We would like to thank the reviewers and the editor for their helpful and constructive comments on the manuscript “The Largest Crop Production Shocks: Magnitude, Causes and Frequency ”.

We found all of the feedback to be useful in improving and sharpening the research. We have now updated the manuscript to address the comments. We believe it has been significantly strengthened and is now suitable for publication. Below we have listed the reviewer and editor comments in black, along with our responses in light green. Text that has been added to the manuscript is in *italics and darker green*.

Reviewer #1

Thanks to the authors for their diligent revisions. I am satisfied with most of the revisions and/or responses to my comments from the first round. I have just a few more comments, mostly a conversational response, with some recommendations for minor revisions. Will be good to see your paper published.

We thank the reviewer for their positive assessment.

Comments

1. I was originally still not entirely convinced by the argument that humans could eat crops intended for livestock feed, but granted I am not the expert on this. When I first made that comment, I was thinking of the #2 yellow corn predominantly grown in the US for livestock feed, which is a different variety from the sweet corn that people normally eat. But having done some more reading, it does seem like one could eat #2 corn in situations such as famine (e.g., https://tinyurl.com/47e85wk9; feel free to add this or other citations to make your argument stronger). I also tried to review the citations you added to support your argument that livestock feed has historically been diverted to human consumption during times of crises. Two of your citations (Collingham, Offer) are books I could not easily access. But I was able to quickly review the Meng et al. (2015) paper, which does say that peasants in rural China ate green crops (illegally) from the fields. So I accept this point. And besides, you have added a sensitivity analysis by removing cotton, soy and rapeseed.

Thank you for engaging with this point and for the additional reference on #2 corn during famine conditions. We appreciate that the Meng et al. (2015) evidence and our sensitivity analysis address your initial concern.

2. I would suggest adding a short sentence (either to lines 213-215 or right after) saying that the spearman correlation also accounts for different production magnitudes of countries.

Changed as proposed.

3. Regarding pests & diseases, I see your point. But there is a potentially important message here that you could touch upon in your discussion if you wish to. It is widely believed that pests and diseases are a major source of crop losses (e.g., see https://doi.org/10.1038/s41559-018-0793-y). This is in fact why farmers apply a lot of pesticides and why a lot of the GM crops are focussed on traits to reduce pest burden. But while the burden of pests and disease for crop production may be high, maybe it just remains at a high baseline without too much year-to-year fluctuation, and therefore does not create the types of episodic shocks that you are examining in this paper. In other words, some drivers are of high magnitude but without large fluctuations, and therefore we don’t see their effects in terms of shocks.

We thank the reviewer for their helpful point here and added the following description to the discussion of pests and diseases:

Pests and diseases are often one of the largest sources of crop losses (Savary et al., 2019). However, given that we do not find them here as one of the main causes of the largest shocks, this implies that they cause damage on a high magnitude but without large fluctuations.

Reviewer #2

Thank you for the responses and the changes. I still find it hard difficult to distinguish novel fatures of this article as it stands. Of the three objectives listed in the introduction (L83) my impression is that these already exist in the literature, so this paper needs to argumentatively and/or technically assess a bit more.

I) I understand that the paper aims to focus on "the worst production shock" L78 and therefore aggregates, however, certain crops are particularly important for certain regions, and temporal dimensions also crucial i.e. alternating cropping seasons across hemispheres, as well as particular trade routes of particular crops. Furthermore, despite the justifcation on 'worst' shocks, the paper then goes on to describe production/shock distributions and coutry profiles, making the analysis extremely similar to Cottrell.

Furthermore, compared to Cottrell 2019 for instance, this paper is not only more aggregated, but actually covers less data, as Cottrell disaggregate between livestock, crops but also marine sources of food, which especially for small island states which this paper highlights crop shocks in, can be quite improtant. The famine framing allows this paper to assume all diversion of feed to direct consumption, but caloric content and calories available for human consumption can end up quite different (although I would disagree with soybean being one of the more difficult ones), and livestock may still play some roles in such cases, in any case, it is a strong simplification that already existing studies actually did not make. Especially as global disaster events could imaginably onset suddenly, adaptations of processing sectors may not happen very fast to make feed crops consumable to the general population.

We would like to thank the Reviewer for their perspective; however, we disagree that our paper does not bring novelty to the existing literature. The key distinction is that our study aggregates crop production by calories at the country level. In addition our focus is the most severe shocks to reflect the full historical spread of worst case calorie reduction. Calories rather than tonnes provide a more meaningful measure of food availability for human consumption, as calorie content differs widely between different foods and therefore covers actual human energy requirements better than tonnes.

Cottrell et al. analyze food shocks in general, while we are looking at the worst ones. It is important to note that the general distribution of shocks and the worst shocks are not caused by the exact same kind of events or concentrated in the same areas. In fact, we find clear differences such as Southern African countries experiencing the largest shocks in our analysis and South Asian countries experiencing the most frequent shocks in Cottrell et al. We have now noted this difference in section 4.2. Additionally, our analysis presents many insightful visualizations that were not present in Cottrell et al., like maps showing where to expect the largest shocks, a comparison of the magnitude of shocks, recurrence rates of shocks of different sizes and the synchronization analysis. In summary, Cottrell et al. describe **how often** food systems fluctuate; this paper asks **how bad it can get** when they fail. Therefore, we argue that both analyses, Cottrell et al and ours is a very worthwhile undertaking. We have outlined this explicitly now in our manuscript that these two studies complement each other.

This study here complements Cottrell et al. (2019). While the earlier study focussed on how often the food system shows shocks in general, this study here explicitly focuses on how bad these shocks can get and why these most extreme shocks happen.

In the first round of revisions, we have added a sensitivity analysis excluding crops most difficult to convert to human food (cotton, soy, rapeseed) showing robust results to our original analysis (see Figure S1). In addition, history has proven that feed-to-food conversion during crises has occurred several times and is well-documented during both World Wars and the Great Chinese Famine (already discussed in section 2.1)

We acknowledge that different crops vary in importance across regions. However, our analysis already includes all major crop types, which are aggregated consistently into a calorie-based measure of food availability. As a result, regionally dominant crops necessarily drive the aggregated calorie signal: a substantial decline in total calories can only occur if the crops that are most important in a given region experience a significant shock. In this sense, the aggregation preserves regional crop relevance rather than obscuring it.

This study is based on FAOSTAT data, which are available at annual temporal resolution. Consequently, we are unable to explicitly resolve seasonal cropping patterns or alternating growing seasons across hemispheres. While this is a limitation of the available data, annual aggregation remains appropriate for our focus on the worst historical calorie shocks, which by definition reflect cumulative production losses at the country level.

Lastly, we acknowledge that Cottrell et al. includes sources besides crops; however, our paper is explicitly focused on crop production shocks, as stated in the title. Also note that a large majority (82%, https://www.nature.com/articles/s43016-022-00573-0) of calories consumed by humans come directly from crops.

II) Regarding the synchrony analysis, how do the correlation coefficients change over time i.e. could analyse with a moving window, and see how the (a)synchrony that cannot be qualitatively ascribed to climate factors match with conflict or policy or other aspects? Or does synchrony change with different crops, northern hemisphere maize could be compensated by south american production for instance, but perhaps wheat is harder to do so?

Our analysis is meant to describe the synchronization across the whole time period. Nevertheless, we added an additional Figure S7 to showcase that there are changes in the synchronization over time, but that many of the broad patterns stay consistent.

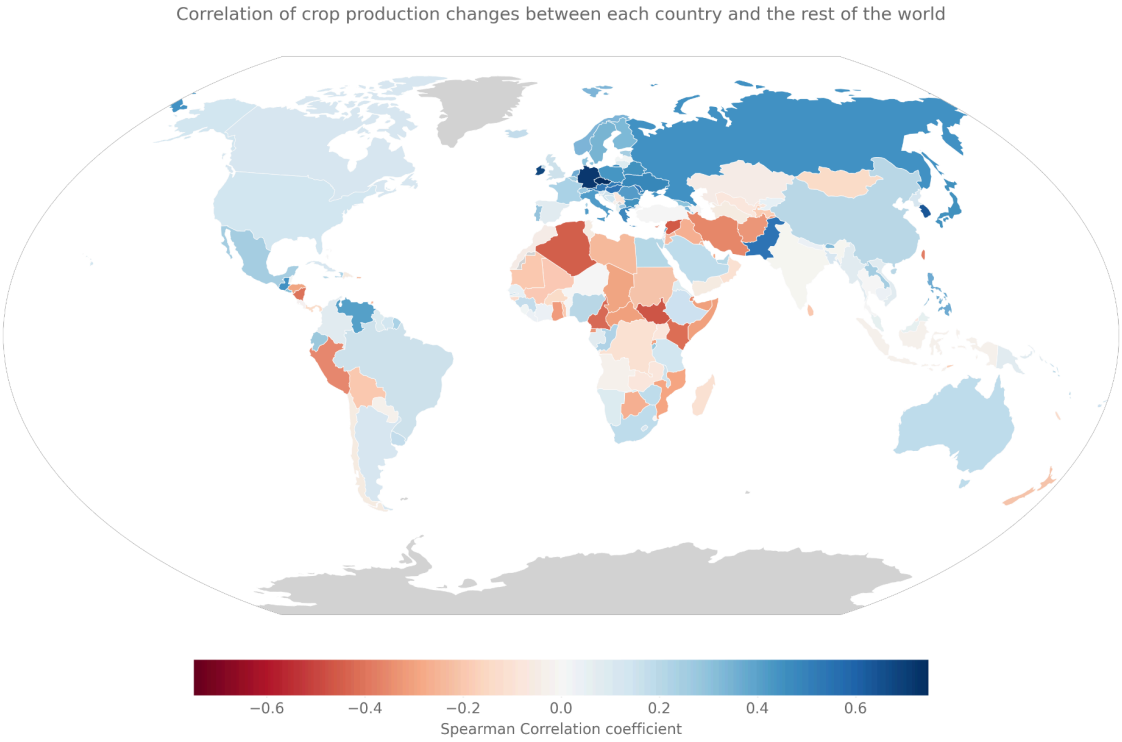

Figure S7: Figure 8: Correlation of crop production changes between each country and the rest of the world for the years 2003-2023. A positive correlation (blue) means a country's crop production tends to move in the same direction as global production. A negative correlation (red) means the country's production tends to move opposite to global trends.

We have also extended the results section to mention this:

These patterns of synchronization mostly stay consistent over time, but there are changes. For example, if we only look at the data from 2003 to 2023 (Figure S7), we can see that Brazil's relatively strong negative correlation becomes positive, while Europe's generally positive correlation and Africa's generally negative correlation persist. This likely reflects changes in agricultural practices and the dominance of certain regions when it comes to crop production.

We further did the analysis for wheat only and also extended the results section for this:

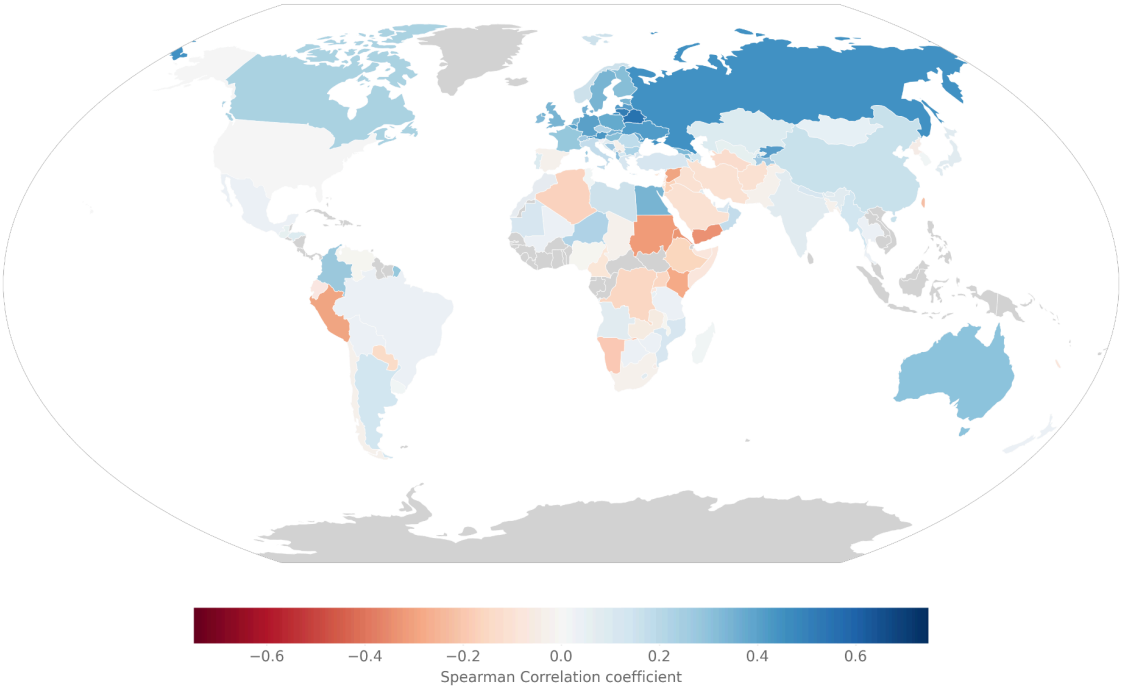

Figure S8: Figure 8: Correlation of only wheat production changes between each country and the rest of the world for the years 1961-2023. A positive correlation (blue) means a country's crop production tends to move in the same direction as global production. A negative correlation (red) means the country's production tends to move opposite to global trends. Light grey indicates no data.

Much of the synchronicity in global crop production is driven by wheat, which makes up a substantial share of total output (Figure S8). This means wheat's year-to-year variation can overshadow more localized patterns in other crops. For our purposes, this is not a limitation—we aim to understand global crop production as a whole and identify which regions might serve as buffers when others fail. The wheat-driven pattern shifts dominance toward major wheat exporters like Russia and Ukraine, whose production swings carry outsized weight in global totals. The United States, despite being a major wheat producer, shows weaker correlation with global trends. Possibly because its production variability operates independently of the factors driving Eurasian wheat yields, as both continents experience different climate impacts.

III) Regarding the qualitative aspect, is there any indication that an AI-assisted search provides more systematic answers than a human-based one, and in which ways? Access to untranslated documents perhaps? but this needs to be validated, and/or mentioned, beyond the fact that it certainly saves time for the authours.

While we do not claim an AI-assisted approach is superior to manual search per se, it is efficient and adequate for our purposes while, at the same time, allowing for higher coverage than humans could. Specifically, our validation workflow was as follows: Claude was used to surface potentially relevant documents, but *all* attributions were verified by human researchers reading the original sources. Sources returned include phrases such as "worst drought year since the mid-15th century" or "most violent and bloody period of the entire armed confrontation," suggesting our method successfully identifies extreme

events. For cases where AI did not surface any relevant information, we conducted additional manual searches. This methodology is described in detail in section 2.2.

Finally, I don't necessarily think closing the loop from driver to shock to (qualitative)human impact is beyond the scope of this analysis, especially given the framing of this paper on 'worst', human disaster, catastrophe etc, and how it aims to 'ground' such framings in historical data. I think it's extremely pertinent to then see if, given the 'worst' shock a country has faced in quantified history, how human impacts appeared, i.e. via price shocks, distributional impacts (on producers vs urban poor for instance), shifts in consumption, etc? And if these did not appear, what mechanisms allowed countries to avoid such impacts, i.e. was storage or trade particularly important under certain policy constraints. This would also allow for providing specific understandings of possible recommendations beyond the generalized ones given in section 4.4, would strenghten the framing and novelty of the paper much more beyond repeating an existing analysis.

We agree that tracing pathways from production shocks through price effects, distributional impacts, and consumption shifts would be valuable. However, this would constitute a substantial research programme in its own right— requiring detailed case studies for each of the country-level shocks identified. This is beyond the scope of a revision and represents a direction for future work, which we now note in Section 4.4:

The dataset produced by this study opens several avenues for future research. First, tracking not only the largest shock but all detectable shocks for each country would reveal how sequential or compound events interact—for instance, whether mismanagement in one year amplifies vulnerability to drought the next. Second, and perhaps most policy-relevant, would be systematic case studies tracing each major shock from production loss through to human welfare outcomes. Key questions include: How did prices respond? Did trade partners maintain exports or impose bans? Which population groups bore the burden? What interventions (if any) mitigated impacts? Answering these questions would substantially improve our understanding of food system resilience and the conditions under which production shocks become humanitarian crises.

Our paper's contribution is to systematically identify the largest shocks and explain why they occurred. Understanding how societies responded to each shock is an important but distinct research question.