Response to Reviewer #1 Comments

Thank you very much for your valuable comments on our manuscript. These comments are highly appreciated and have been very helpful in improving the quality of our manuscript and strengthening the overall research.

General Comments:

Comment 1: Cropland mask is too coarse and does not isolate crop signals

The study masks non-cropland grids using the MODIS MCD12C1 land-cover product (0.05° resolution). This spatial resolution is insufficient for isolating crop vegetation signals.

Within "cropland" pixels in the NCP, substantial non-crop components remain, including shelterbelts and windbreak trees, orchards and agroforestry vegetation, grasses and weeds, rural settlements and impervious surfaces, bare soil and fallow land, etc.

Because NDVI and GPP respond to all photosynthetically active vegetation, these components may contribute significantly to the observed anomalies. Consequently, the reported NDVI/GPP signals cannot be confidently attributed to crop growth.

Higher-resolution land-cover products (10–30 m) are available for NCP and should be used to better isolate cropland signals.

Moreover, all datasets were resampled to 0.1° (~10 km) resolution to match the climate data. This coarse aggregation is not strictly necessary. NDVI and productivity metrics could be derived from higher-resolution sensors (e.g., Landsat or Sentinel), which would better represent field-scale variability.

At the 0.1° scale, even when retaining only cropland pixels (from the 0.05° mask), each grid cell still contains heterogeneous land cover and management conditions. The vegetation signal therefore represents **landscape-level vegetation activity**, rather than crop-specific responses.

Reply: Thank you very much for this insightful and constructive comment. The issue of sub-grid heterogeneity and signal contamination is indeed a common challenge in regional-scale studies, especially in a landscape as complex as the North China Plain (NCP). We have carefully considered your suggestions and made the following improvements to our analysis:

First, we agree that the previous 0.05° cropland mask could be improved. To address this, we have now incorporated the 10 m resolution Dynamic World V1 dataset from Sentinel-2 imagery (Fig. R1). Instead of simply masking, we used this high-resolution product to calculate the fractional cropland area within each 0.1° grid cell. In the revised analysis, we only consider "cropland pixels" to be those where crop coverage is dominant. This rigorous filtering significantly minimizes the influence of non-crop components like shelterbelts and rural settlements, ensuring that the vegetation signals we discuss are primarily driven by agricultural activity.

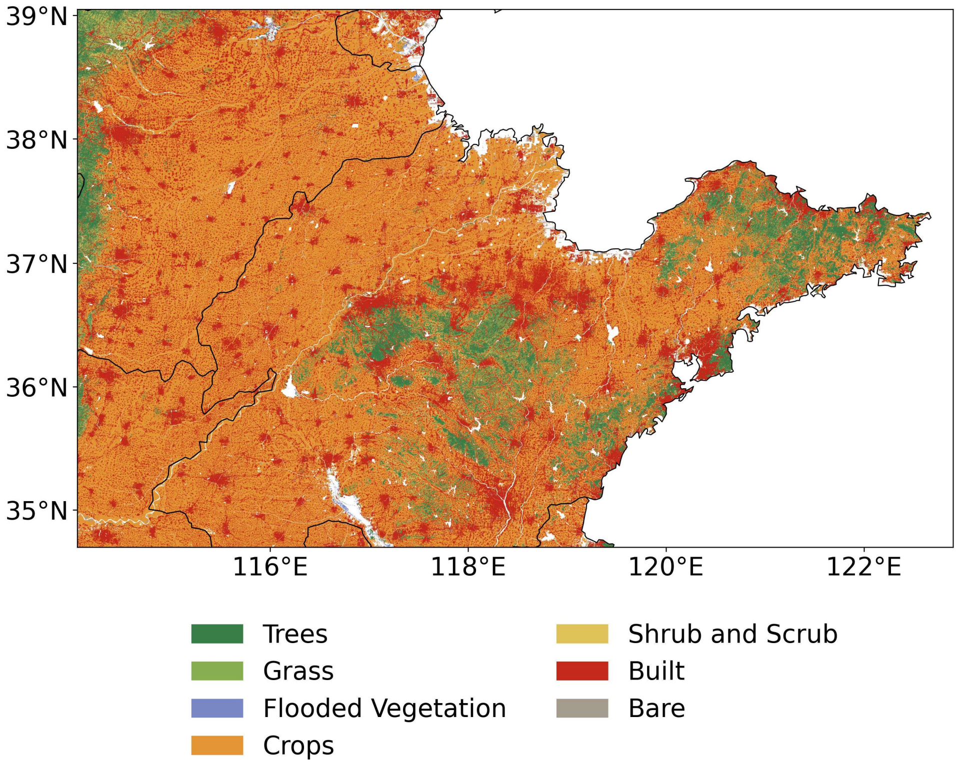
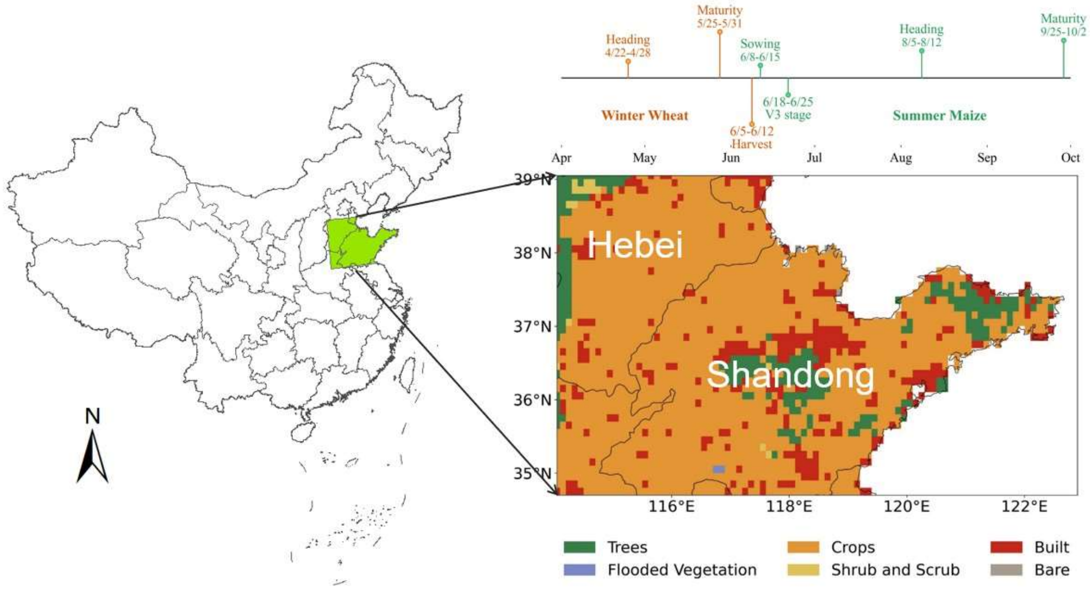


Figure R1. Original Land use distribution of the study area derived from the Dynamic World V1 land cover product.

Regarding the resolution of the analysis (0.1°), we would like to clarify our rationale. While we recognize the value of field-scale observations from Landsat or Sentinel, our study is fundamentally **a climate attribution analysis**. The ERA5-Land climate data used to drive our diagnostic has a native resolution of 0.1°. In the field of climate-vegetation interactions, it is generally preferred to maintain spatial consistency between the biological response (NDVI/GPP) and the meteorological forcing. Downscaling climate data would introduce substantial interpolation uncertainties, which could potentially bias the attribution of dry-hot impacts.

Furthermore, we did test the feasibility of using Landsat-derived NDVI. However, as we detailed in our response to **Comment 16**, the frequent cloud cover and longer revisit cycles in the NCP during the growing season led to significant data gaps and artifacts. For capturing the rapid development and regional impact of compound extremes, the high temporal frequency and radiometric stability of the MODIS product provide a more robust and continuous signal.

We have updated the manuscript and Figures (below is the example for the revised Figure 1 in the main text) to reflect these changes in the masking procedure. We believe this refined approach strikes a necessary balance between utilizing high-resolution land-cover information and maintaining the physical consistency required for large-scale climate attribution.



Revised Figure 1. Geographic location and land use distribution of the study area, derived from the Dynamic World V1 land cover product. A detailed phenological timeline for 2024 was shown according to the 2024 Weekly Agricultural Meteorological Intelligence (National Meteorological

Center 2024) and official progress reports from the Shandong Provincial Department of Agriculture and Rural Affairs.

Comment 2: Lack of discrimination among winter wheat, maize, and other crops

The NCP follows a winter wheat–summer maize rotation, but this study does not distinguish crop types within pixels.

At any time, pixels may include: wheat at heading or maturity stage; harvested wheat fields; newly sown maize; other crops (vegetables, peanuts, cotton, orchards); fallow land.

Because NDVI and GPP vary strongly among crop types and phenological stages, the vegetation signal cannot be uniquely attributed to wheat or maize dynamics.

Reply: We appreciate the reviewer’s point regarding the complexity of crop types and their respective phenological stages in the NCP. In a 0.1° grid, sub-grid heterogeneity is inevitable. However, we believe that the regional-scale signal captured in our study is robust for the following reasons.

The winter wheat–summer maize rotation is the overwhelmingly dominant agricultural system in the NCP, accounting for the vast majority of the region's cultivated area and primary productivity. At a monthly resolution, the seasonal trajectory of NDVI and GPP is primarily governed by the synchronized development of these two staple crops. While other vegetation types (e.g., orchards or vegetables) exist, they lack the large-scale spatial and temporal synchrony of the wheat-maize system and thus contribute more to the background "noise" than to the regional-scale anomalies we identified.

To better support this interpretation and address the reviewer’s concern regarding sub-grid heterogeneity, we have provided additional evidence from both spatial and temporal perspectives:

（1） We have incorporated an independent 1 km resolution crop-type dataset to characterize the spatial distribution of agricultural land across the NCP. This high-resolution map (now provided as Fig. S1 in the Supplementary Material) confirms that the winter wheat-summer maize rotation system is the overwhelmingly dominant landscape component. This spatial evidence justifies that the 0.1° grid signals are primarily driven by this synchronized rotation system rather than fragmented localized crops.

(2) To link the observed 0.1° anomalies to actual biological processes, we have added a detailed phenological timeline for 2024 directly into **Fig. 1** of the manuscript. Unlike a multi-year average calendar, this timeline is derived from the 2024 Weekly Agricultural Meteorological Intelligence (National Meteorological Center 2024) and official progress reports from the Shandong Provincial Department of Agriculture and Rural Affairs. This allows for a more precise alignment between climate stressors and crop sensitivity in the specific year of the event.

(3) Since our analysis is conducted at a monthly resolution, it captures the integrated landscape-level response. The synergy between the 1 km spatial reference and the 2024-specific timeline allows us to clarify the biological significance of these signals (Figs. 1 and S1). For instance, an anomaly in May predominantly reflects the status of winter wheat during its heat-sensitive grain-filling stage.

(4) We have some sentences in the Discussion section to explicitly acknowledge the influence of mixed-pixel signals. We clarify that our findings represent the **landscape-level response** of the dominant agricultural system to climate extremes, which is consistent with the objectives of a regional-scale biogeochemical assessment. We added some descriptions in the discussion as follows: "*In addition, the relatively coarse spatial resolution may introduce mixed-pixel effects, where non-cropland signals are inadvertently included within cropland pixels, potentially confounding the interpretation of vegetation responses.*"

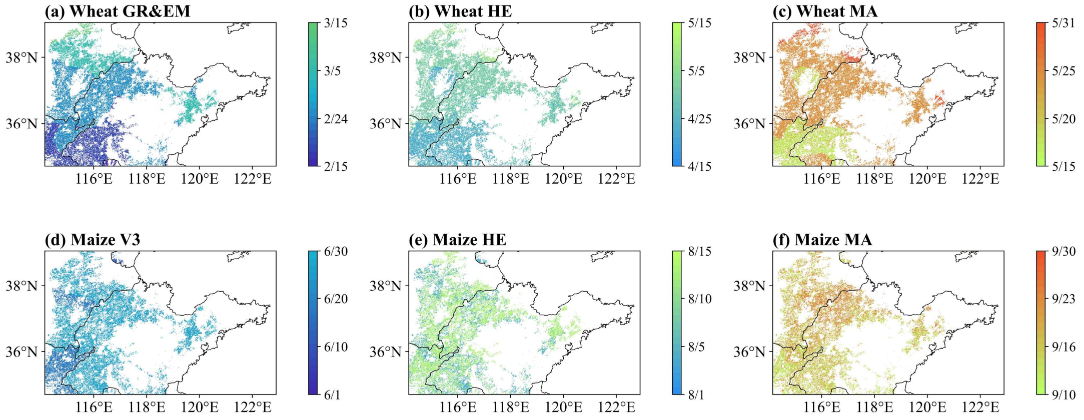


Figure S1. Phenological Map of Winter Wheat Summer Maize in the NCP. GR&EM represents Green up & Emergence date for wheat; V3 represents V3 stage for maize; HE represents Heading date for all the crops; MA represents Maturity date for all the crops.

Comment 3: Harvest transition and bare soil exposure confound June signals

The manuscript acknowledges that winter wheat harvest and maize sowing occur in early June. During this transition, canopy removal reduces NDVI, bare soil exposure dominates reflectance, and early maize emergence contributes little greenness. Thus, the June NDVI/GPP collapse may partly reflect land surface transition rather than physiological drought stress. This confounding effect is not adequately quantified.

Reply: We sincerely appreciate the reviewer's insightful comment regarding the potential confounding effects of the harvest-sowing transition. In the NCP, the winter wheat-summer maize rotation (Cui et al., 2018) is indeed tightly coupled, with the transition typically occurring in early to mid-June. However, we contend that the identified anomalies are primarily driven by physiological stress rather than the transition itself, for the following reasons:

(1) Our analysis focuses on monthly anomalies relative to the multi-year climatological mean, rather than raw NDVI/GPP values. The harvest-induced canopy removal and bare soil exposure are recurring seasonal events that occur every June. By subtracting the long-term mean (climatology) for June, the systematic signal of land surface transition is largely neutralized. Therefore, a negative anomaly in June 2024 indicates a productivity decline that exceeds the "normal" harvest-related reduction,

pointing toward additional stress factors specific to that year.

(2) According to the 2024 Weekly Agricultural Meteorological Intelligence (National Meteorological Center 2024), the extreme heat and soil moisture deficit in June 2024 significantly hindered the emergence and early growth of summer maize. While a normal June transition involves a brief period of bare soil followed by rapid green-up, the 2024 drought led to lower emergence rates, uneven seedling stands, and scorched leaves. This physiological impairment of the newly established maize canopy is the primary driver of the significant negative anomalies we observed, which go beyond the typical harvest transition.

(3) While the physical harvest and sowing may last only a few days, the monthly integrated signal captures the cumulative carbon assimilation capacity. The synergy between the heat-accelerated senescence of wheat in early June and the drought-stressed emergence of maize in mid-to-late June collectively suppressed the monthly GPP. This is further supported by our 2024-specific phenological timeline (Fig. 1), which shows that the most intense phase of the 2024 heatwave overlapped directly with the sensitive maize emergence window.

In summary, while we acknowledge that land surface transitions contribute to the background signal, the use of climatological anomalies effectively isolates the abnormal physiological response triggered by the 2024 extreme event.

Comment 4: Uncertainty of FluxSat GPP over croplands is insufficiently addressed

The FluxSat GPP product is derived from MODIS reflectance and flux tower calibration. While useful, it carries uncertainties in intensively managed croplands, especially under irrigation and rapid phenological transitions. I would expect the author to quantify expected GPP uncertainty over croplands in NCP; performance in irrigated agroecosystems as NCP region is heavily irrigated.

Reply: Thanks for your comments. To address this question, we first quantified the uncertainty associated with the FluxSat GPP product by utilizing the uncertainty estimates provided within the dataset itself. Specifically, we calculated the ratio between GPP anomalies and its uncertainties (i.e., the Signal-to-Noise Ratio) for April, May, and June 2024 (Fig. R2). The results indicate that the Signal-to-Noise Ratio (SNR) exceeds 1 in the major cropland regions, particularly in the northwestern and eastern parts of the study area, during April and May, and remains above 1 across most regions in June. This suggests that during the key growing season, the signal is sufficiently strong, and the detected GPP anomalies are generally distinguishable from the associated uncertainties.

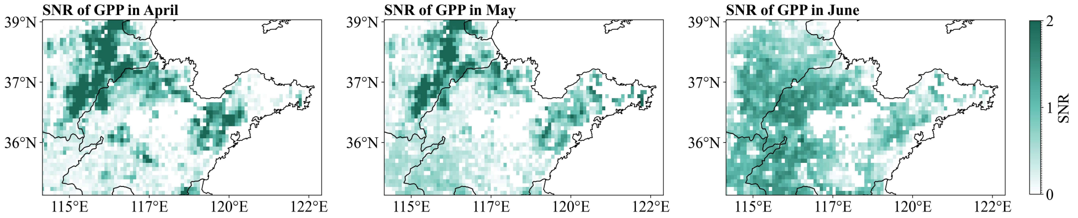


Figure R2. Spatial distribution map of the SNR between FluxSat GPP anomalies and their uncertainties from April to June 2024.

Second, we evaluated the FluxSat GPP against in situ measurements from the Yucheng Agro-ecosystem Station in Shandong Province. As a key site in the Chinese Ecosystem Research Network (CERN), Yucheng is highly representative of the NCP, featuring the dominant winter wheat–summer maize rotation and intensive irrigation management. Our validation, using monthly data from 2003 to 2011, demonstrates that FluxSat GPP captures the seasonal and interannual dynamics of irrigated croplands with high fidelity ($R^2 = 0.86$, $p < 0.001$; see Fig. S5). While FluxSat GPP exhibits a slight systematic underestimation compared to site-based observations, the strong linear coupling confirms its reliability for anomaly-based analysis.

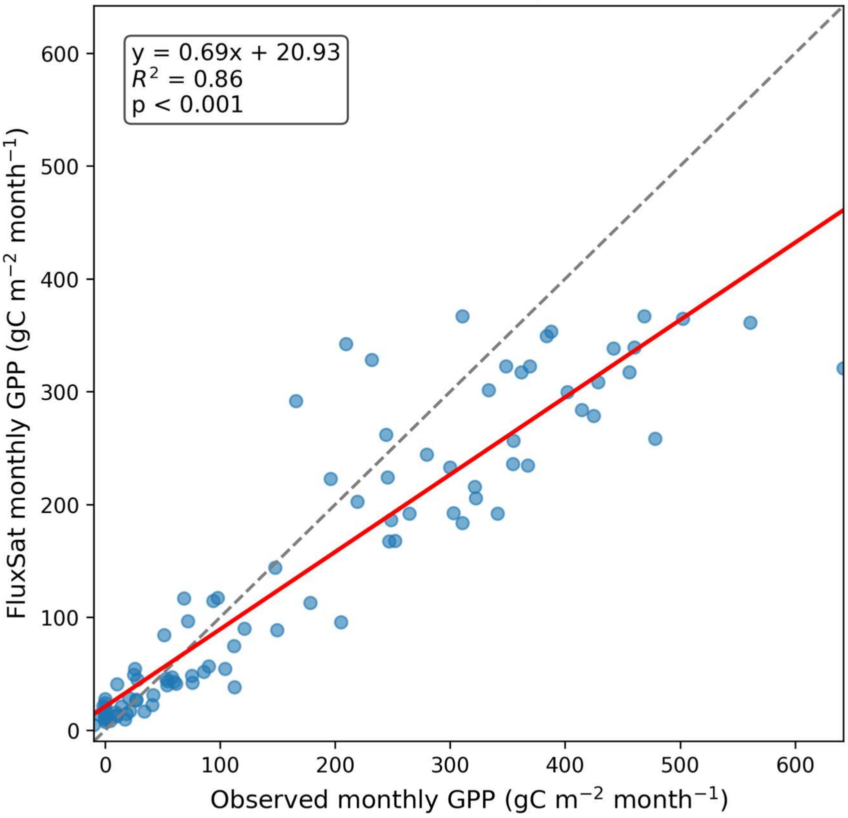


Figure S5. Scatter plot of FluxSat GPP versus in situ measurements at the Yucheng Agro-Ecosystem Station in Shandong Province.

Comment 5: Attribution risks conflating correlation with causation

The analysis demonstrates statistical associations between climate anomalies and vegetation indices but does not establish causal mechanisms. Given mixed pixels, heterogeneous crop composition, and non-crop vegetation contributions, attributing anomalies directly to crop growth responses remains inferential rather than mechanistic.

Reply: We acknowledge the importance of distinguishing between correlation and causation. However, the associations identified in this study are grounded in well-established biological mechanisms rather than being purely inferential. In agricultural ecosystems, the physiological response of crops to extreme heat and water deficit is governed by clear physical determinism. For instance, high temperatures and VPD directly trigger stomatal closure and impair photosynthetic enzymes. In the context of the 2024 NCP event, these climate anomalies act as the direct drivers of physiological impairment rather than mere statistical coincidences.

The validity of our statistical attribution is further reinforced by a priori biological knowledge. By precisely aligning the timing of these climate stressors with the critical growth stages in our 2024 phenological timeline (Fig. 1)—such as the grain-filling of wheat and the emergence of maize—we ensure that the statistical sensitivity reflects the known mechanistic vulnerability of the crops. This convergence between statistical significance and biological timing provides a physically-consistent causal link. While we acknowledge that a 0.1° resolution introduces sub-grid heterogeneity, the overwhelming dominance of the wheat-maize rotation provides a clear signal-to-noise ratio, allowing us to capture the integrated landscape response to climatic forcing. This approach is consistent with established regional-scale biogeochemical assessments where the objective is to quantify first-order impacts on primary productivity.

Additionally, we added some descriptions in the discussion as follows: "*In addition, the relatively coarse spatial resolution may introduce mixed-pixel effects, where non-cropland signals are inadvertently included within cropland pixels, potentially confounding the interpretation of vegetation responses.*".

Comment 6: Multiple Linear Regression framework is inadequate for extreme events

The study applies a linear regression model to characterize responses under an extreme compound event. However, the vegetation responses to heat and drought are nonlinear; Thresholds and tipping points are well documented; Interactions among VPD, soil moisture, and temperature are nonlinear; Phenological transitions introduce structural discontinuities. A linear model is therefore unlikely to capture system behavior under extreme stress.

Reply: We thank the reviewer for the insightful comments regarding the suitability of the modeling framework. We agree that vegetation responses to compound extreme events, such as heat and drought, are inherently nonlinear, involving threshold effects, nonlinear interactions among variables (e.g., VPD, soil moisture, and temperature), and structural changes associated with phenological transitions.

Nevertheless, multiple linear regression has been widely applied in studies of extreme climate impacts and agroecosystems. Statistical regression approaches are commonly used to link crop yield or vegetation variability with a set of key drivers, such as meteorological variables and remotely sensed vegetation indices (Feng et al., 2020; Feng et al., 2018; Lobell and Asseng 2017). The main advantages of MLR lie in its simplicity, interpretability, and its ability to quantify the relative contributions of different driving factors.

In this study, the primary objective of applying the MLR model is attribution rather than process-based simulation. Specifically, the model is intended to identify the relative importance of different climate factors during extreme events, rather than to fully represent the complex nonlinear ecological responses.

At the same time, we acknowledge the limitations of linear models in capturing nonlinear system behavior under extreme stress. We added the results of using random forests for analysis in Comment 21.

Comment 7: Model performance reporting and interpretation are insufficient

The manuscript reports $R^2 \approx 0.6$ only for GPP, while performance metrics for NDVI are not provided.

The interpretation over such metrics are unclear. In the regression formulation, is the response variable and is included as a predictor. Normally the R2 is to quantify the predicted Y and the observed Y, but in the Xs, is also there, which is a prediction from the model, I am not sure how this R2 is calculated and what the R2 represents, this requires additional elaboration.

Moreover, multicollinearity assessment need to be done and show how this could affect the model explanation power under extreme conditions.

Reply: We sincerely thank the reviewer for the insightful comments on model performance reporting and statistical interpretation.

First, regarding model performance, we have revised the manuscript to provide a more complete and transparent evaluation. The original manuscript presented the R^2 values for GPP in Figs. 6a-c. After replacing the soil moisture dataset with SMAP root-zone soil moisture, the R^2 for GPP improves to 0.74-0.88. In addition, we now report the model performance for NDVI in the Supplementary Material (Figs. S4a-c), with R^2 values around 0.6, ensuring that both response variables are consistently evaluated. Regarding the calculation and interpretation of R^2, we have clarified this point in the revised manuscript. The reported R^2 is calculated based on the agreement between estimated values and observed values of the response variable, following the standard definition. To avoid ambiguity, we have added the following statement in the manuscript: “Here, R^2 represents the agreement between estimated and observed values of the response variable.”

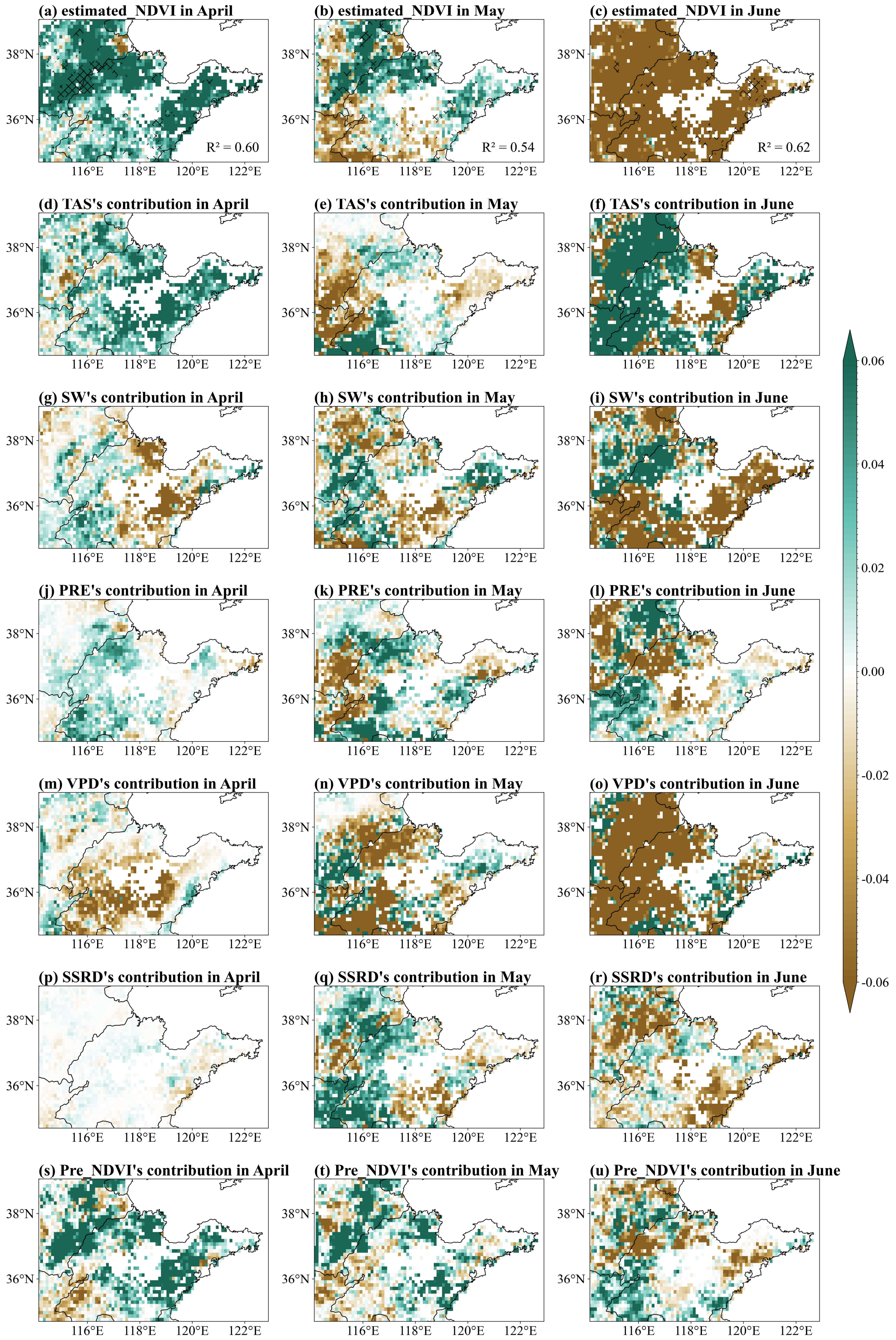


Figure S4. Geographical distributions of individual contributions from different driving factors. (a-c) Estimated NDVI during April–June 2024 based on the multiple linear regression model. Shaded areas labeled “XX” indicate grid cells where the model's F-test p-value is < 0.05. Panels (d–f), (g–i), (j–l), (m–

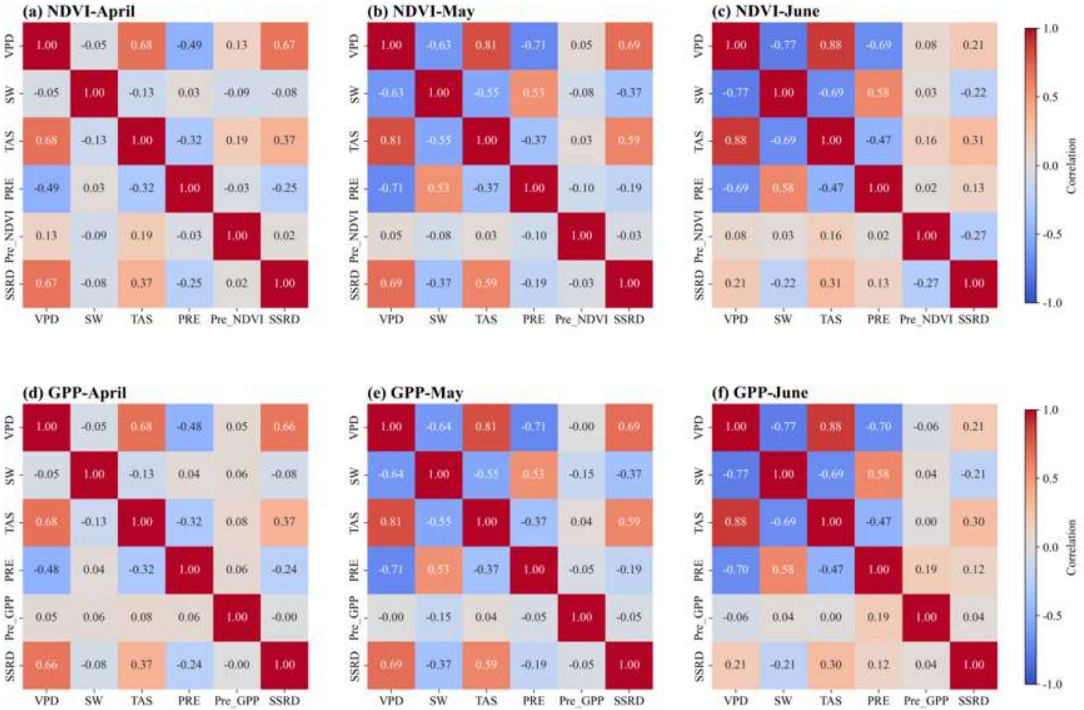
o), (p–r), and (s–u) present the separated contributions from TAS, SW, PRE, VPD, SSRD, and Pre_NDVI, respectively.

Regarding multicollinearity, we appreciate the reviewer's valuable suggestion. We calculated the Variance Inflation Factor (VIF) for all explanatory variables, and the results have been added to the Supplementary Material (Table S3). The results show that most explanatory variables exhibit VIF values below 5, indicating generally low multicollinearity. However, in June, the VIF values of TAS and VPD increase to approximately 7-9, suggesting a moderate level of collinearity. This is mainly attributable to their inherent physical coupling, as high temperature conditions are typically associated with increased atmospheric water demand, leading to elevated VPD, especially during dry-hot events.

Table. S3. VIF list of TAS, VPD, SW, PRE, SSRD, Pre_NDVI(Pre_GPP) in multiple linear regression model.

	Month	**TAS**	**VPD**	**SW**	**PRE**	**SSRD**	**Pre_NDVI(Pre_GPP)**
NDVI	April	3.50	8.60	2.16	2.82	3.76	3.10
	May	10.16	42.36	4.57	21.14	7.36	3.37
	June	19.00	23.77	6.32	5.18	2.87	3.21
GPP	April	3.37	7.56	2.13	2.68	3.57	2.67
	May	12.31	49.85	5.34	24.34	7.54	3.36
	June	13.18	22.18	6.71	5.35	2.98	3.48

To further diagnose this issue, we analyzed the correlation matrices across months (Fig. R3): (1) April: Overall correlations are relatively weak, with only moderate correlation between TAS and VPD (r ≈ 0.68), indicating limited collinearity effects. (2) May: The correlation between TAS and VPD becomes strong (r ≈ 0.81), while VPD shows a strong negative correlation with precipitation (PRE, r ≈ −0.71), reflecting enhanced coupling among climate variables. (3) June: The TAS–VPD correlation further strengthens (r ≈ 0.88), and VPD exhibits a strong negative correlation with soil moisture (SW, r ≈ −0.77), indicating tightly coupled atmospheric and soil moisture processes under hot and dry conditions.



(a) NDVI-April

	VPD	SW	TAS	PRE	Pre_NDVI	SSRD
VPD	1.00	-0.05	0.68	-0.49	0.13	0.67
SW	-0.05	1.00	-0.13	0.03	-0.09	-0.08
TAS	0.68	-0.13	1.00	-0.32	0.19	0.37
PRE	-0.49	0.03	-0.32	1.00	-0.03	-0.25
Pre_NDVI	0.13	-0.09	0.19	-0.03	1.00	0.02
SSRD	0.67	-0.08	0.37	-0.25	0.02	1.00

(b) NDVI-May

	VPD	SW	TAS	PRE	Pre_NDVI	SSRD
VPD	1.00	-0.63	0.81	-0.71	0.05	0.69
SW	-0.63	1.00	-0.55	0.53	-0.08	-0.37
TAS	0.81	-0.55	1.00	-0.37	0.03	0.59
PRE	-0.71	0.53	-0.37	1.00	-0.10	-0.19
Pre_NDVI	0.05	-0.08	0.03	-0.10	1.00	-0.03
SSRD	0.69	-0.37	0.59	-0.19	-0.03	1.00

(c) NDVI-June

	VPD	SW	TAS	PRE	Pre_NDVI	SSRD
VPD	1.00	-0.77	0.88	-0.69	0.08	0.21
SW	-0.77	1.00	-0.69	0.58	0.03	-0.22
TAS	0.88	-0.69	1.00	-0.47	0.16	0.31
PRE	-0.69	0.58	-0.47	1.00	0.02	0.13
Pre_NDVI	0.08	0.03	0.16	0.02	1.00	-0.27
SSRD	0.21	-0.22	0.31	0.13	-0.27	1.00

(d) GPP-April

	VPD	SW	TAS	PRE	Pre_GPP	SSRD
VPD	1.00	-0.05	0.68	-0.48	0.05	0.66
SW	-0.05	1.00	-0.13	0.04	0.06	-0.08
TAS	0.68	-0.13	1.00	-0.32	0.08	0.37
PRE	-0.48	0.04	-0.32	1.00	0.06	-0.24
Pre_GPP	0.05	0.06	0.08	0.06	1.00	-0.00
SSRD	0.66	-0.08	0.37	-0.24	-0.00	1.00

(e) GPP-May

	VPD	SW	TAS	PRE	Pre_GPP	SSRD
VPD	1.00	-0.64	0.81	-0.71	-0.00	0.69
SW	-0.64	1.00	-0.55	0.53	-0.15	-0.37
TAS	0.81	-0.55	1.00	-0.37	0.04	0.59
PRE	-0.71	0.53	-0.37	1.00	-0.05	-0.19
Pre_GPP	-0.00	-0.15	0.04	-0.05	1.00	-0.05
SSRD	0.69	-0.37	0.59	-0.19	-0.05	1.00

(f) GPP-June

	VPD	SW	TAS	PRE	Pre_GPP	SSRD
VPD	1.00	-0.77	0.88	-0.70	-0.06	0.21
SW	-0.77	1.00	-0.69	0.58	0.04	-0.21
TAS	0.88	-0.69	1.00	-0.47	0.00	0.30
PRE	-0.70	0.58	-0.47	1.00	0.19	0.12
Pre_GPP	-0.06	0.04	0.00	0.19	1.00	0.04
SSRD	0.21	-0.21	0.30	0.12	0.04	1.00

Figure R3. Correlation coefficient matrix heatmap, variables include VPD, SW, TAS, PRE, SSRD, Pre_NDVI (Pre_GPP)

These patterns are fully consistent with the VIF results and demonstrate that the observed collinearity primarily arises from the physical coupling of climate variables during periods of intensified heat and drought, rather than from model specification issues.

We acknowledge that collinearity is a potential concern in multivariate attribution. However, we have calculated the VIFs for all climatic drivers included in the model to ensure the reliability of our results. All VIF values remain within a widely accepted range (typically <10, indicating that multi-collinearity does not severely inflate the standard errors or destabilize the coefficient estimates.

Comment 8: Definition issue for anomaly

Another very important issue is the definition of anomaly in this study, for example L154 (formula 2), the anomaly is defined as the yield in 2024 minus the mean yield within 2001-2023 and then divide by mean yield, this is more of the deviation from the mean condition.

Within the years 2001-2023, the yields are varying too, there will be a "normal" range of the variation, the part for the 2024 deviation outside this "normal" range is the real anomaly.

I would ask the author to redo the calculation by defining a true anomaly signal - Standardized anomaly (z-score).

This should apply for all the variables: yield, GPP, NDVI, etc.

Reply: We sincerely thank the reviewer for this insightful comment regarding the definition of anomalies. We agree that standardized anomalies (e.g., z-scores) are useful for quantifying deviations relative to interannual variability and for enabling comparisons across variables with different units.

In this study, however, we intentionally defined anomalies as relative deviations from the long-term mean, as our primary objective is to quantify proportional changes under anomalous climate conditions. This formulation directly reflects the relative magnitude of change compared to the climatological baseline, which is more consistent with our focus on assessing the intensity of vegetation and productivity responses.

From a mathematical perspective, the difference between this definition and the standardized anomaly mainly lies in the normalization factor. While z-scores emphasize deviations relative to background variability, both approaches preserve the temporal and spatial patterns of anomalies. Therefore, the key findings of this study—particularly those related to the spatial consistency and relative magnitude of responses—are not expected to be sensitive to the choice of anomaly definition.

Given this, and considering that the mean-based anomaly definition has been widely adopted in similar studies (Finch et al., 2025; Li et al., 2023; Robert et al., 2020), we retain the current formulation.

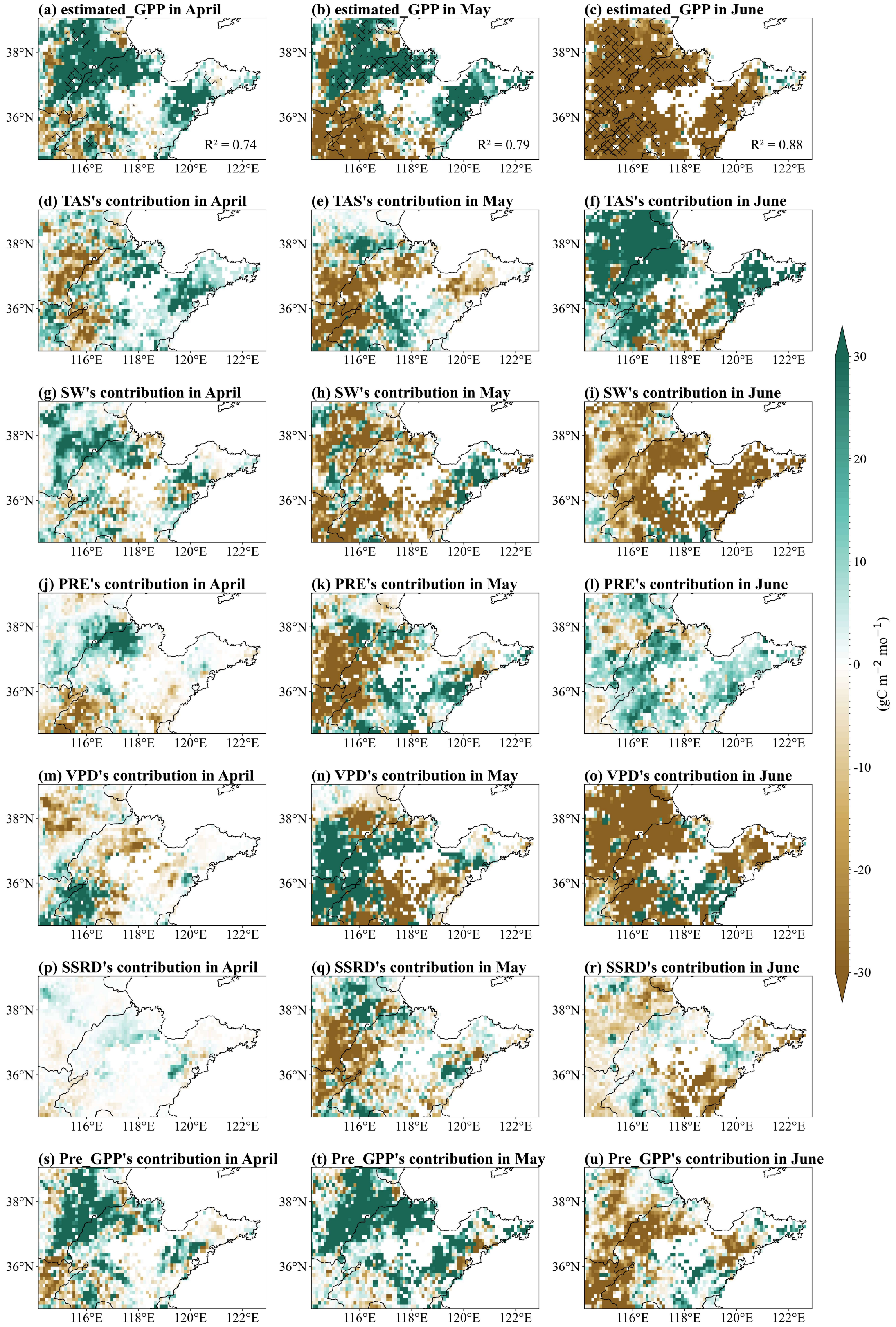
Comment 9: Irrigation effects omitted

Irrigation is a dominant control on crop productivity in the NCP, yet it is not explicitly represented.

Reply: We sincerely thank the reviewer for highlighting the important role of irrigation in crop productivity in the NCP. We agree that irrigation is a key management factor influencing crop growth. In the NCP winter wheat-summer maize rotation system, irrigation practices are primarily scheduled according to critical crop growth stages (e.g., green-up, jointing, heading, and grain filling), rather than dynamically adjusted based on short-term climate variability. This phenology-driven irrigation regime has been documented in previous studies (Zhang 2018). In addition, existing research indicates that the total irrigation amount tends to remain relatively stable (typically around 120 – 150 mm with two irrigation events) across dry, normal, and wet years to achieve optimal winter wheat yield (Zhou et al., 2022). These findings suggest that irrigation in this region is largely governed by crop developmental requirements, with limited responsiveness to short-term climate fluctuations.

Nevertheless, we agree that irrigation may still influence GPP estimates by modifying soil moisture conditions. To better account for this effect, we improved our modeling framework in the revised manuscript. Specifically, we replaced the original ERA5-Land soil moisture dataset with SMAP root-zone soil moisture (2015-2024). This dataset is derived from L-band microwave radiometer and radar observations, which can penetrate clouds and vegetation to directly capture soil moisture conditions, thereby providing improved representation of irrigation-induced moisture variability.

Based on this dataset, all subsequent attribution analyses were recalculated using SMAP, and the corresponding figures (Fig. 2, Figs. 5-7) have been updated in the revised manuscript. The figure below presents an example of the relative contributions of different drivers to GPP derived from the multiple linear regression after incorporating SMAP data:


(a) estimated_GPP in April
(b) estimated_GPP in May
(c) estimated_GPP in June
(d) TAS's contribution in April
(e) TAS's contribution in May
(f) TAS's contribution in June
(g) SW's contribution in April
(h) SW's contribution in May
(i) SW's contribution in June
(j) PRE's contribution in April
(k) PRE's contribution in May
(l) PRE's contribution in June
(m) VPD's contribution in April
(n) VPD's contribution in May
(o) VPD's contribution in June
(p) SSRD's contribution in April
(q) SSRD's contribution in May
(r) SSRD's contribution in June
(s) Pre_GPP's contribution in April
(t) Pre_GPP's contribution in May
(u) Pre_GPP's contribution in June

Revised Figure 6. Geographical distributions of individual contributions from different driving factors. (a-c) Estimated GPP during April-June 2024 based on the multiple linear regression model. Shaded areas labeled “XX” indicate grid cells where the model's F-test p-value is < 0.05. Panels (d-f), (g-i), (j-l), (m-o), (p-r), and (s-u) present the separated contributions from TAS, SW, PRE, VPD, SSRD, and Pre_GPP, respectively.

Comment 10: Phenology inference without direct phenology data

Crop stages are inferred rather than validated using phenological datasets.

Reply: We thank the reviewer for highlighting the importance of explicitly incorporating phenological information. To address this, we have added a detailed phenological timeline for 2024 directly into **Figure 1** of the manuscript and an independent 1 km resolution crop-type dataset in supplementary Figure S1 (see **Comment 2** for details). These phenological patterns are consistent with established agronomic studies (Qin et al., 2012) and provide independent support for the analysis of key growth stages and their variations in this study.

Comment 11: Assumed crop species over the field

Same as above, for the crop species over the land, they are based on the assumptions (maize-wheat rotation).

Reply: We sincerely thank the reviewer for this valuable comment. We would like to clarify that this study is not conducted at the field scale, but rather aims to identify the dominant relationships and first-order response characteristics between climate anomalies and vegetation dynamics at a regional scale. The winter wheat-summer maize rotation is the dominant cropping system in the NCP and has been widely documented in previous studies (Qin et al., 2012; Wang et al., 2018). We also drew a phenological map of the NCP, please refer to comment 2 for details

Specific Comments:

Comment 12: L30-32, NDVI and GPP are still increasing markedly in April and May during the dry-hot year? Which crops is this? Maize or wheat or any other crops? How do you know the decrease in June is solely due to the dry-hot event? Not because of harvesting, tillage, or other management practices.

For NDVI and GPP, as they are from satellite, it is also very likely that the signal in this month is different from other months or other years due to image saturations or local weather conditions. How do you know it is not the satellite issue?

Reply: We sincerely thank the reviewer for these insightful comments. During the 2024 dry-hot event, NDVI and GPP show a significant increase in April and a slightly positive anomaly in May. This period corresponds primarily to the growth stage of winter wheat in the NCP, while summer maize is typically sown after the harvest of winter wheat in early June. Therefore, the vegetation signals in April-May are mainly dominated by winter wheat rather than maize.

Regarding the decline observed in June, our analysis is based on anomalies relative to the 2001-2023 baseline. This anomaly-based approach effectively removes the recurring effects of agricultural management practices (e.g., harvesting, tillage, and sowing) that occur at similar times each year. Moreover, we find that both NDVI and GPP in June 2024 reach their lowest values since 2000. This motivates our further investigation into the associated climate anomalies and attribution analysis. The results indicate that June 2024 is characterized by significantly elevated temperature, pronounced dryness, and high VPD, with VPD exerting a dominant negative effect on vegetation productivity.

Regarding potential satellite-related uncertainties, the MODIS NDVI product used in this study provides near-daily observations and applies compositing algorithms to reduce cloud contamination. In addition, quality assurance (QA) flags are used to identify and mask clouds and shadows, ensuring more reliable NDVI estimates. The FluxSat GPP product is derived from high-frequency observations aggregated to the monthly scale, which helps reduce short-term noise.

Furthermore, meteorological conditions during April-June 2024 were characterized by low precipitation, strong radiation, and relatively low cloud cover, which further minimizes potential cloud-related biases in satellite observations. Therefore, we consider the influence of satellite-related uncertainties on our conclusions to be limited.

Comment 13: L36-37, any proof of this?

Reply: The conclusion that vegetation dynamics in May are largely governed by the carryover effect from the previous month is supported by the convergence of our sensitivity and contribution analyses (Figs. 5-7). Specifically, both NDVI and GPP in May exhibit their highest standardized sensitivity coefficients to the antecedent vegetation state (pre_NDVI and pre_GPP) rather than to concurrent climate drivers. This statistical dominance is further corroborated by the contribution analysis, which identifies the previous month's vegetation state as the primary source of variability for May.

Mechanistically, this carryover effect reflects the biological "memory" of the winter wheat system in the NCP. The biomass and canopy structure established during the rapid vegetative growth phase in April (e.g., jointing and booting) provide the essential physiological foundation and non-structural carbohydrate reserves for the subsequent heading and grain-filling stages in May. The high sensitivity to antecedent states captured by our model is a direct manifestation of this robust biological coupling between successive growth stages.

Comment 14: L97-98, Land cover data with 0.05x0.05 degree is not acceptable as you are using it to keep only the cropland pixels.

Reply: We sincerely thank the reviewer for this important comment. We agree that the use of 0.05° land cover data in the original manuscript may not be sufficient for accurately identifying cropland pixels. In the revised manuscript, we have replaced this dataset with the 2024 Dynamic World land V1 cover product at 10 m spatial resolution to better delineate cropland areas (see **Comment 1** for details). Specifically, cropland pixels were first identified using the high-resolution dataset, and then aggregated to 0.1° resolution to generate a cropland mask for subsequent analyses. All analyses in the revised manuscript are based on this updated cropland mask, which substantially reduces mixed-pixel effects associated with coarse-resolution land cover data and improves the robustness of our results.

Comment 15: L101-104, this is a big problem, you cannot simply assume that there are only maize and wheat on the field at a certain month, please use crop species masks at fine resolution. From your current setting, inside a 0.05x0.05 degree pixel or 0.1x0.1 degree pixel, there are too many fields, they can be covered with different crops other than maize/wheat, and even for maize/wheat, they might be in different growing stages, with different management (different sowing time, irrigation, weed/pest control, tillage, harvest), which could lead to a very different optical characteristics that were captured by the satellite, and all of them are mixed together, you cannot simply attribute them to a simple maize/wheat at the same growing stage.

Reply: We appreciate the reviewer's critical point regarding sub-grid heterogeneity. At the field scale, variations in crop species, management practices, and micro-phenology undeniably exist. However, for a regional-scale assessment of the 2024 extreme event, we contend that the identified signals remain robustly representative of the dominant wheat-maize system for the following reasons:

First, we have replaced the previous land cover dataset with the high-resolution Dynamic World Land Cover (V1) product to refine our cropland masking. Crucially, we have incorporated an independent 1 km resolution crop-type dataset as a spatial reference (Fig. S1). This high-resolution evidence confirms that the winter wheat–summer maize rotation is the overwhelmingly dominant landscape component in the NCP, justifying our focus on this specific system. While minor crops and varied management practices exist, they lack the large-scale spatial and temporal synchrony of the wheat-maize system; thus, their contributions to the 0.1° grid are largely secondary to the coherent regional-scale anomalies identified in this study.

Second, the concern about varying growth stages is addressed by the highly synchronized nature of the NCP agricultural calendar. As illustrated in our 2024-specific phenological timeline (Fig. 1)—derived from the 2024 Weekly Agricultural Meteorological Intelligence (National Meteorological Center 2024) and official progress reports from the Shandong Provincial Department of Agriculture and Rural Affairs—the crop development in 2024 was exceptionally uniform due to regional-scale climatic forcing. For instance, the persistent heat in early June 2024 triggered a rapid, synchronized "harvest-to-sowing" window. This large-scale synchronization ensures that the monthly anomalies at 0.1° resolution are not a random mixture of signals, but an integrated landscape-level response of the dominant rotation system at critical physiological stages.

Finally, while we acknowledge that sub-pixel heterogeneity is inevitable at a 0.1° resolution, our focus is on capturing the first-order response patterns of regional productivity. Previous studies (Cui et al., 2018; Qin et al., 2012) have demonstrated that the core phenological progression in the NCP (e.g., jointing to heading in April-May and harvest in early June) provides a stable basis for interpreting seasonal vegetation signals at this scale. By utilizing climatological anomalies, we effectively factor out the "normal" background heterogeneity, isolating the abnormal physiological impairment caused by the 2024 dry-hot event.

Comment 16: L128-129, for NDVI, why not derive from landsat or sentinel with 10/30m resolution.

Reply: We appreciate the reviewer's suggestion to utilize higher-resolution products such as Landsat or Sentinel. To assess the robustness of our results, we examined Landsat 8 NDVI (30 m resolution) anomalies (Fig. R4). While the Landsat-based data successfully captured the general patterns of the 2024 event—specifically the green-up in April and the pronounced collapse in June—it also introduced significant limitations that justify our reliance on MODIS.

First, the Landsat-based results exhibit clear spatial inconsistencies and "striping" artifacts, particularly during the critical transition in May. These artifacts, coupled with persistent cloud contamination, result in a fragmented and noisy spatial signal at the 30 m scale. Due to its 16-day revisit cycle, Landsat is highly sensitive to transient land-surface changes; a single contaminated or missing scene in May can lead to an unrepresentative monthly mean, which explains the observed discrepancies between Landsat and the more temporally stable MODIS composites during that month.

Second, MODIS's daily revisit frequency provides the temporal fidelity necessary to generate a high-quality, cloud-free monthly integral. This is essential for defining the climatological baseline (2001-2023) required for accurate anomaly calculation. Using a coarser-resolution but temporally dense sensor like MODIS significantly improves the Signal-to-Noise Ratio (SNR) for regional-scale studies, effectively filtering out the sensor-specific noise and orbital artifacts prevalent in 30 m products. Finally, since the driving climatic factors (e.g., VPD, temperature) are inherently coarse-grained (0.1°), conducting the attribution at 30 m resolution would introduce substantial scale-mismatch uncertainties during resampling without providing additional mechanistic depth. Therefore, while Landsat provides useful localized snapshots, MODIS remains the most reliable and physically consistent choice for capturing the continuous, integrated landscape-level response across the NCP.

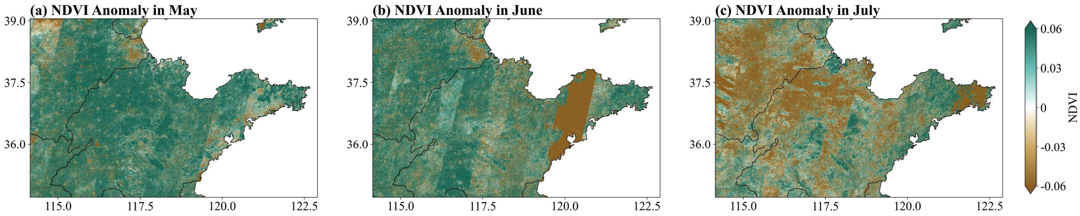


Figure R4. Spatial distribution map of Landsat 8 NDVI anomalies from April to June 2024.

Comment 17: L129-131, need to validate the GPP data over croplands in NCP, with/without irrigation.

Reply: Regarding the validation of GPP over NCP croplands, we have addressed this in our Response to Comment 4. Specifically, we compared the FluxSat GPP against in situ measurements from the Yucheng Agro-ecosystem Station, which represents the typical irrigated winter wheat–summer maize rotation of the NCP. The results (detailed in Fig. S5) show a high temporal consistency with an R^2 of 0.86 ($p<0.001$), confirming that FluxSat GPP reliably captures the productivity dynamics of irrigated Agro-ecosystems.

Comment 18: L136, with such coarse resolution, I doubt the model validity under phonological transitions, vegetation indices during harvest transitions may reflect canopy removal, bare soil exposure, land surface changes, rather than climate-driven physiological stress. How do you know if the trend you find in your data and model really reflects climate-driven physiological stress, rather than canopy removal, bare soil exposure, or land surface changes.

Reply: We sincerely thank the reviewer for this insightful and important comment. We fully agree that, at coarse spatial resolution and during phenological transitions (e.g., harvest periods), vegetation indices may be influenced by canopy removal, bare soil exposure, and land surface changes, rather than purely reflecting climate-driven physiological stress.

In response to this concern, we have taken several steps to improve the robustness of our analysis. First, we replaced the previous land cover dataset with the higher-resolution Dynamic World Land Cover V1 product to better constrain cropland extent. To ensure consistency with the climate datasets, the land cover data were subsequently resampled to a 0.1° spatial resolution.

We acknowledge that sub-pixel heterogeneity and management-induced signals cannot be completely eliminated. However, the use of monthly anomaly data (relative to a multi-year baseline) helps suppress the influence of recurrent and seasonally consistent management practices (e.g., harvest cycles), which are largely embedded in the climatology. This approach allows us to better isolate abnormal deviations associated with extreme climate conditions.

Furthermore, both our anomaly patterns and attribution analysis consistently reveal a strong and coherent response of vegetation to concurrent heat and drought conditions. The spatial and temporal correspondence between vegetation anomalies and climate extremes provides additional evidence that the observed signals are not solely driven by harvest-related surface changes, but are strongly linked to climate-induced stress.

Taken together, although we acknowledge the potential confounding effects during harvest transitions, the combined evidence from anomaly-based analysis and climate attribution supports our interpretation that the detected trends predominantly reflect climate-driven physiological responses at the regional scale.

Comment 19: L139-140, only 1 site?

Reply: We acknowledge the reviewer’s comment regarding the number of sites. While the field-scale validation relies on one primary experimental location (36.1565°N, 117.1607°E), we emphasize that this dataset is derived from rigorous, multi-year controlled experiments rather than a simple point observation.

First, this experimental plot provides high-quality data through replicated trials (two to three replicates for each variety), covering the region's dominant varieties for both winter wheat (Shimai, Jimai) and summer maize (Lainong14, Zhengdan958). Such multi-variety, multi-replicate experiments are exceptionally labor-intensive and time-consuming, as they require consistent long-term management and standardized measurement protocols. In this study, all measurements were conducted by the same operator to minimize procedural inconsistencies—a level of precision that is often unattainable in larger, multi-site networks.

Second, this controlled experimental approach is specifically designed to isolate the impact of climatic stressors from management noise. By maintaining a conventional fertilization regime and consistent procedures across different years (2020–2025), the dataset provides a "gold standard" for evaluating the model’s ability to capture year-to-year yield fluctuations driven by extreme events.

Finally, the fact that the yield variations observed at this intensive experimental site align with province-level statistical data further confirms its regional representativeness (Fig. 4). Therefore, we contend that the high fidelity and rigorous control of this single-site experimental data offer a robust and meaningful independent validation of our model's performance at the field scale.

Comment 20: L145-156, I am not sure if you should call this anomaly, as I also mentioned before, this is more of the derivation from the mean condition. Please revise it with z-score. This should apply for GPP and NDVI too, as you are discussing the anomaly of them too.

Reply: Regarding the definition of anomalies and the use of z-scores, we have provided a comprehensive clarification and mathematical justification in our Response to **Comment 8**.

Comment 21: L157-180, I strongly doubt the capability of a linear model can capture these relations under extreme weather conditions. Please compare with other machine learning algorithms.

Reply: We sincerely thank the reviewer for this valuable suggestion. We agree that under extreme weather conditions, the relationships between variables may exhibit nonlinear characteristics, and linear models may have limitations in capturing such complexities. In this study, the multiple linear regression model was primarily adopted due to its interpretability and its widespread use in climate-vegetation studies. Following the reviewer’ s suggestion, we further introduced a Random Forest model as a representative machine learning approach for comparison. In addition, SHAP (Shapley Additive Explanations) values were calculated to quantify the contribution of each feature in 2024. As shown in the figure below:

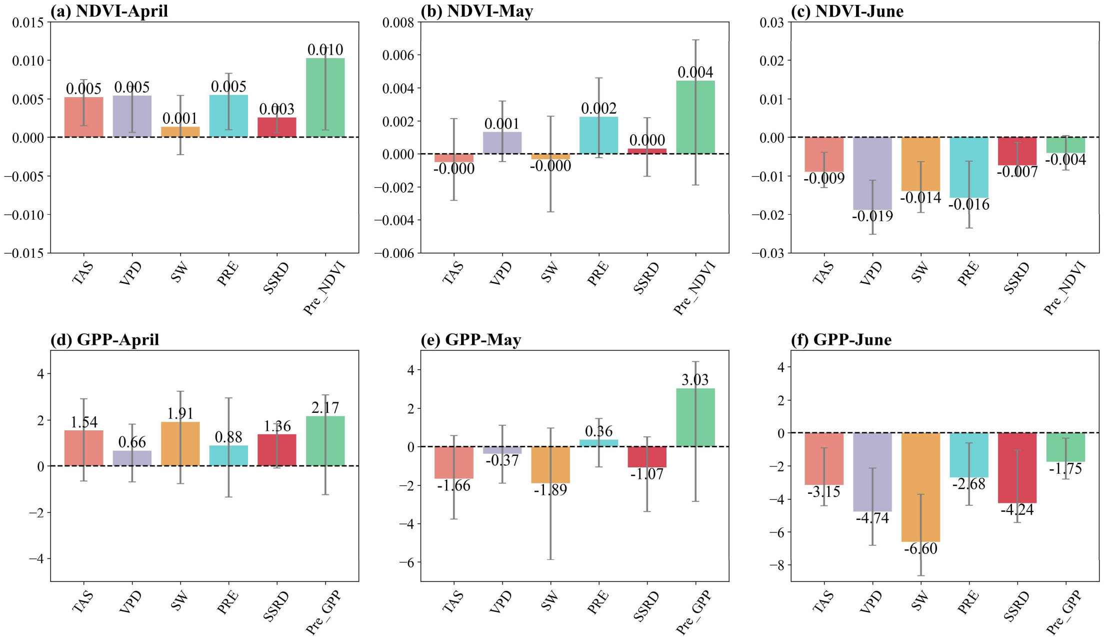


Figure R5. SHAP values of (a-c) NDVI and (d-f) GPP to multiple driving factors from April to June over the NCP. The driving factors include TAS, VPD, SW, PRE, SSRD, and the previous month's NDVI or GPP (Pre_NDVI or Pre_GPP).

we compared the SHAP-derived feature contributions from the Random Forest model with the contributions reported in Figure 7. The overall patterns are largely consistent.

Specifically, in April, both NDVI and GPP are strongly positively influenced by previous-month vegetation conditions and TAS, while SW also shows a strong positive contribution to GPP. In May, previous-month vegetation conditions emerge as the dominant contributing factor for both NDVI and GPP, and SW exhibits a strong negative contribution to GPP. In June, VPD shows a strong negative contribution to both NDVI and GPP. However, TAS exhibits a negative contribution, which differs somewhat from our original results.

It should be noted that SHAP values represent feature contributions based on a different methodological framework than that used in our linear regression analysis, and therefore exact agreement is not expected. Overall, the linear regression model captures the primary relationships among variables, and the use of a machine learning approach does not alter the main conclusions of this study.

Comment 22: L172, sensitivities represent statistical associations rather than direct physiological responses.

Reply: We agree with the reviewer that the sensitivities derived in this study reflect statistical associations. However, we contend that these associations are physically grounded and serve as a robust proxy for understanding regional-scale ecosystem responses. In the context of this study, sensitivity is used as a diagnostic metric to quantify the relative coupling between climatic forcing and productivity shifts. While statistical models do not explicitly simulate leaf-level gas exchange or enzymatic kinetics, the strong and consistent sensitivity to VPD and temperature we identified is mechanistically consistent with established plant physiological principles.

Furthermore, this statistical approach is a standard and essential tool in satellite-based Earth system studies (De Keersmaecker et al., 2015; Liu et al., 2024) , as it allows for the attribution of dominant drivers across vast spatial domains where complex process-based modeling may introduce higher parameterization uncertainty.

Comment 23: L171-174, not clear how this sensitivity is defined and calculated, please elaborate more.

Reply: We sincerely thank the reviewer for pointing out this issue. In this study, a multiple linear regression model was employed to attribute the relationships between climate variables and vegetation dynamics. As shown in Eq. (3), all variables were standardized prior to regression, and thus the corresponding coefficients (β_i) represent standardized regression coefficients. Based on this, we define the standardized regression coefficient β_i as the sensitivity of NDVI/GPP to each variable. Specifically, it represents the change in NDVI/GPP associated with a one standard deviation change in the predictor, while holding other variables constant. This definition has two main advantages: (1) Standardization removes the effects of differing units and scales among variables, making the regression coefficients directly comparable and allowing them to reflect the relative strength of different climatic factors on vegetation dynamics. (2) Standardized regression coefficients help reduce the influence of scale differences on model stability, thereby improving the robustness and interpretability of the results.

Comment 24: L175-180, about the contribution, as in the model, previous-month vegetation anomalies are also included in the model as the predictor, can you clarify the ecological meaning of this term and discuss its potential influence on attribution results.

Reply: We thank the reviewer for this insightful comment. The inclusion of the previous-month vegetation anomaly in our model is motivated by the well-documented time-lag effects in vegetation responses to climatic forcing.

A substantial body of studies has demonstrated that vegetation responses to climate variability are often not instantaneous, but instead exhibit pronounced temporal delays due to ecosystem internal processes and spatial heterogeneity. For instance, in grassland ecosystems, extreme heat events have been shown to reduce net ecosystem productivity (NEP) not only in the same year but also persisting for up to 1-2 subsequent years (Arnone et al., 2008); similarly, studies across the Australian continent have reported approximately one-month lagged responses of vegetation to soil moisture variability (Chen et al., 2014). These findings collectively highlight that climate-vegetation interactions are characterized by intrinsic time delays, which represent a key mechanism governing vegetation dynamics (Chen et al., 2014).

Accordingly, we included the previous-month vegetation anomaly as an explanatory variable to acknowledge that current vegetation conditions are influenced not only by contemporaneous climatic drivers but also by recent vegetation states. This formulation improves the model's ability to capture vegetation dynamics under climate variability.

Comment 25: L184, For figure 2a, i, m, q, the shading area is the condition within 1σ? The lines indicated the condition for different years? From what we can observe, there are many years that outside this 1σ, why not included all those years for analysis the extreme condition's impact on crop growth and yield?? Otherwise, you should exclude those years in your baseline condition.

Reply: We sincerely thank the reviewer for the careful observation and constructive suggestion. Yes, in Figs. 2(a), (i), (m), and (q), the shaded area represents the baseline conditions within ±1σ, and the lines correspond to individual years. We agree that including all years beyond the ±1σ range to systematically analyze the impacts of extreme conditions is a valuable approach. However, the primary objective of this study is to investigate a specific compound hot-dry event that occurred in 2024 over the NCP, with a focus on understanding its driving factors and impacts on crop growth and yield. Therefore, our analysis is event-based rather than a comprehensive long-term assessment of all extreme years. The inclusion of detrended anomalies from other years in the figure is intended to provide a historical context, allowing for a clearer visualization of the relative extremity of the 2024 event. This approach, using historical variability as a reference to highlight a specific extreme event, has been adopted in previous studies (Yan et al., 2024).

Comment 26: L221-222, there might be correlations, but not the causations?

Reply: We sincerely thank the reviewer for this important comment. We agree that this statement is better interpreted in terms of correlation rather than causation. We have revised the relevant wording in the Results section to avoid any potential misinterpretation of causal relationships. The revised sentence is as follows: "The growth of crops in April may be related to the increase in temperature."

Comment 27: L232-233, How do you separate the impact of "practice" from the results? June includes wheat harvest; bare soil exposure; maize sowing period. The remote sensing signals may reflect land surface change; canopy removal; crop rotation timing and not purely physiological stress. This is the most important interpretational risk.

Reply: Regarding the potential confounding effects of harvest transitions and bare soil exposure, we have addressed this concern in detail in our **Response to Comment 3**. To summarize, our use of climatological anomalies and detrending effectively filters out recurring seasonal practices. The negative signals identified in June 2024 represent a productivity decline that exceeds the normal harvest-induced reduction. This anomalous loss is primarily driven by the extreme atmospheric VPD and soil drought, which triggered forced ripening and impaired maize emergence.

Comment 28: L240-242, the seeding density in different years are the same? This is the dry weight or what? What is the water content when balancing for different years? Is there irrigation? What is this error bar indicating? Error from what? If this is the site observation, why is there this error bar for a single site?

Reply: We sincerely thank the reviewer for carefully examining the experimental details. We clarify the following points:

(1) The field observation data used in this study were collected by the same person under consistent management practices. The seeding density was kept constant across years, with 1.8×10^6 seeds·hm^{-2} for winter wheat and 5.25×10^4 plants·hm^{-2} for summer maize. Fertilization and irrigation followed standard local agricultural practices, which are representative of typical field management conditions in the region.

(2) Although the experiment was conducted at a single site, the field was divided into several sub-plots with different crop varieties of wheat and maize, and observations were conducted over three growing seasons (see Table S2 for details). This design

allows the dataset to capture both spatial heterogeneity within the field and interannual variability.

(3) The error bars shown in the figure represent the standard deviation of yield across different sub-plots and years, reflecting within-site spatial variability and interannual fluctuations, rather than measurement error or replicate sampling uncertainty.

Comment 29: L298, what are these R2 showed in figure 6a-c? The predicted GPP vs. the observed GPP? Please plot the scatterplot with density as the color, instead of the maps. In your formula, you have Y_t as the response variable, and Y_t-1 as the predicter, how do you calculate this R2? And how to interpret it?

Reply: We sincerely thank the reviewer for these insightful questions. The R^2 values shown in Figs. 6a-c represent the agreement between the predicted GPP and the observed GPP. Specifically, we constructed a multiple linear regression model for each pixel across the study region. The R^2 reported in the figure corresponds to the spatial average of the pixel-wise R^2 values, which reflects the overall model performance at the regional scale. Regarding Eq. (3), Y_t denotes the vegetation condition (or GPP) at the current time step, while Y_{t-1} represents the vegetation condition in the previous time step. We would like to clarify that Y_{t-1} is not treated differently from other predictors; rather, it is included as an explanatory variable to account for the lagged effect of vegetation dynamics.

In terms of visualization, the original spatial maps were intended to demonstrate the consistency between modeled and observed GPP in terms of spatial patterns across the North China Plain. However, we agree that scatterplots provide a more direct and quantitative comparison. Following the reviewer's suggestion, we have added scatterplots of predicted versus observed GPP, with point density represented by color, to better illustrate their relationship and improve the clarity of model evaluation (Figs.

S6).

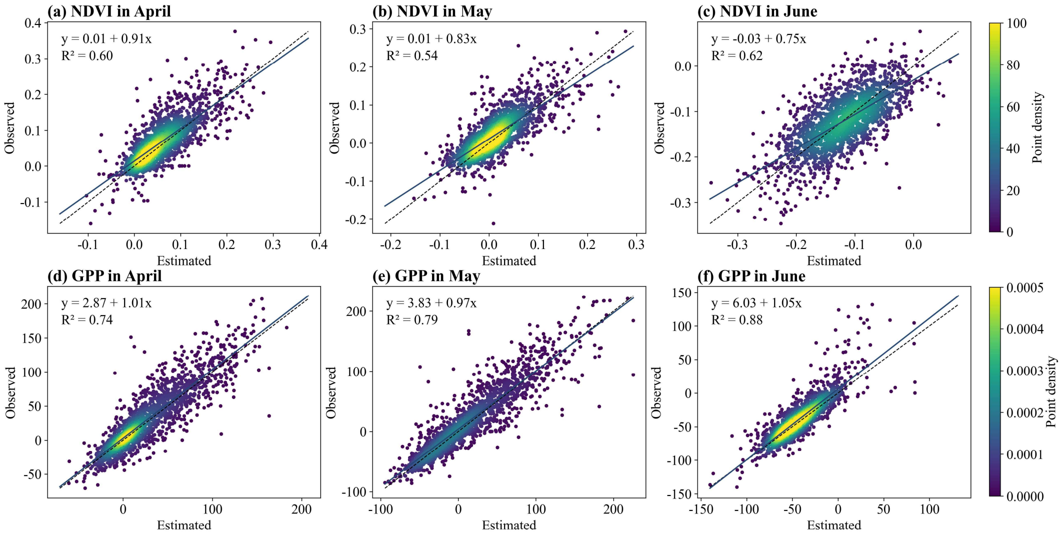


Figure S6. Scatter plot of (a-c) NDVI and (d-f) GPP anomalies estimated by the linear regression model versus the observed NDVI and GPP anomalies.

References

Arnone, J., Verburg, P., Johnson, D., Larsen, J., Jasoni, R., Lucchesi, A., Batts, C., Von Nagy, C., Coulombe, W., Schorran, D., et al.: Prolonged suppression of ecosystem carbon dioxide uptake after an anomalously warm year, Nature, 455, 383–386, https://doi.org/10.1038/nature07296, 2008.

Chen, T., De Jeu, R., Liu, Y., Van Der Werf, G. and Dolman, A.: Using satellite based soil moisture to quantify the water driven variability in NDVI: A case study over mainland Australia, Remote Sens. Environ., 140, 330–338, https://doi.org/10.1016/j.rse.2013.08.022, 2014.

Cui, Y., Xiao, D., Liu, S., Li, N., Jiang, L., Shi, X., Liu, X., Li, J., Lu, J. and Qin, Y.: Growth periods variation of summer maize and winter wheat and their correlations with hydrothermal conditions in recent years in China, Chinese Journal of Eco-Agriculture, 26, 388–396, 2018.

De Keersmaecker, W., Lhermitte, S., Tits, L., Honnay, O., Somers, B. and Coppin, P.: A model quantifying global vegetation resistance and resilience to short-term climate anomalies and their relationship with vegetation cover, Global Ecol. Biogeogr., 24, 539–548, https://doi.org/10.1111/geb.12279, 2015.

Feng, P., Wang, B., Liu, D., Waters, C., Xiao, D., Shi, L. and Yu, Q.: Dynamic wheat yield forecasts are improved by a hybrid approach using a biophysical model and machine learning technique, Agric. For. Meteorol., 285,https://doi.org/10.1016/j.agrformet.2020.107922, 2020.

Feng, P., Wang, B., Liu, D., Xing, H., Ji, F., Macadam, I., Ruan, H. and Yu, Q.: Impacts of rainfall extremes on wheat yield in semi-arid cropping systems in eastern Australia, Clim. Change, 147, 555–569, https://doi.org/10.1007/s10584-018-2170-x, 2018.

Finch, E., Chang, C., Kucharski, A., Sim, S., Ng, L. and Lowe, R.: Climate variation and serotype competition drive dengue outbreak dynamics in Singapore, Nat. Commun., 16,https://doi.org/10.1038/s41467-025-66411-6, 2025.

Li, X., Piao, S., Huntingford, C., Peñuelas, J., Yang, H., Xu, H., Chen, A., Friedlingstein, P., Keenan, T., Sitch, S., et al.: Global variations in critical drought thresholds that impact vegetation, Natl. Sci. Rev., 10,https://doi.org/10.1093/nsr/nwad049, 2023.

Liu, M., Trugman, A., Peñuelas, J. and Anderegg, W.: Climate-driven disturbances amplify forest drought sensitivity, Nature Climate Change, 14,https://doi.org/10.1038/s41558-024-02022-1, 2024.

Lobell, D. and Asseng, S.: Comparing estimates of climate change impacts from process-based and statistical crop models, Environ. Res. Lett., 12,https://doi.org/10.1088/1748-9326/aa518a, 2017.

National Meteorological Center 2024. Weekly Agricultural Meteorological Intelligence.

Qin, X., Liu, K., Zhou, L., Zhou, S., Lu, L. and Wang, R.: Characteristics of Annual Water Utilization in Winter Wheat- Summer Maize Rotation System in North China Plain, Scientia Agricultura Sinica, 45, 4014–4024, 2012.

Robert, M., Stewart-Ibarra, A. and Estallo, E.: Climate change and viral emergence: evidence from Aedes-borne arboviruses, Curr. Opin. Virol., 40, 41–47, https://doi.org/10.1016/j.coviro.2020.05.001, 2020.

Wang, Z., Chen, J., Tong, W., Xu, C. and Chen, F.: Impacts of Climate Change and Varietal Replacement on Winter Wheat Phenology in the North China Plain, Int. J. Plant Prod., 12, 251–263, https://doi.org/10.1007/s42106-018-0024-0, 2018.

Yan, R., Wang, X., Wang, J., Tan, J. and Ju, W.: Warm Early Summer Compensated Reduction in Photosynthesis Caused by 2022 Late Summer Extreme Drought Over the Tibetan Plateau, Journal of Geophysical Research-Biogeosciences, 129,https://doi.org/10.1029/2023jg007859, 2024.

Zhang, X.: Water use and water-saving irrigation in typical farmlands in the North China Plain, Chinese Journal of Eco-Agriculture, 26, 1454–1464, https://doi.org/10.13930/j.cnki.cjea.180636, 2018.

Zhou, L., Sun, S., Zhang, Z., Zhang, F., Guo, S., Shi, Y. and Yang, X.: High Yield and Water Use Efficiency Synergistical Improvement Irrigation Scheme of Winter Wheat in North China Plain Based on Meta-Analysis, Chin. J. Agrometeorol., 43, 515–526, https://doi.org/10.3969/j.issn.1000-6362.2022.07.001, 2022.