Response to reviewer 1

We thank you for your careful and constructive review. We have addressed each of your comments below. The manuscript has been revised accordingly; changes in the text are noted by section and line numbers when applicable.

1. **Reviewer Comment:** A central assumption of the study is that the IOB-yield relationship reflects a signal that is distinct from ENSO. The manuscript notes that Gram-Schmidt orthogonalization was applied to remove the influence of Niño3.4 from the IOB index, which is a reasonable approach. However, given the apparent Pacific SST anomalies shown in Fig. 6, it would be helpful to clarify the extent to which the IOB signal is statistically independent of ENSO. Was the orthogonalization applied only to the IOB index, or also to the meteorological predictors used in the regression (e.g., Tmax, SMroot)? Could the Pacific anomalies still reflect residual ENSO influence? Since ENSO and Indian Ocean warming often co-evolve, further clarification would be helpful. Aconditional correlation or partial regression analysis (yield vs IOB, controlling for ENSO) would more directly test their statistical independence. If such analyses were not feasible due to sample size or other constraints, a short note acknowledging this would suffice.

Response: In our analyses, we applied Gram-Schmidt orthogonalization specifically when examining (i) the correlations between climate indices and U.S. soybean yield, and (ii) the relationships between climate indices, meteorological factors, and atmospheric circulation fields. In both cases, we removed the linear component associated with Niño3.4 from each target variable. In other words, the IOB index, other climate indices, meteorological predictors (e.g., Tmx, SMroot, Pre, VPD), and large-scale circulation variables (e.g., SLP, GPH200, and 925 hPa winds) were all orthogonalized against Niño3.4 before entering the correlation and regression analyses. This ensured that the variability we attributed to the Indian Ocean was linearly independent of ENSO at the same time step.

We acknowledge, however, that this approach only guarantees independence at zero lag. Because ENSO and Indian Ocean warming often co-evolve and can exert lead-lag influences on each other across seasons, some residual ENSO-related effects may remain after orthogonalization. We have now added a statement in Section 2.2 to make this limitation explicit. Importantly, the IOB-yield relationship remains significant even after orthogonalization, suggesting that the IOB contributes predictive information beyond ENSO co-variability.

Manuscript changes: In Section 2.2 (Lines 97-99), we revised the text as follows:

B Before conducting the specific analyses, we employed Gram-Schmidt orthogonalization to remove the linear influence of ENSO (represented by Niño3.4) from the IOB index, other climate indices, meteorological factors, and large-scale circulation fields. This method transforms correlated variables into orthogonal sets by sequentially projecting each target variable onto the space orthogonal to ENSO. The ENSO-independent component of a variable X was calculated as:

$$\mathrm{X}_{\perp \mathrm{E}} = \mathrm{X} - \left(\frac{\langle \mathrm{X}, \mathrm{E} \rangle}{\langle \mathrm{E}, \mathrm{E} \rangle} \right) \mathrm{E} \tag{5}$$

Where X is the original variable, E the ENSO signal, and $\langle \cdot, \cdot \rangle$ denotes the inner product. Through this procedure, only the variability linearly independent of ENSO is retained, enabling a clearer attribution of Indian Ocean-related effects. We note that while this approach ensures zero-lag statistical independence from ENSO, lead-lag influences cannot be fully eliminated, as ENSO and Indian Ocean warming often co-evolve and interact across seasons. Similar approaches have been applied in recent climate studies (Hou et al., 2024).

2. **Reviewer Comment:** In Fig. 3, Tmax and SMroot emerge as the most influential predictors of soybean yield anomalies. Given that dry soils can lead to elevated Tmax via reduced evaporative cooling, these two variables are often physically and statistically linked. This raises the question of whether they contribute independent information to the regression model or reflect overlapping aspects of the same underlying drought process. Have the authors assessed their correlation or examined variance inflation among predictors? Even a brief note on whether these variables act jointly or additively would help clarify their interpretation within the ridge regression framework.

Response: We assessed the correlation between Tmax and SMroot across the U.S. soybean production regions and found a mean Pearson correlation coefficient of -0.42, indicating moderate association rather than strong collinearity. Ridge regression, by design, mitigates multicollinearity through coefficient shrinkage, allowing both predictors to retain meaningful contributions. Tmax primarily captures atmospheric heat stress, whereas SMroot reflects water availability, making them complementary indicators of drought impacts. We have added a clarification in the Discussion to emphasize this interpretation.

Manuscript change: Ridge regression identifies Tmax and SMroot as the most influential meteorological variables for soybean yield anomalies. Tmax primarily reflects atmospheric heat stress, while SMroot measures subsurface water availability and drought persistence. These variables are physically interdependent: dry soils reduce evaporative cooling, increasing Tmax, while higher Tmax enhances evapotranspiration, further reducing SMroot and creating a reinforcing feedback. Despite this interdependence, ridge regression mitigates multicollinearity through coefficient shrinkage, allowing both variables to retain meaningful contributions and jointly characterize compound climate risks. Many other climate variables affecting crop growth also interact, such as soil moisture, precipitation, and temperature, which together determine the water and heat conditions experienced by crops. These complex interactions make yield variability difficult to attribute to any single factor and highlight the importance of modeling approaches that explicitly capture interacting and reinforcing climatic influences on crop production.

3. **Reviewer Comment:** The persistence of soil moisture from winter to summer is a key element of the proposed mechanism. Supplementary Fig. S3 appears to illustrate this, but the discussion could benefit from making more use of it. Would the authors consider highlighting which regions show the strongest ND(-1)J-JAS soil moisture correlation? A short mention in the main text would help readers better understand the spatial aspects of this memory effect..

Response: In the original manuscript, we only referenced Supplementary Fig. S3 without explicitly describing the key regions where soil moisture persistence is strongest. In the revised version, we have expanded this section of the Results to highlight that the strongest positive correlations between ND(-1)J (November-December of the year preceding harvest and January of the harvest year) and JAS (July-August-September) soil moisture anomalies are primarily concentrated in the U.S. Midwest and Northern Plains regions that represent the core soybean production areas in the United States. This revision strengthens the connection between early-season hydrological conditions and subsequent summer drought risk, emphasizing the "memory effect" of soil moisture on climate-crop interactions.

Manuscript change: In Section 3.3, we revised the paragraph:

In this study, we found that soil moisture anomalies during ND(-1)J are significantly positively correlated with soil moisture anomalies in JAS (Supplementary Fig. S3). Specifically, the strongest positive correlations are observed over the U.S. Midwest and Northern Plains, which represent the core soybean production regions, indicating that early-season soil moisture conditions in these areas persist into the summer. This "memory effect" suggests that reductions in ND(-1)J soil moisture can influence root-zone moisture during the reproductive stage, thereby affecting soybean yields and highlighting the role of early-season hydrological conditions in modulating summer drought risk.

4. **Reviewer Comment:** The use of a 5-year running mean to detrend soybean yield is a standard choice to remove technological and management-related trends. However, this method may also suppress low-frequency climate variability, such as decadal SST modes, and reduce the number of effective degrees of freedom. Was the sensitivity of the results to this detrending method evaluated? For example, how do key correlations or regression outcomes compare when using a linear detrending approach instead? A brief justification for selecting the 5-year running mean, or a short note on whether this choice meaningfully affects the results, would help readers assess the robustness of the teleconnection signal.

Response: We conducted a sensitivity analysis using linear detrending of soybean yield anomalies instead of the 5-year running mean. The correlation with the IOB index changed from -0.41 to -0.389, and the spatial regression results remained largely consistent, albeit with slightly different amplitude. Specifically, the linear detrending yields significant responses in Nebraska, Oklahoma, Mississippi, and Alabama, while the 5-year running means show stronger signals in South Dakota and New York. These differences are modest and do not alter the main conclusions about the IOB's teleconnection pattern. We chose the 5-year running mean as it effectively removes long-term agronomic and technological trends while retaining interannual climate variability; this method is widely used in crop-climate studies (Iizumi et al., 2021)

We are also prepared to include the linear detrended regression map as a Supplementary Figure if the reviewer believes it would strengthen the presentation.

5. **Reviewer Comment:** Pearson correlation is used extensively throughout the manuscript to assess relationships among SST indices, meteorological variables, and anomalies in soybean yield. While this is a standard approach, Pearson correlation assumes linearity and normality, and can be sensitive to outliers. Were these assumptions checked in the analysis? For key relationships such as IOB-yield or SMroot-yield, would the results be consistent if Spearman rank correlation were used instead? Even a brief mention of this in the methods or supplement would help confirm the robustness of the reported associations.

Response: To ensure that our conclusions are not biased by the assumptions of Pearson correlation (linearity, normality, and sensitivity to outliers), we repeated the key analyses using Spearman rank correlation, which is non-parametric and more robust to outliers. The results were highly consistent: for instance, the correlation between the ND(-1)J IOB index and U.S. soybean yield anomalies was -0.38 (p=0.017) with Spearman correlation, closely matching the Pearson result (-0.41, p=0.0098). These results confirm that the reported teleconnection is not sensitive to the choice of correlation metric.

A supplementary figure summarizing the Spearman results can be provided upon request if the editor deems it useful.

6. **Reviewer Comment:** The manuscript describes a compelling multi-step pathway: IOB warming leads to changes in atmospheric circulation, reduced soil moisture, increased summer heat and drought, and ultimately, yield loss. Would the authors consider adding a simple schematic to summarize this mechanism? This could help readers from interdisciplinary fields quickly grasp the whole story.

Response: We agree that a schematic would be valuable and will add it as a new figure in the revised manuscript to illustrate the multi-step mechanism linking IOB anomalies to U.S. soybean yield variability.

Specific comments

L59. “food securety” → “food security”.

We have revised this word.

L66. “Political units” could be ambiguous to international readers. Please specify that this refers to U.S. states.

To avoid ambiguity, we revised the wording in the text to explicitly state “U.S. states” instead of “political units.”

L97. The Gram-Schmidt procedure is mentioned but not described in detail. Clarify whether Niño3.4 was regressed from IOB or vice versa and consider including a short equation or citing a standard reference.

We have expanded the Methods (Section 2.2).

L126. When selecting ND(-1)J as the optimal window, indicate whether a formal selection criterion (e.g., max correlation, statistical threshold) or multiple testing adjustment was applied.

We have clarified our selection procedure. The ND(-1)J period was chosen because it showed the highest absolute Pearson correlation with U.S. soybean yield anomalies among all tested 3-month windows. We did not apply multiple testing corrections because the selection step served as a screening procedure rather than a formal hypothesis test. This clarification has been added to the Methods.

Manuscript change: We revised the sentence as:

“The year-to-year anomalies in soybean yield exhibit the strongest Pearson correlation (-0.41) with the IOB index during ND(-1)J (November and December of the year preceding harvest and January of the harvest year), which was identified as the optimal 3-month window following an exhaustive correlation screening across all possible periods; this relationship is statistically significant at the 99% confidence level (two-tailed t-test; Fig. 1).”

L201-209. Please update figure references to follow the standard format:

L201. “Fig. 4(b) and 4(c)” → “Figs. 4(b) and 4(c)”

L208. “Fig. 4(b) and 4(e)” → “Figs. 4(b) and 4(e)”

L209. “Fig. 4(c)-(e)” → “Figs. 4(c)- (e)”

L255. The text refers to “Fig. 6(g)”, but the panels go only to (f). This should be corrected.

We have revised these figure references.

Units and labeling: Colorbars in Figs. 3-6 should include clear units (e.g., “% per σ” or “°C per σ”). Consistent labeling will improve readability.

The units of the colorbars are already specified in the figure captions. If the reviewer considers it necessary, we are happy to additionally include the units directly on the colorbars in the revised figures to further improve clarity and consistency.