. & Kamsi, N. S. Does climate change only affect food availability? What else matters? *Cogent Food Agric* **5**, 1707607 (2019).
3. Schmidhuber, J. & Tubiello, F. N. Global food security under climate change. *Proc. Natl Acad. Sci. USA* **104**, 19703–19708 (2007).
4. Food and Agriculture Organization. *Food Security. Policy Brief* (FAO, 2006); https://www.fao.org/fileadmin/templates/faoitaly/documents/pdf/pdf_Food_Security_Cocept_Note.pdf
5. Davis, K. F., Downs, S. & Gephart, J. A. Towards food supply chain resilience to environmental shocks. *Nat. Food* **2**, 54–65 (2021).
6. Wheeler, T. & von Braun, J. Climate change impacts on global food security. *Science* **341**, 508–513 (2013).
7. Lobell, D. B., Schlenker, W. & Costa-Roberts, J. Climate trends and global crop production since 1980. *Science* **333**, 616–620 (2011).
8. Pacetti, T., Caporali, E. & Rulli, M. C. Floods and food security: a method to estimate the effect of inundation on crops availability. *Adv. Water Res.* **110**, 494–504 (2017).
9. Allen, M. R. et al. in *Special Report on Global Warming of 1 °C* (eds Masson-Delmotte, V. et al.) (WMO, 2018).
10. Coumou, D. & Robinson, A. Historic and future increase in the global land area affected by monthly heat extremes. *Environ. Res. Lett.* **8**, 034018 (2013).
11. Jones, A. W. & Phillips, A. Historic food production shocks: quantifying the extremes. *Sustainability* **8**, 427 (2016).
12. Coumou, D. & Rahmstorf, S. A decade of weather extremes. *Nat. Clim. Change* **2**, 491–496 (2012).
13. Robert, M., Thomas, A. & Bergez, J.-E. Processes of adaptation in farm decision-making models. A review. *Agron. Sustain. Dev.* **36**, 64 (2016).
14. Iizumi, T. & Ramankutty, N. How do weather and climate influence cropping area and intensity. *Glob. Food Sec.* **4**, 46–50 (2015).
15. Kotera, A., Nguyen, K. D., Sakamoto, T., Iizumi, T. & Yokozawa, M. A modeling approach for assessing rice cropping cycle affected by flooding, salinity intrusion, and monsoon rains in the Mekong Delta, Vietnam. *Paddy Water Environ.* **12**, 343–354 (2014).
16. Beausang, C., Hall, C. & Toma, L. Food waste and losses in primary production: qualitative insights from horticulture. *Resour. Conserv. Recycl.* **126**, 177–185 (2017).
17. Gunders, D. & Bloom, J. *Wasted: How America Is Losing up to 40 Percent of Its Food from Farm to Fork to Landfill* (Natural Resources Defense Council, 2017).

18. Leng, G. & Hall, J. Crop yield sensitivity of global major agricultural countries to droughts and the projected changes in the future. *Sci. Total Environ.* **654**, 811–821 (2019).
19. Renard, D. & Tilman, D. National food production stabilized by crop diversity. *Nature* **571**, 257–260 (2019).
20. Vogel, E. et al. The effects of climate extremes on global agricultural yields. *Environ. Res. Lett.* **14**, 054010 (2019).
21. Zampieri, M., Weissteiner, C., Grizzetti, B., Toreti, A., van den Berg, M. & Dentener, F. Estimating resilience of crop production systems: from theory to practice. *Sci Total Environ.* **735**, 139378 (2020).
22. Zipper, S. C., Qiu, J. & Kucharik, C. J. Drought effects on US maize and soybean production: spatiotemporal patterns and historical changes. *Environ. Res. Lett.* **11**, 094021 (2016).
23. Ray, D. K., Gerber, J. S., MacDonald, G. K. & West, P. C. Climate variation explains a third of global crop yield variability. *Nat. Commun.* **6**, 5989 (2015).
24. Rezaei, E. E., Ghazaryan, G., Moradi, R., Dubovyk, O. & Siebert, S. Crop harvested area, not yield, drives variability in crop production in Iran. *Environ. Res. Lett.* **16**, 064058 (2021).
25. Lesk, C., Rowhani, P. & Ramankutty, N. Influence of extreme weather disasters on global crop production. *Nature* **529**, 84–87 (2016).
26. Cohn, A. S., VanWey, L. K., Spera, S. A. & Mustard, J. F. Cropping frequency and area response to climate variability can exceed yield response. *Nat. Clim. Change* **6**, 601–604 (2016).
27. Food and Agriculture Organization of the United Nations. *FAOSTAT Statistical Database* (FAO, 2022).
28. United States Department of Agriculture National Agricultural Statistics Service. *Quick Stats Database* (USDA NASS, 2022); https://www.nass.usda.gov/Quick_Stats/Lite/index.php
29. Gephart, J. A., Deutsch, L., Pace, M. L., Troell, M. & Seekell, D. A. Shocks to fish production: identification, trends, and consequences. *Glob. Environ. Change* **42**, 24–32 (2017).
30. Ang, B. W. LMDI decomposition approach: a guide for implementation. *Energy Policy* **86**, 233–238 (2015).
31. Savary, S. et al. Mapping disruption and resilience mechanisms in food systems. *Food Secur.* **12**, 695–717 (2020).
32. Bebber, D. P. & Gurr, S. J. Crop-destroying fungal and oomycete pathogens challenge food security. *Fungal Genet. Biol.* **74**, 62–64 (2015).
33. Xia, L. & Robock, A. Impacts of a nuclear war in South Asia on rice production in mainland China. *Clim. Change* **116**, 357–372 (2013).
34. Aday, S. & Aday, M. S. Impact of COVID-19 on the food supply chain. *Food Qual. Saf.* **4**, 167–180 (2020).
35. Brinkman, H.-J., de Pee, S., Sanogo, I., Subran, L. & Bloem, M. W. High food prices and the global financial crisis have reduced access to nutritious food and worsened nutritional status and health. *J. Nutr.* **140**, 153S–161S (2010).
36. De Schutter, O. *Report of the Special Rapporteur on the Right to Food. Final Report: the Transformative Potential of the Right to Food (A/HRC/25/57)* (United Nations, General Assembly, Human Rights Council, 2014).
37. Fraser, E. D. G., Legwegoh, A. & Krishna, K. C. Food stocks and grain reserves: evaluating whether storing food creates resilient food systems. *J. Environ. Stud. Sci.* **5**, 445–458 (2015).
38. Heslin, A. et al. Simulating the cascading effects of an extreme agricultural production shock: global implications of a contemporary US dust bowl event. *Front. Sustain. Food Syst.* https://doi.org/10.3389/fsufs.2020.00026 (2020).
39. Porkka, M., Kummu, M., Siebert, S. & Varis, O. From food insufficiency towards trade dependency: a historical analysis of global food availability. *PLoS ONE* **8**, e82714 (2013).
40. Kearney, J. Food consumption trends and drivers. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **365**, 2793–2807 (2010).
41. Marchand, P. et al. Reserves and trade jointly determine exposure to food supply shocks. *Environ. Res. Lett.* **11**, 095009 (2016).
42. Suweis, S., Carr, J. A., Maritan, A., Rinaldo, A. & D'Odorico, P. Resilience and reactivity of global food security. *Proc. Natl Acad. Sci. USA* **112**, 6902–6907 (2015).
43. Boyer, J. S. et al. The U.S. drought of 2012 in perspective: a call to action. *Glob. Food Secur.* **2**, 139–143 (2013).
44. Herrero, M. et al. Innovation can accelerate the transition towards a sustainable food system. *Nat. Food* **1**, 266–272 (2020).
45. Bowles, T. M. et al. Long-term evidence shows that crop-rotation diversification increases agricultural resilience to adverse growing conditions in North America. *One Earth* **2**, 284–293 (2020).
46. Mahaut, L., Violle, C. & Renard, D. Complementary mechanisms stabilize national food production. *Sci. Rep.* **11**, 4922 (2021).
47. Walker, B. et al. Response diversity as a sustainability strategy. *Nat. Sustain.* https://doi.org/10.1038/s41893-022-01048-7 (2023).
48. *Field Crops: Usual Planting and Harvesting Dates. Agricultural Handbook 628* (USDA National Agricultural Statistics Service, 2010).
49. Guilpart, N., Iizumi, T. & Makowski, D. Data-driven projections suggest large opportunities to improve Europe's soybean self-sufficiency under climate change. *Nat. Food* **3**, 255–265 (2022).
50. Yang, Y. et al. Characterizing spatiotemporal patterns of crop phenology across North America during 2000–2016 using satellite imagery and agricultural survey data. *ISPRS J. Photogramm. Remote Sens.* **170**, 156–173 (2020).
51. Ang, B. W., Huang, H. C. & Mu, A. Properties and linkages of some index decomposition analysis methods. *Energy Policy* **37**, 4624–4632 (2009).
52. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
53. Hoffman, A. L., Kemanian, A. R. & Forest, C. E. The response of maize, sorghum, and soybean yield to growing-phase climate revealed with machine learning. *Environ. Res. Lett.* **15**, 094013 (2020).
54. Bergstra, J. & Bengio, Y. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* **13**, 281–305 (2012).


## Acknowledgements

This study was supported in part by the Gerard J. Mangone Climate Change Science & Policy Hub at the University of Delaware. K.F.D. acknowledges support from a USDA National Institute of Food and Agriculture grant no. 2022-67019-37180 and University of Delaware General University Research Fund.

## Author contributions

D.W. and K.F.D. conceptualized the study. D.W., J.A.G., T.I., N.R. and K.F.D. devised the methodology. D.W. validated and visualized the data. D.W. and K.F.D. wrote the original manuscript draft. D.W., J.A.G., T.I., N.R. and K.F.D. reviewed and edited the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s41893-023-01152-2.

**Correspondence and requests for materials** should be addressed to Dongyang Wei.

https://doi.org/10.1038/s41893-023-01152-2

**Peer review information** *Nature Sustainability* thanks the anonymous reviewers for their contribution to the peer review of this work.

**Reprints and permissions information** is available at www.nature.com/reprints.

nature portfolio

Corresponding author(s): Dongyang Wei

Last updated by author(s): May 9, 2023

# Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☒ | ☐ | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☒ | ☐ | A description of all covariates tested |
| ☐ | ☒ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ | For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☒ | ☐ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒ | ☐ | Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |

*Our web collection on statistics for biologists contains articles on many of the points above.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | No software was used. |
| Data analysis | The data analyses were performed using R 4.2.1 (shock detection and decomposition) and Python 3.11.1 (random forests). Details were reported in the section of the Methods. R code for shock detection is derived from Gephart et al. (https://github.com/jagephart/Shock_Detection). R code for decomposition and Python code for Random Forest analysis are available at https://github.com/Dongyang2020/US_Shock. |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All underlying raw data are publicly available online. County-level harvestable fraction, planted area, yield, and production data for the US are available at https://

www.nass.usda.gov/Quick_Stats/Lite/index.php. Climatic data from PRISM database are available at https://prism.oregonstate.edu/. Growing period data are available at https://www.nass.usda.gov/Publications/Todays_Reports/reports/fcdate10.pdf.

## Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

| | |
|---|---|
| Reporting on sex and gender | N/A |
| Population characteristics | N/A |
| Recruitment | N/A |
| Ethics oversight | N/A |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

# Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Study description | We use county-level data on seven major crops (barley, corn, cotton, sorghum, soybeans, spring wheat, and winter wheat) in the United States to evaluate the relative importance of disruptions in the three components of production (i.e., planted area, harvestable fraction (i.e., the ratio of planted area to harvested area), and yield) to production instability. |
| Research sample | N/A |
| Sampling strategy | N/A |
| Data collection | N/A |
| Timing and spatial scale | Annual, county-level data covering the years 1977-2020 |
| Data exclusions | N/A |
| Reproducibility | All data are publicly available to replicate the analysis |
| Randomization | N/A |
| Blinding | N/A |

Did the study involve field work? ☐ Yes ☒ No

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials & experimental systems

| n/a | Involved in the study | |
|---|---|---|
| ☒ | ☐ | Antibodies |
| ☒ | ☐ | Eukaryotic cell lines |
| ☒ | ☐ | Palaeontology and archaeology |
| ☒ | ☐ | Animals and other organisms |
| ☒ | ☐ | Clinical data |
| ☒ | ☐ | Dual use research of concern |

## Methods

| n/a | Involved in the study | |
|---|---|---|
| ☒ | ☐ | ChIP-seq |
| ☒ | ☐ | Flow cytometry |
| ☒ | ☐ | MRI-based neuroimaging |



March 2021