remote sensing

MDPI

Article

Phenological Artifacts in Satellite-Based Agricultural Drought Assessment: Limitations of Detrended and Deseasonalized NDVI Metrics for Crop Stress Evaluation

Yumeng Wang [1] and Zhengfeng Zhang [2,*]

[1] School of Agricultural Economics and Rural Development, Renmin University of China, Beijing 100872, China; wymmyw@ruc.edu.cn
[2] School of Public Administration and Policy, Renmin University of China, Beijing 100872, China
* Correspondence: zhangzhengfeng@ruc.edu.cn

Highlights

What are the main findings?

- Raw NDVI drought-damage metrics track phenological phase, not damage (r = 0.07)
- On cropland, deseasonalized metrics recover a growing-season grain yield signal

What are the implications of the main findings?

- Out-of-sample damage predictability is limited (linear $R^2 = 0.18$; RF 0.44)
- County greenness remains blind to oilseed yield losses
- Crop-type maps are required for crop-specific drought monitoring

Abstract

Satellite drought monitoring commonly assesses agricultural damage from vegetation-index drops, but such metrics are easily contaminated by seasonal phase and greening trends. Using monthly SPEI-3, NDVI, and nighttime light data for China's 13 major grain-producing provinces (2000–2023), we identify 1433 severe drought events and compare a raw baseline-minus-minimum NDVI metric with a detrended, deseasonalized vegetation anomaly. The two correlate at only 0.07, and the explanatory power of raw-metric models ($R^2 = 0.77$) comes almost entirely from onset-month seasonality. Once the yield-loss definition is corrected and the vegetation signal is restricted to cropland pixels using the 30 m China Land Cover Dataset, the deseasonalized anomaly is positively associated with detrended county grain losses (growing-season Spearman r = 0.33, robust to province-clustered inference), whereas the raw metric carries no positive yield signal. Recovery half-life is likewise phenology-dominated; censoring-aware Weibull survival models show that severity variables accelerate half-recovery attainment, mechanically, because the recovery target is anchored to the event minimum, while nighttime light has no robust effect. Out-of-sample predictability remains limited ($R^2 = 0.18$ linear, 0.44 random forest). Greenness metrics measure county-dominant vegetation; crop-specific monitoring requires crop-type maps.

Keywords: agricultural drought; NDVI; SPEI; phenological artifact; deseasonalization; detrending; cropland masking; survival analysis; nighttime light; crop yield validation

Check for updates

Academic Editor: Won-Ho Nam

Received: 20 October 2026
Revised: 19 November 2026
Accepted: 24 November 2026
Published: 28 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Under climate change, the frequency, duration, and intensity of extreme meteorological droughts have increased, posing a growing threat to agricultural production and global

food security [1–3]. In China, the 13 major grain-producing provinces account for more than 70% of national grain output [4], and their drought exposure is markedly heterogeneous: droughts occur most frequently in the northern and central provinces, where crop-yield risk is also highest [1,5].

Remote sensing has become an indispensable tool for large-scale drought impact assessment. The Standardized Precipitation Evapotranspiration Index (SPEI), which incorporates both precipitation and atmospheric evaporative demand, is considered superior to precipitation-only indices for characterizing agricultural drought impacts [6]. Satellite vegetation indices—particularly the Normalized Difference Vegetation Index (NDVI)—enable continuous monitoring of vegetation responses to drought across spatial scales [7,8].

A key but underexplored distinction in drought impact research is that between the magnitude of drought-induced vegetation decline (resistance) and the speed of post-drought recovery (resilience). Recovery dynamics have received considerable theoretical and empirical attention [2,3], whereas the determinants of the initial damage magnitude—how much vegetation greenness actually declines during a drought—remain comparatively understudied [9,10]. This asymmetry matters: identifying the factors that drive damage magnitude may be more actionable for near-term disaster response than predicting recovery trajectories, which are confounded by post-event management interventions and seasonal phenological cycles [11,12].

Existing evidence on damage predictors is fragmented. Drought characteristics (duration and intensity) are expected to influence damage magnitude, but their independent contributions are rarely separated because composite indices of cumulative water deficit conflate the two dimensions [13]. Pre-drought vegetation condition may also modulate damage susceptibility: systems with higher baseline greenness may be more vulnerable (more biomass at risk) or more resistant (deeper roots, irrigation access) [14,15]. Seasonal timing further complicates this relationship, as droughts occurring during the active growing season may produce different damage profiles from dormant-season droughts [16,17]. Adjudicating these competing hypotheses requires empirical evidence that remains limited, especially at the spatial scales relevant to operational monitoring.

Nighttime light (NTL) data are increasingly used as a proxy for socioeconomic development, correlating with GDP, urbanization, and energy consumption at fine spatial scales [18–20]. Whether NTL can capture dimensions of agricultural adaptive capacity—such as irrigation infrastructure, crop insurance penetration, or farm-level management—and thereby buffer drought-induced vegetation damage remains uncertain [21,22]. To date, no large-scale study has tested whether an NTL-based socioeconomic baseline moderates the relationship between drought severity and vegetation damage in agricultural systems.

This study makes three contributions. First, methodologically, we systematically quantify the phenological artifact embedded in absolute-level NDVI damage metrics. The raw "baseline-minus-minimum" construction is the most direct way to quantify an event-scale vegetation drop from satellite series, yet it has not, to our knowledge, been systematically benchmarked against detrended, deseasonalized alternatives at regional scale; difficulties with seasonal handling in absolute-level constructions are documented across seasonal-change detection [12,16] and productivity estimation studies [23]. We show that it correlates with a detrended, deseasonalized metric at only 0.07, and that the high explanatory power of models fit to the raw metric ($R^2 = 0.77$) essentially predicts seasonal phase rather than damage. This finding has broad implications for the drought impact assessment literature: any damage or recovery metric constructed from absolute vegetation-index levels embeds seasonal bias, and anomalies relative to a full-period climatology additionally embed residual greening trends. Second, empirically, supported by strict event identification (continuity and independence constraints), explicit detrending and deseasonalization, cropland mask-

ing, survival analysis of censored recovery times, mathematical-coupling diagnostics, and few-cluster-robust inference (wild cluster bootstrap), we provide credible estimates of the predictors of deseasonalized agricultural vegetation damage: peak intensity is robustly significant; baseline greenness is significant under the standard construction but sensitive to the baseline's seasonal alignment (suggestive rather than conclusive); duration is consistent in sign (and significant in the growing season and in larger-sample specifications); and the moderating role of nighttime light is not robust to specification. Third, application-wise, county-grouped cross-validation shows that standard remote sensing variables have limited out-of-sample predictability for deseasonalized damage (linear $R^2 = 0.18$; random forest 0.44), and the corrected, crop-masked yield validation delineates the validity boundary of greenness metrics: on cropland pixels, the deseasonalized anomaly recovers a positive, cluster-robust yield signal for county-dominant grain production in the growing season, whereas county-mean greenness—raw or deseasonalized—remains blind to the secondary oilseed category, giving direct empirical support to the claim that crop-specific monitoring in mixed cropping systems must overlay crop-type maps.

2. Materials and Methods

2.1. Study Area

The study area covers China's 13 major grain-producing provinces (Hebei, Inner Mongolia, Liaoning, Jilin, Heilongjiang, Jiangsu, Anhui, Jiangxi, Shandong, Henan, Hubei, Hunan, and Sichuan), which together account for more than 70% of national grain output, spanning climate zones from the temperate continental climate of the north to the subtropical monsoon climate of the south. Annual precipitation ranges from about 300 mm in Inner Mongolia to over 1500 mm in Sichuan. The growing season is defined as April–September, based on provincial crop calendars (maize April–September, rice May–October, wheat October–June of the following year) and the county-level NDVI phenological peak (May–August). We acknowledge that this is a general calendar-based classification: winter wheat (sown in October, harvested the following June) and double-cropped rice span different windows, and a sensitivity analysis using region-specific growing-season definitions (Section 3.4) leaves the core results unchanged. The basic spatial unit is the county-level administrative unit, totaling 1546 counties (2020 administrative division codes; all datasets are aggregated to this common boundary set, Section 2.2). The spatial distribution of county-level drought-event frequency and mean deseasonalized damage is shown in Figure 1 (county-centroid scatter; no administrative boundaries).

2.2. Data

This study uses six county-level panel datasets, all derived from publicly available academic data products, covering the 2000–2023 study period. Raster products are uniformly aggregated to county-level administrative units (pixel averages within each county) before use.

Precipitation and potential evapotranspiration (PET). Monthly precipitation grids are taken from the 1 km monthly precipitation dataset for China (1901–2024; the cited data release covers 1901–2023, and the 2024 extension is not used in this study) released by Peng Shouzhang via the National Tibetan Plateau Data Center [24]: a 1 km grid generated by Delta spatial downscaling from the CRU TS 0.5° reanalysis climate field and the WorldClim high-resolution climatology, validated against precipitation observations from 496 independent meteorological stations [25]. Monthly PET grids are taken from the companion 1 km monthly PET product of the same series (National Tibetan Plateau Data Center) [26], estimated from the companion 1 km monthly temperature dataset using the Hargreaves equation [27]. We clip both products to the 13 major grain-producing provinces

for 2000–2023 and aggregate them to county-level monthly series by arithmetic averaging of pixels within each county boundary (1546 counties × 288 months), with no additional spatial smoothing.

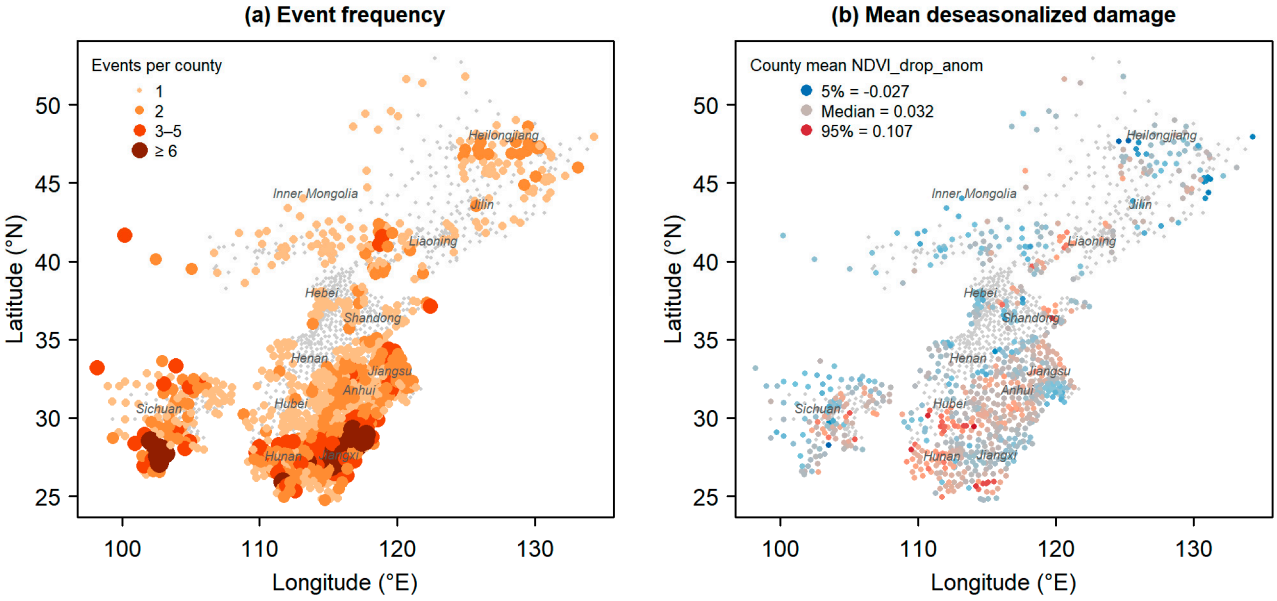

Figure 1. Spatial distribution of county-level drought-event frequency (**a**) and mean deseasonalized damage NDVI_drop_anom (**b**). County-centroid scatter without administrative boundaries; gray points are counties without events; province names are orientation annotations only; 1522 of 1546 counties (98.4%) have coordinate records.

Vegetation index. NDVI is taken from the China regional 250 m Normalized Difference Vegetation Index dataset (2000–2023, National Tibetan Plateau Data Center) [28]. Produced by the Satellite Application Center for Ecology and Environment, Ministry of Ecology and Environment, based on Aqua/Terra MODIS MOD13Q1 250 m/16-day NDVI and pixel reliability data, the product pre-processes noisy pixels of the same land-cover type in each image, applies long time-series Savitzky–Golay filtering to retain high-quality pixels, and then composites the 16-day products into monthly maximum value compositing (MVC) products covering the whole country [28]; the monthly MVC effectively suppresses cloud and atmospheric contamination. The valid pixel range is −0.2 to 1, scaled by 10^4 (raw integer values −2000 to 10,000; fill value −3000). We aggregate pixel-level monthly values to county-level monthly NDVI by arithmetic averaging within county boundaries and rescale by dividing by 10,000 to obtain standard NDVI. Pre-processing effects on short drought dips are assessed in Section 4.2.

Cropland mask and built-up fraction. Annual 30 m land cover is taken from the China Land Cover Dataset (CLCD), which provides nationwide classification at 30 m resolution from Landsat imagery processed on the Google Earth Engine platform [29]. We use the cropland class to restrict the NDVI aggregation to cropland pixels (Section 2.3.5) and to derive each county's built-up (impervious) fraction, used as a control variable to test whether nighttime light effects are confounded with urban land (Section 3.3). The CLCD maps cover 2000–2022; the 2022 map is applied to 2023.

Nighttime light. NTL is taken from the 1986–2024 VIIRS-like nighttime light reconstruction dataset constructed by Tian et al. (2025) [30] (arXiv:2508.00590): the modeling framework reconstructs historical DMSP-OLS observations into a radiance series compatible with VIIRS, eliminating the cross-sensor break around 2013; the county-aggregated panel

provides mean, min, max, sum, and sd pixel statistics. We use the county-level annual mean brightness (mean statistic) and take the natural logarithm ln(mean + 1) in the analysis.

Crop yields. Grain and oilseed output are taken from the digitized panel of the China Statistical Yearbook (County Level) (compiled by the National Bureau of Statistics; China Statistics Press) [31–54]. We emphasize that the yearbook variable is total production (tons), not yield per hectare; an area-normalized yield proxy (production per unit of cropland area) is constructed in Section 2.3.7. Note: Each volume reports data for the preceding year (e.g., the 2024 volume reports 2023 data). County-level grain production data are available through the 2017 volume (covering 2016 data) and are discontinued in subsequent volumes; oilseed production, by contrast, is reported in the 2001–2024 volumes (data years 2000–2023, fully covered without interruption). The data also contain 68 county-level socioeconomic indicators—including administrative land area and year-end population—used for the crop-category validation of the damage metrics (grain and oilseeds; see Section 2.3.7). Under the official Chinese statistical definitions, "grain" is an aggregate output across multiple crops, including cereals such as rice, wheat, and maize, pulses (including soybean), and tubers (fresh tubers converted to grain at a 5:1 ratio); "oilseeds" is likewise an aggregate, covering herbaceous oil crops such as peanuts, rapeseed, sesame, sunflower seed, and flaxseed, excluding soybean (counted as grain) and woody oil crops (e.g., camellia seed). The 13 major grain-producing provinces also cover China's two major oilseed belts—the Yangtze River winter rapeseed belt (Hubei, Hunan, Jiangxi, Anhui, Jiangsu, Sichuan) and the spring-sown oilseed belt of the Huang–Huai–Hai region and Northeast China (Hebei, Shandong, Henan, Liaoning)—so no change to the study area is required.

Administrative boundaries. County-level boundaries and codes are taken from the public-edition 1:1,000,000 fundamental geographic information data (Ministry of Natural Resources, 2021) [55], updated to the 2020 version of the codes using the Ministry of Civil Affairs' published changes in administrative divisions at county level and above, totaling 1546 counties. All source products are aggregated to this common 1546-county, 2020-code boundary set; differences in the vintage of boundary attachments bundled with individual products were reconciled to this set.

Data processing and statistical analysis were performed in R 4.5.2 [56] using the data.table 1.18.4 [57], fixest 0.14.2 [58], haven 2.5.5 [59], survival 3.8.3 [60], and ranger 0.18.0 [61] packages; the cropland-masked NDVI aggregation was implemented in Google Earth Engine [62].

2.3. Methods

2.3.1. Drought Index

The monthly water balance is computed for each county:

$$D_i = P_i - PET_i \quad (1)$$

where P_i and PET_i are precipitation and potential evapotranspiration (mm) in month i. The D series is accumulated over a 3-month sliding window (including the current month); a three-parameter log-logistic distribution $F(\cdot)$ is fitted to each county-month series by the method of L-moments, and the standardized precipitation evapotranspiration index at the 3-month scale (SPEI-3) is obtained via the inverse standard normal transform:

$$SPEI3_i = \Phi^{-1}(F(D_i^{(3)})) \quad (2)$$

where $D_i^{(3)}$ is the 3-month accumulated water balance. The 3-month scale is chosen to match the typical duration of root-zone water memory of crops. The SPEI-3 series were precomputed by the data provider from the above 1 km precipitation and PET grids, following the SPEI method of Vicente-Serrano et al. [6], with distribution parameters estimated by L-moments [63].

2.3.2. Drought Event Identification

A drought event is defined as a run of months satisfying two conditions: Condition 1, SPEI-3 ≤ -1.5 (severe drought) for at least 2 consecutive months; Condition 2, no drought month (SPEI > -1) in the 6 months preceding onset. The six-month drought-free requirement guarantees that consecutive events in the same county are separated by at least six non-drought months, ensuring mutual independence of events—a stricter and more explicit independence guarantee than a minimum inter-event gap, because it also separates the pre-event baselines. A sensitivity analysis relaxes or tightens this window to 3 or 12 months: the event count is unchanged at 1433 with the 3-month window and decreases to 1079 with the 12-month window, and the core Model B coefficients remain stable (duration 0.0057–0.0083; peak intensity 0.0379–0.0417; baseline greenness 0.0323–0.0506; Supplementary Table S3). A stricter threshold of SPEI-3 ≤ -2 yields only 29 events satisfying the independence constraint; we therefore adopt -1.5, with threshold sensitivity analyses reported in the Supplementary Materials. For each event, the NDVI series is extracted from 6 months before onset to 24 months after termination.

2.3.3. Damage and Recovery Metrics

Two damage metrics are constructed. The raw metric:

$$NDVI_drop = median(NDVI_{t-6}, \ldots, NDVI_{t-1}) - min(NDVI_{t0}, \ldots, NDVI_{t1}) \quad (3)$$

where t0 and t1 are the onset and termination months of the event; the baseline is the median NDVI of the six pre-onset months (requiring at least 3 valid months; NDVI must be complete within the event window). The deseasonalized metric proceeds in two steps. First, given the overall greening trend of Chinese vegetation over the study period (county-month mean trend slope +0.0032 NDVI/year), a linear trend is fitted and removed separately for each county and month:

$$NDVI_dt(c, m, y) = NDVI(c, m, y) - \beta(c, m) \cdot (y - \bar{y}) \quad (4)$$

where $\beta(c, m)$ is the linear trend slope for county c and month m over 2000–2023, and $\bar{y}$ is the mean sample year; the trend term vanishes at the group-mean year, so detrending does not alter the long-run mean greenness of that county-month. Second, anomalies are taken relative to the detrended climatology:

$$NDVI_anom(c, m, y) = NDVI_dt(c, m, y) - mean_y[NDVI_dt(c, m, y)] \quad (5)$$

NDVI_drop_anom is then constructed from NDVI_anom in the same way as in Equation (3); throughout the paper we refer to this quantity as the detrended, deseasonalized vegetation anomaly. A non-detrended, full-period climatology anomaly is retained as a sensitivity comparison (Supplementary Materials).

Three sensitivity constructions test the robustness of the metric itself; full results are reported in Supplementary Table S7. Seasonal baseline. Because the six-month pre-onset baseline can retain seasonal variation, we recompute the deseasonalized drop using a seasonally matched baseline (the same six calendar months of the previous year, t−18...t−13; 1376 events have a complete previous-year window). This variant correlates at only 0.679 (Spearman) with the standard construction, so it is a noticeably different metric rather than a cosmetic redefinition. Its Model B coefficients remain positive and significant for duration (0.0162, $p = 0.030$) and marginal for peak intensity (0.0471, $p = 0.053$), while the baseline-greenness coefficient loses significance (0.0171, $p = 0.548$)—consistent with the coupling analysis in Section 3.3, in which part of the baseline effect reflects seasonal phase alignment rather than biology. The growing-season grain-yield correlation remains positive (r = 0.165),

preserving the sign and approximate magnitude of the main metric (r = 0.214), although no county-mean construction is significant under province-clustered bootstrap (all $p > 0.3$; Table S7)—consistent with the main yield-validation result that the province-cluster-robust yield signal of the deseasonalized metric emerges only after cropland masking. Quadratic detrending. Replacing the linear trend in Equation (4) with a quadratic trend changes the anomaly metric only marginally (correlation 0.947 with the linearly detrended version) and leaves the Model B coefficients essentially unchanged (duration 0.0058, $p = 0.019$; peak intensity 0.0295, $p = 0.059$; baseline greenness 0.0487, $p = 0.014$; vs. 0.0057, 0.0379, and 0.0506 for the corresponding main-specification coefficients), indicating that linear detrending adequately captures vegetation change over 2000–2023; the growing-season yield correlation is likewise preserved (r = 0.158, vs. 0.214 for the main metric; Table S7). Consistent-window construction. For the coupling diagnostics in Section 3.3, both the damage metric and the baseline predictor are recomputed on non-overlapping windows.

Recovery half-life T_{50}. The recovery target is defined as

$$target = baseline - 0.5 \times (baseline - min) \quad (6)$$

Starting from the month after drought termination, T_{50} is the number of months until NDVI first reaches the Equation (6) target; events not recovering within the 24-month window are right-censored at 24 months and handled with survival models (Section 2.3.6) in addition to the complete-case regressions.

2.3.4. Predictors

Predictors include: drought duration (number of consecutive months with SPEI-3 ≤ -1.5 within the event); peak intensity (absolute value of the lowest monthly SPEI-3 within the event); pre-drought NDVI (the baseline above); seasonal timing (trigonometric encoding of onset month m: $\sin(2\pi m/12)$ and $\cos(2\pi m/12)$); and the nighttime light baseline

$$NTL_base = mean[ln(NTL_{y-1} + 1), ln(NTL_{y-2} + 1), ln(NTL_{y-3} + 1)] \quad (7)$$

where y is the onset year, requiring at least 2 valid years. In the confounding check of Section 3.3, the county built-up fraction (impervious share of county land area from CLCD) is added as a control.

2.3.5. Regression Inference, Machine Learning, and Robustness

With events as the unit of observation, the linear response of damage to the predictors is estimated as

$$Y = \beta_0 + \beta_1 \cdot Duration + \beta_2 \cdot Peak + \beta_3 \cdot NDVI_pre + \beta_4 \cdot sin(2\pi m/12) + \beta_5 \cdot cos(2\pi m/12) + \beta_6 \cdot NTL_base + \varepsilon \quad (8)$$

where the dependent variable Y is NDVI_drop and NDVI_drop_anom, and ε is the error term. Standard errors are clustered at the province level; given only 13 provinces (few clusters), all core conclusions are verified with wild cluster bootstrap-t tests (Rademacher weights, 999 replications) [64] and further checked against province × start-year two-way clustering and Conley spatial HAC (200 km cutoff; 100–500 km sensitivity). Robustness analyses include: a growing-season (April–September onset) subsample estimated on the deseasonalized metric; a county fixed-effects plus start-year fixed-effects model; a province × start-year interaction fixed-effects model; regional stratified models for the four macro-regions (Northeast, North China, Central China, South); sensitivity to the SPEI threshold (±0.2); an overlap sample sensitivity analysis excluding events followed by a new event in the same county within 12 months; Benjamini–Hochberg FDR correction for multiple testing; and county-grouped 5-fold cross-validation (events of the same county never span

the training and test sets, avoiding information leakage), with out-of-sample R^2 measured as $1 - SSE/SST$.

Machine learning benchmark. A random forest regression is fit on the same predictors and the same county-grouped folds (ranger implementation [61]: 500 trees, mtry = 2, minimum node size 5, impurity-based variable importance; seeds are fixed and reported in the code).

Mathematical coupling and window-length diagnostics. Because the pre-drought baseline enters both the damage metric and the predictor set, we re-estimate the damage models with (i) a non-overlapping baseline from months t−12…t−7 and (ii) a consistent-window construction in which the predictor baseline matches the window used in the outcome. To test whether longer events mechanically reach lower minima, we correlate event duration with the raw drop, and we correlate each event-window metric with a variant computed on a fixed 12-month window.

Forward-in-time validation. Because the full-period climatologies in Equations (4) and (5) use observations after an event, the deseasonalized metric is retrospective. We therefore recompute it using only information available before each event (forward anomaly: trends and climatologies estimated from pre-event years only) and re-estimate the damage model and the county-grouped cross-validation on this real-time-feasible variant, separating spatial transferability from operational predictability.

Cropland masking. All county-level analyses are repeated on a cropland-restricted NDVI panel. For each county and year, only MODIS pixels classified as cropland by CLCD in that year (2022 map for 2023) are retained, after the same quality screening and monthly maximum value compositing as the source product; county-month values are arithmetic means over cropland pixels, and all event metrics are reconstructed on this panel. The aggregation is implemented in Google Earth Engine [39]; the median cropland pixel count per county-year is 717 (P5–P95: 73–1361).

2.3.6. Recovery-Time Survival Analysis

Because 519 events (36.2%) under the raw metric and 270 (18.8%) under the deseasonalized metric—percentages of all 1433 events—are censored at 24 months without recovery, and a further 4.6% and 0.8% are administratively censored at the study end, complete-case regressions of T_{50} are biased. Recovery times are therefore analyzed with right-censored survival methods [60,65,66], distinguishing three outcome classes: events observed to recover (848 raw-metric, 1152 deseasonalized), events censored at 24 months without recovery (519 and 270), and events administratively censored at the study end (66 and 11; of the 235 events terminating in 2022–2023, the remainder recovered within their shorter follow-up). The modeled event is recovery (reaching the target), so Cox hazard ratios above 1 indicate faster target attainment. We fit Cox proportional-hazards models [67] for both metrics and test the proportional-hazards assumption with Schoenfeld residuals; where the assumption is rejected, Weibull accelerated failure time (AFT) models are reported as primary, with time ratios (TR) interpreted as multiplicative effects on recovery time (TR > 1 indicates slower recovery).

2.3.7. Crop-Category Validation of the Damage Metrics (Grain and Oilseeds)

The construct validity of the damage metrics is examined against county-level crop production, separately for the grain and oilseed categories (both are aggregate outputs of multiple crops; crop composition is defined in Section 2.2). Grain. Each event is first assigned to a harvest year by crop calendar: events starting in January–September are assigned to the same year's harvest (summer maize, single-season rice), and events starting in October–December to the following year's harvest (winter wheat, sown in October–

November and harvested the following June). This assignment replaces the previous convention of always using the year after event termination, which misaligned summer droughts with the following harvest. The production change rate relative to the pre-disaster three-year mean is:

$$loss = [mean(Y_{h-3}, Y_{h-2}, Y_{h-1}) - Y_h]/mean(Y_{h-3}, Y_{h-2}, Y_{h-1}) \times 100\% \quad (9)$$

where Y denotes production (tons), and h is the crop calendar-matched grain harvest year, requiring valid records in at least two of the three pre-disaster years and in the harvest year. Oilseeds. Oilseed validation retains the crop calendar matching of the original manuscript: winter rapeseed (the main oilseed in southern and central counties) is sown in October and harvested in April–May of the following year, with critical water-demand periods at the budding stage (February–April) and the seedling stage (September–November); spring-sown oilseeds (in northern and northeastern counties) are sown in April and harvested in September–October, with the critical period at flowering and pegging (June–August). Each drought event is mapped to the harvest year whose window intersects the oilseed growing period: for southern and central counties, an event window intersecting February–April is assigned to the current year's harvest and one intersecting September–November to the following year's harvest, while events with summer windows (no oilseed in the field) are dropped; for northern and northeastern counties, an event window intersecting April–September is assigned to the current year's harvest, and winter events are dropped. The loss rate is then defined by the harvest year as in Equation (9), with h the oilseed harvest year.

Detrended loss rate. For both categories, a linear trend is fitted to production by county (with at least 6 valid years), and the detrended loss rate compares the observed harvest-year production with the trend-expected value:

$$loss_dt = [\hat{Y}(h) - Y(h)]/\hat{Y}(h) \times 100\% \quad (10)$$

where $\hat{Y}(h)$ is the county-year linear trend fit of production, and $Y(h)$ is the observed harvest-year production. Positive values indicate production below trend. (In the original manuscript, the trend fit erroneously replaced the observed post-event production in the numerator, so that the detrended loss contained no harvest-year information; Equation (10) corrects this.)

Area-normalized yield proxy. Because the yearbook variable is total production rather than yield per hectare, we construct a yield proxy as production divided by the county's CLCD cropland stock (tons per hectare; cropland stock includes non-grain crops, so the proxy is an approximation, stated as such), and repeat the validation with proxy-based loss rates. Inference. Spearman rank correlations are used to examine the monotonic association of the two damage metrics with the two categories and the loss-rate definitions (production is affected by trend growth and statistical conventions, so rank correlation is more robust than linear correlation); correlation inference is verified with province-level cluster bootstrap (999 replications, resampling provinces with replacement) to mitigate few-cluster bias. Grain validation covers harvest years 2002–2016 (production reported through data year 2016), N = 426 events; oilseed validation covers all harvest years 2000–2023 (N = 684 events after crop calendar matching), including events corresponding to the 2019 and 2022 major droughts.

3. Results

3.1. Drought Event Characteristics

A total of 1433 independent severe drought events were identified during 2000–2023, involving 902 counties; 59.5% of events began in the growing season (April–September). The regional distribution is 670 events in Central China, 498 in the South, 135 in the Northeast,

and 130 in North China (provincial distribution in Figure 2). Event characteristics are summarized in Table 1.

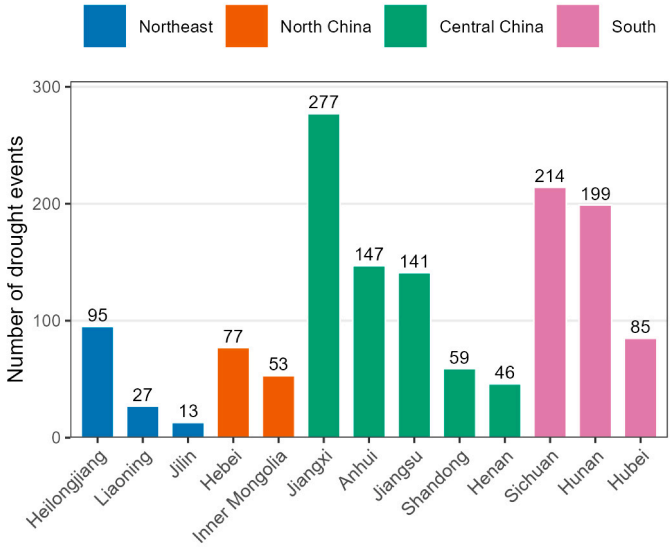

Figure 2. Number of identified severe drought events (SPEI-3 ≤ −1.5, 2000–2023) by province, colored by region (Northeast: Heilongjiang, Jilin, Liaoning; North China: Hebei, Inner Mongolia; Central China: Shandong, Henan, Jiangsu, Anhui, Jiangxi; South: Sichuan, Hunan, Hubei).

Table 1. Descriptive statistics of drought events and key variables.

Variable	Mean	SD	Median	IQR	Min	Max	N
Duration (months)	2.77	1.14	2	2–3	2	10	1433
Peak intensity |SPEI|	1.92	0.20	1.92	1.76–2.07	1.52	3.07	1433
Mean event SPEI	−1.77	0.13	−1.75	−1.86 to −1.67	−2.38	−1.50	1433
Pre-drought baseline NDVI	0.517	0.154	0.543	0.438–0.622	0.062	0.848	1433
NDVI_drop (raw)	0.062	0.190	0.047	−0.069 to 0.198	−0.564	0.571	1433
NDVI_drop_anom (deseasonalized)	0.033	0.044	0.029	0.006–0.057	−0.127	0.258	1433
T_{50} (raw, months)	3.77	—	3	2–5	1	10	848
T_{50}_anom (deseasonalized, months)	2.48	2.32	2	1–3	1	22	1152
Share of growing-season events (%)	59.5	—	—	—	—	—	1433

IQR is the interquartile range. T_{50} statistics cover only events observed to recover within the 24-month window (N = 848 raw, 1152 deseasonalized); non-recovered and administratively censored events are excluded here and retained in the survival analysis.

3.2. Conventional Damage Metrics Are Dominated by the Phenological Cycle

The interquartile range of the raw metric NDVI_drop is −0.069 to 0.198: in more than a quarter of "drought events", NDVI actually increases—this cannot reflect physiological damage and is instead an artifact of event windows falling in the green-up period; after detrending and deseasonalization, the interquartile range narrows to 0.006–0.057. The correlation between the two metrics is only 0.069. Figure 3 illustrates the mechanism with two real events of nearly identical deseasonalized damage but very different raw drops, depending only on the onset month. Group-level composite trajectories by onset season show the same phase artifact (Figure 4). Regression results further confirm the dominance of phenology (Table 2): the model with NDVI_drop as the dependent variable (Model A) has an adjusted R^2 = 0.77, in which the onset-month cosine term (coefficient 0.215, t = 15.6,

$p < 0.001$) contributes the vast majority of explanatory power; with NDVI_drop_anom as the dependent variable (Model B), the model R^2 falls to 0.18 and the cosine term is no longer significant. Five-fold cross-validation shows an out-of-sample $R^2 = 0.77$ for the raw metric (seasonal phase is spatially transferable) but only 0.18 (linear) for the deseasonalized damage; a random forest on the same folds achieves $R^2 = 0.44$—nonlinearity and interactions more than double predictability, yet about 56% of the damage variance remains unpredictable (see Section 4.4).

Retrospective vs. real-time predictability. Because the full-period climatologies in Equations (4) and (5) use observations after an event, the deseasonalized metric is retrospective, and county-grouped cross-validation assesses spatial transfer rather than operational predictability at drought onset. When we recompute the anomaly with pre-event information only (forward anomaly; N = 1274 events), the forward metric correlates at 0.827 with the full-period metric; the forward-specification damage model shows duration clearly significant (0.0089, $p = 0.002$) and peak intensity not significant ($p = 0.11$), and the county-grouped out-of-sample R^2 (0.224) is, if anything, higher than the retrospective value (0.179). The conclusion that standard severity variables identify the direction and broad magnitude—but not the precise level—of damage therefore holds, and is, if anything, stronger, for real-time-feasible implementation.

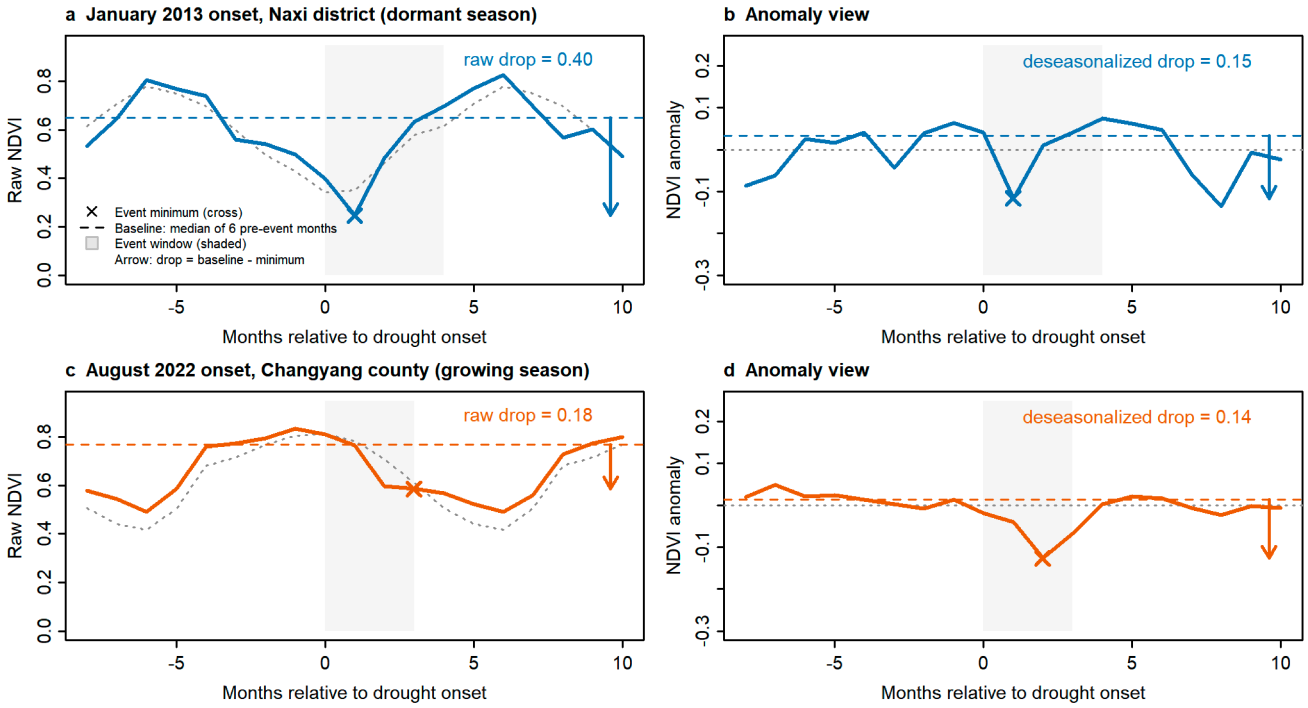

Figure 3. The phenological artifact in the raw damage metric, illustrated by two real drought events with nearly identical deseasonalized damage (0.149 vs. 0.140). (**a**,**b**) January 2013 onset in Naxi District, Sichuan (dormant season, 5 months): the raw drop is as large as 0.40; (**c**,**d**) August 2022 onset in Changyang County, Hubei (growing season, 4 months): the raw drop is only 0.18. The discrepancy arises entirely from the seasonal cycle, not from drought damage. Shaded bands mark event duration; dotted line = monthly climatology (**a**,**c**) or zero anomaly (**b**,**d**); dashed line = baseline (median of the six pre-event months); cross = event minimum.

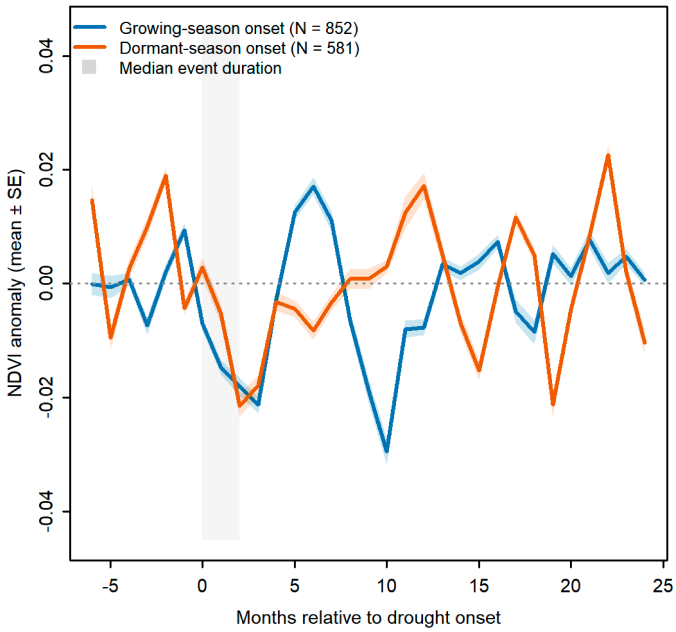

Figure 4. Composite NDVI anomaly trajectories relative to drought onset (months −6 to +24; mean ± SE), separately for growing-season (N = 852) and dormant-season (N = 581) onsets. The shaded band marks the median event duration. The dip during the event window is modest and is followed by rebounds and oscillations of comparable amplitude (Section 4.5), indicating that damage signals in deseasonalized NDVI are subtle relative to interannual variability.

Table 2. Predictors of the damage metrics: raw specification (Model A), deseasonalized specification (Model B), and growing-season subsample on the deseasonalized metric (Model C).

Variable	Model A: NDVI_drop	Model B: NDVI_drop_anom	Model C: NDVI_drop_anom (Growing Season)
Duration	0.037 *** (0.010)	0.006 + (0.003)	0.0084 ** (0.0023)
Peak intensity	0.004 (0.028)	0.038 *** (0.012)	0.0228 (0.0151)
Pre-drought NDVI	0.291 *** (0.069)	0.051 ** (0.018)	0.0737 * (0.0288)
$\sin(2\pi m/12)$	−0.010 (0.014)	−0.011 *** (0.003)	−0.0064 (0.0038)
$\cos(2\pi m/12)$	0.215 *** (0.014)	−0.010 (0.007)	−0.0076 (0.0236)
NTL baseline	0.018 (0.011)	0.003 (0.004)	0.0043 (0.0050)
Constant	−0.204 ** (0.073)	−0.085 *** (0.016)	−0.072 (0.041)
N	1433	1433	852
Adjusted R^2	0.772	0.180	0.163
Out-of-sample R^2 (5-fold, county-grouped)	0.772	0.179	—

Province-clustered standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$. Wild cluster bootstrap (999 replications) verification for Model B: peak intensity $p = 0.020$, pre-drought NDVI $p = 0.012$, sine term $p < 0.001$; duration is marginal under clustered standard errors ($p = 0.054$) and not significant under wild cluster bootstrap ($p = 0.148$). Model C is estimated on the deseasonalized metric as requested in the review; the raw-metric growing-season specification (adjusted $R^2 = 0.639$, onset-month cosine 0.284***, NTL 0.035*) is reported in Supplementary Table S4. Benjamini–Hochberg FDR correction across all 42 tests in Tables 2 and 3: in Model B, the sine term (q = 0.06) and peak intensity (q = 0.08) remain significant, and baseline greenness is marginal (q = 0.12) at the 10% FDR level; duration in Model C (q = 0.05) and in the South regional model (q = 0.01) is significant (Supplementary Table S5).

Table 3. Regional deseasonalized damage models (dependent variable: NDVI_drop_anom; province-clustered standard errors).

Variable	Northeast (N = 135)	North China (N = 130)	Central China (N = 670)	South (N = 498)
Duration	0.006 (0.003)	0.007 (0.009)	−0.002 (0.011)	0.010 *** (0.0003)
Peak intensity	0.013 (0.019)	−0.041 (0.011)	0.033 (0.017)	0.025 (0.011)
Pre-drought NDVI	0.100 (0.047)	0.026 (0.036)	0.043 (0.034)	0.137 (0.047)
sin(2πm/12)	0.016 (0.023)	0.006 (0.005)	−0.014 + (0.006)	−0.008 (0.004)
cos(2πm/12)	−0.048 * (0.010)	−0.029 * (0.002)	0.003 (0.009)	−0.019 (0.007)
NTL baseline	0.012 + (0.004)	0.005 (0.002)	−0.004 (0.008)	0.015 (0.008)
Adjusted R^2	0.259	0.308	0.164	0.279

The dependent variable is NDVI_drop_anom; province-clustered standard errors in parentheses. *** $p < 0.001$, * $p < 0.05$, + $p < 0.1$. Northeast: Heilongjiang, Jilin, Liaoning; North China: Hebei, Inner Mongolia; Central China: Jiangsu, Anhui, Jiangxi, Shandong, Henan; South: Hubei, Hunan, Sichuan. The Northeast and South contain only three provinces each and North China two; clustered standard errors are of limited use at this scale, and the regional models are presented as exploratory (heteroskedasticity-robust comparisons in the text).

3.3. Predictors of Deseasonalized Damage, Coupling Diagnostics, and Fixed Effects

Model B shows that each 1-unit deepening of peak intensity increases damage by 0.038 (clustered $p < 0.01$; wild cluster bootstrap $p = 0.020$), making it the most robust predictor; under the standard construction the pre-drought NDVI coefficient is 0.051 (clustered $p < 0.01$; wild bootstrap $p = 0.012$), so higher baseline greenness implies greater damage; the duration coefficient is 0.006, positive in sign, marginal under clustered standard errors ($p = 0.054$) and not significant under wild cluster bootstrap ($p = 0.148$), recovering significance in the growing-season subsample (0.0084, $p = 0.005$) and under the larger event sample of the SPEI-3 ≤ -1.3 specification ($p = 0.016$; sensitivity analysis in Section 4.5)—the independent contribution of duration is weaker than that of peak intensity, and its significance is sample-dependent. The onset-month sine term (−0.011, $p < 0.001$) survives deseasonalization, indicating that droughts with onset in spring cause less damage than autumn-onset droughts—pointing to real physiological sensitivity to crop phenological stages rather than a metric artifact. The nighttime light baseline is not significant ($p = 0.61$).

Mathematical coupling. Because the pre-drought baseline enters both the outcome and the predictor set, part of the baseline coefficient is mechanical. In Model A, the baseline coefficient is 0.290 ($p < 0.001$) under the standard overlapping construction; replacing the predictor with a non-overlapping baseline (months t−12...t−7) reverses the sign to −0.219 ($p = 0.015$); and under a consistent-window construction in which predictor and outcome use the same baseline definition, it is 0.103 ($p = 0.019$). Roughly two-thirds of the raw baseline coefficient therefore reflects construction rather than biology. In Model B, the baseline coefficient is smaller and stable in sign across all three constructions (0.051–0.078), although a mechanical component cannot be excluded; we accordingly interpret the baseline-greenness effect as suggestive of a "biomass-at-risk" mechanism [14,15] rather than as conclusive evidence. Greener counties do have larger absolute anomaly amplitudes, and the forward-anomaly robustness check (Section 3.2) leaves the qualitative predictor ranking unchanged. Under the seasonally matched baseline variant (Supplementary Table S7), the baseline-greenness coefficient loses significance altogether, reinforcing the suggestive (rather than established) status of the biomass-at-risk reading.

Window length. Longer events provide more opportunities to reach a low minimum, which could manufacture a duration–damage association. In practice, duration correlates with the raw drop at only r = −0.042, and each metric computed on a fixed 12-month window correlates at 0.924 (raw) and 0.843 (deseasonalized) with its event-window version: mechanical minimum-sampling does not drive the duration or damage results.

Fixed effects and confounding checks. In the county fixed-effects plus start-year fixed-effects model (N = 890, adjusted $R^2 = 0.39$), drought duration (0.007, $p = 0.017$) and pre-drought NDVI (0.119, $p = 0.041$) remain significant, while peak intensity and nighttime light are not significant. Duration thus acts mainly through within-county, within-event variation, strengthening the "longer deficit → greater damage" mechanism. For nighttime light, adding the county built-up fraction as a control (median 7.39% of county land area) flips the (already insignificant) NTL coefficient from +0.003 to −0.004, and the built-up coefficient itself is +0.046 ($p = 0.11$): the sign instability of NTL across specifications is noise rather than structure, and we refrain from interpreting it as an adaptive-capacity effect (Section 4.3).

3.4. Regional Heterogeneity

The regional deseasonalized models (Table 3) reveal qualitative differences in damage mechanisms. The South is dominated by drought duration (0.010, $p < 0.001$), consistent with the region's multiple-cropping and long growing-season structure; the Northeast is characterized by the onset-month cosine term (−0.048, $p = 0.04$) and nighttime light (0.012, $p = 0.089$, marginal), with duration consistent in sign (0.006) but not significant; in Central China only the onset-month sine term is marginally significant (−0.014, $p = 0.085$), with the lowest explanatory power (adjusted $R^2 = 0.16$); in North China the cosine term is significant (−0.029, $p = 0.04$) and the peak intensity coefficient is negative. Inference caveat. The Northeast and South regions comprise only three provinces each, so province-clustered inference is unreliable at regional scale; heteroskedasticity-robust standard errors give the same qualitative picture (e.g., South duration 0.010, $p < 0.001$; Northeast nighttime light 0.012, $p = 0.047$; Northeast duration 0.006, $p = 0.001$), and North China contains only two provinces. We therefore treat the regional models as exploratory and refrain from significance interpretation of individual regional coefficients. A region-specific growing-season sensitivity (winter wheat-based definition for the Huang–Huai–Hai counties of Hebei, Shandong, and Henan; rice-based definitions for the southern provinces) leaves the coefficient pattern and the core conclusions unchanged for both metrics—in the deseasonalized growing-season model, duration goes from 0.0084 (N = 852) to 0.0072 (N = 983) and baseline greenness from 0.074 to 0.083, with significance intact (Supplementary Table S6). The regional adjusted R^2 values (0.16–0.31) are far below the spuriously high explanatory power of the raw metric, and the regional differences in coefficient structure support the necessity of region-specific model calibration.

3.5. Recovery Half-Life: Artifact and Survival Analysis

Regression of raw T_{50} on the predictors (complete cases, N = 848) yields an adjusted $R^2 = 0.67$, but it is dominated by the onset-month sine term (−2.77, t = −15.3, $p < 0.001$): events with onset in spring "recover" fastest—a green-up artifact (Table 4). The deseasonalized T_{50}_anom model (N = 1152) has $R^2 = 0.19$, and the sine term is no longer significant ($p = 0.49$); the cosine term remains significant (1.167, $p < 0.001$), but its explanatory power is far smaller than the fake "recoverability" manufactured by the sine term in the raw specification. Because censoring differs between the two metrics, the correlation between them is computed on the same-event subsample with both recoveries observed (N = 687): r = −0.032 (Pearson) and −0.015 (Spearman)—the two metrics measure almost entirely different objects. Moreover, of all 1433 events, 36.2% do not reach the recovery target within 24 months under the raw metric, versus 18.8% after deseasonalization (a further 4.6% and 0.8%, respectively, are administratively censored at the study end), showing that the raw metric mistakes seasonal lows for non-recovery.

Table 4. Phenological artifact in the recovery half-life: complete-case regressions of raw T_{50} and T_{50}_anom.

Variable	Raw T_{50} (N = 848)	T_{50}_anom (N = 1152)
Duration	−0.688 *** (0.136)	−0.335 *** (0.069)
Peak intensity	−0.063 (0.754)	−0.854 (0.672)
Pre-drought NDVI	−0.183 (0.882)	−1.156 + (0.538)
sin(2πm/12)	−2.773 *** (0.181)	−0.144 (0.205)
cos(2πm/12)	−0.229 (0.293)	1.167 *** (0.198)
NTL baseline	−0.106 (0.117)	0.032 (0.143)
Adjusted R^2	0.670	0.194

*** $p < 0.001$, + $p < 0.1$. Non-recovered events are excluded here and retained in the survival models of Table 5: censored at 24 months (519 raw, 36.2% of all events; 270 deseasonalized, 18.8%) and administratively censored at the study end (66, 4.6%; 11, 0.8%). Same-event correlation between the two recovery metrics (N = 687): r = −0.032.

Table 5. Survival analysis of recovery time (all 1433 events; right-censored at 24 months or at the study end, whichever comes first).

Variable	Raw T_{50}: Cox HR (Descriptive)	Raw T_{50}: Weibull AFT TR	T_{50}_anom: Cox HR	T_{50}_anom: Weibull AFT TR
Duration (per month)	1.670 *** (1.550–1.799)	0.706 *** (0.670–0.744)	1.225 *** (1.153–1.302)	0.832 *** (0.753–0.919)
Peak intensity	2.128 *** (1.483–3.053)	0.503 *** (0.378–0.670)	1.876 *** (1.369–2.570)	0.293 *** (0.180–0.477)
Pre-drought NDVI	17.169 *** (9.511–30.993)	0.129 *** (0.080–0.209)	2.160 ** (1.241–3.757)	0.249 *** (0.106–0.587)
sin(2πm/12)	1.099 (0.960–1.259)	1.366 *** (1.220–1.530)	0.758 *** (0.673–0.854)	1.891 *** (1.587–2.252)
cos(2πm/12)	7.396 *** (6.343–8.624)	0.175 *** (0.161–0.190)	0.725 *** (0.657–0.801)	1.506 *** (1.279–1.774)
NTL baseline	1.403 *** (1.269–1.552)	0.797 *** (0.737–0.863)	1.041 (0.949–1.143)	0.894 (0.770–1.036)
Recovered events (n)	848	848	1152	1152
Censored at 24 months, no recovery (n, %)	519, 36.2%	519, 36.2%	270, 18.8%	270, 18.8%
Administratively censored (n, %)	66, 4.6%	66, 4.6%	11, 0.8%	11, 0.8%
N	1433	1433	1433	1433

HR = hazard ratio (Cox); the modeled event is recovery, so HR > 1 indicates faster target attainment. TR = time ratio (Weibull AFT; TR < 1 indicates faster target attainment); 95% confidence intervals in parentheses. *** $p < 0.001$, ** $p < 0.01$. The proportional-hazards assumption is rejected for both Cox models (Schoenfeld global test $p < 0.001$), so the AFT columns are primary and the raw Cox column is descriptive. Censoring counts and percentages are shares of all 1433 events; administrative censoring applies to the 235 events terminating in 2022–2023 that did not recover within their shorter available follow-up (66 raw, 11 deseasonalized).

Survival analysis accounting for censoring. Complete-case exclusion of censored events biases the comparison, so we fit right-censored survival models on all 1433 events (Table 5), distinguishing recovered events (848 raw, 1152 deseasonalized), events censored at 24 months without recovery (519, 270), and events administratively censored at the study end (66, 11). The modeled event is recovery (reaching the Equation (6) target), so a Cox hazard ratio above 1 means faster target attainment and a Weibull time ratio (TR) below 1 means less time to attainment. Schoenfeld tests reject the proportional-hazards

assumption for both Cox models (global $p < 0.001$), so the Cox columns are reported as descriptive and the Weibull AFT time ratios are the interpretable estimates. For the raw metric, the AFT model confirms overwhelming phase dominance (cosine TR = 0.175, 95% CI 0.161–0.190); the severity variables are likewise associated with faster target attainment (duration TR = 0.71, peak intensity TR = 0.50, baseline greenness TR = 0.13, all $p < 0.001$), but because the raw recovery model is a description of seasonal phase throughout, we do not interpret its individual coefficients and draw no socioeconomic moderation claim from it. For T_{50}_anom, longer duration (TR = 0.83 per month, $p < 0.001$), deeper peak intensity (TR = 0.29, $p < 0.001$), and higher baseline greenness (TR = 0.25, $p = 0.002$) are all associated with faster target attainment, the onset-month sine and cosine terms with slower attainment (TR = 1.89 and 1.51, both $p < 0.001$), and NTL is null (TR = 0.89, $p = 0.14$). Two points follow. First, the direction is not "slower recovery after more severe drought" because the T_{50} target is anchored to the event minimum (Equation (6)); deeper events face lower targets and cross them sooner, so severity variables mechanically accelerate half-recovery attainment; the negative duration coefficients in the complete-case regressions of Table 4 carry the same sign and are consistent with this mechanism rather than in conflict with it. Second, beyond this anchoring, the deseasonalized survival model identifies timing effects (residual seasonal phase) but no socioeconomic moderation, while the raw metric's recovery model remains a description of seasonal phase throughout.

3.6. Crop-Category Yield Validation of the Damage Metrics

Corrected loss definition and grain. Grain validation (N = 426 events with computable loss rates; harvest years 2002–2016, extending the previously validatable window because harvest years are now assigned by crop calendar) yields a pattern that differs fundamentally from the original manuscript (Table 6). The raw metric NDVI_drop is negatively associated with grain losses under every loss-rate definition (e.g., r = −0.287, $p < 0.001$ with detrended losses; province-cluster bootstrap $p = 0.028$), and significantly so in the growing-season subsample (r = −0.362, bootstrap $p = 0.018$). The deseasonalized metric has the opposite (positive) sign but is not significant at the county-mean level (max |r| = 0.214; all province-cluster bootstrap $p > 0.05$). Two candidate mechanisms explain the disappearance of the raw metric's apparent positive signal. First, the positive raw–grain association in the original manuscript was carried by the erroneous Equation (11), in which the trend fit replaced the observed production so that the "loss" contained no harvest-year information; once the actual production enters, the apparent association disappears. Second, the raw metric is mechanically largest in counties and seasons with high NDVI amplitude (Section 3.3), which are not the counties with the largest yield losses: the raw metric embeds absolute greenness and phenological information, not damage. Neither mechanism, however, explains the negative sign of the raw association; a complete decomposition of that sign would require county-level crop composition data that this study does not have, and we therefore report the negative correlation descriptively without claiming a mechanism for it.

Cropland masking and the yield signal of the deseasonalized metric. Restricting NDVI to CLCD cropland pixels (Section 2.3.5; masked metrics correlate with the county-mean versions at 0.928 for the raw drop and 0.700 for the deseasonalized drop) leaves the main damage regressions essentially unchanged (Table 7) but changes the validation: the deseasonalized metric is positively associated with detrended grain losses, and this association is significant and robust to province-clustered inference in the growing-season subsample (r = 0.327, $p < 0.001$; bootstrap $p = 0.008$; Table 8); in the full sample the point estimate is positive but not robust (r = 0.127, bootstrap $p = 0.146$). The area-normalized proxy shows the same pattern in the growing season (r = 0.291, $p < 0.001$). The raw metric, by contrast, remains significantly negative under masking (r = −0.295, bootstrap $p = 0.026$).

The yield signal of greenness damage is therefore real but resides in the deseasonalized anomaly measured over cropland within this boundary: county-mean greenness dilutes the crop signal with non-agricultural vegetation, and the raw metric's apparent signal was an artifact of both its construction and the previous loss definition.

Table 6. Spearman rank correlations between damage metrics (county-mean panel) and crop production losses, corrected loss definition and crop calendar-matched harvest years.

Crop Category	Metric × Loss-Rate Definition	Full Sample	Growing-Season Subsample
Grain (N = 426/249)	NDVI_drop × loss	r = −0.104 * (p = 0.031)	r = −0.165 ** (p = 0.009)
	NDVI_drop × loss_dt	r = −0.287 *** ($p < 0.001$)	r = −0.362 *** ($p < 0.001$)
	NDVI_drop × proxy_loss	r = −0.158 *** (p = 0.001)	r = −0.156 ** (p = 0.014)
	NDVI_drop × proxy_loss_dt	r = −0.185 *** ($p < 0.001$)	r = −0.291 *** ($p < 0.001$)
	NDVI_drop_anom × loss	r = 0.050 (p = 0.299)	r = 0.178 ** (p = 0.005)
	NDVI_drop_anom × loss_dt	r = 0.069 (p = 0.155)	r = 0.214 *** (p = 0.001)
	NDVI_drop_anom × proxy_loss	r = 0.031 (p = 0.528)	r = 0.096 (p = 0.129)
	NDVI_drop_anom × proxy_loss_dt	r = 0.066 (p = 0.175)	r = 0.208 *** (p = 0.001)
Oilseeds (N = 684/524)	NDVI_drop × loss	r = −0.016 (p = 0.683)	r = −0.117 ** (p = 0.007)
	NDVI_drop × loss_dt	r = −0.066 + (p = 0.087)	r = −0.162 *** ($p < 0.001$)
	NDVI_drop_anom × loss	r = −0.026 (p = 0.498)	r = −0.062 (p = 0.154)
	NDVI_drop_anom × loss_dt	r = −0.061 (p = 0.114)	r = −0.124 ** (p = 0.005)

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$. Loss-rate definitions in Section 2.3.7; grain harvest years assigned by crop calendar (265 events to the same year, 161 to the following year); oilseed validation applies the crop calendar matching of Section 2.3.7 (unmatched events dropped). Province-cluster bootstrap (999 replications, resampling provinces with replacement): grain NDVI_drop × loss_dt p = 0.028 (full) and 0.018 (growing); all other grain and oilseed cells $p > 0.05$. The yearbook variable is total production; proxy_loss divides by CLCD cropland stock (tons/hectare).

Oilseeds. Neither metric is positively associated with oilseed losses under any definition, before or after masking (Tables 6 and 8); in the growing season, the associations are significantly negative for the raw metric. This should not be read as evidence that drought monitoring "fails" for oilseeds: oilseeds occupy a minor share of county cropland in most study counties, their planting-area decisions respond to prices and policy, and their phenology (winter rapeseed) is mismatched with the grain-dominated greenness signal. The appropriate inference is narrower: county-aggregate greenness—however processed—cannot yet speak to secondary-category yields without crop distribution maps.

Validity boundary. Taken together, the corrected validation delineates where greenness damage metrics are informative. For the county-dominant crop category (grain), the deseasonalized anomaly measured over cropland carries a positive yield signal that is robust to province-clustered inference within a defined boundary—cropland pixels, grain-dominant counties, and growing-season events; outside that boundary (full-sample events, secondary categories), the signal is not robust, and the raw metric carries no positive yield signal under any specification. For secondary categories (oilseeds), no county-mean metric works. The implication for monitoring practice is direct: operational damage assessment in mixed cropping areas must combine deseasonalized anomalies with crop distribution layers [67,68]—crop classification maps (e.g., ESA WorldCover 10 m, the 30 m China Land Cover Dataset (CLCD)) are mature remote sensing products, and this step involves no technical obstacle.

Table 7. Cropland-masked damage models (dependent variables computed on CLCD cropland pixels; province-clustered standard errors).

Variable	Model A: NDVI_drop_m	Model B: NDVI_drop_anom_m	Model C: NDVI_drop_anom_m (Growing Season)
Duration	0.038 *** (0.009)	0.0078 *** (0.0023)	0.0091 ** (0.0028)
Peak intensity	0.033 (0.034)	0.0389 *** (0.0104)	0.0350 * (0.0128)
Pre-drought NDVI (cropland)	0.449 *** (0.084)	0.065 ** (0.025)	0.012 (0.031)
$\sin(2\pi m/12)$	−0.027 (0.017)	−0.0134 *** (0.0040)	−0.0192 *** (0.0037)
$\cos(2\pi m/12)$	0.194 *** (0.017)	−0.0062 (0.0063)	0.016 (0.011)
NTL baseline	0.019 * (0.009)	0.0011 (0.0046)	−0.0011 (0.0040)
Constant	−0.322 *** (0.071)	−0.090 *** (0.023)	−0.046 (0.043)
N	1408	1408	827
Adjusted R^2	0.791	0.262	0.295

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. N = 1408 county-events with at least one cropland pixel (25 events with negligible cropland coverage are dropped). The masked metrics correlate with the county-mean versions at 0.928 (raw) and 0.700 (deseasonalized); the phenological-artifact conclusions (raw model dominated by the onset-month cosine term, $R^2 = 0.79$ vs. 0.26) are unchanged. Model C is the cropland-masked counterpart of Table 2 Model C.

Table 8. Cropland-masked yield validation (Spearman rank correlations, N as in Table 6).

Crop Category	Metric × Loss-Rate Definition	Full Sample	Growing-Season Subsample
Grain (N = 424/247)	NDVI_drop_m × loss	r = −0.106 * ($p = 0.029$)	r = −0.065 ($p = 0.312$)
	NDVI_drop_m × loss_dt	r = −0.295 *** ($p < 0.001$)	r = −0.290 *** ($p < 0.001$)
	NDVI_drop_m × proxy_loss	r = −0.143 ** ($p = 0.003$)	r = −0.072 ($p = 0.257$)
	NDVI_drop_m × proxy_loss_dt	r = −0.187 *** ($p < 0.001$)	r = −0.185 *** ($p = 0.004$)
	NDVI_drop_anom_m × loss	r = 0.042 ($p = 0.385$)	r = 0.167 ** ($p = 0.009$)
	NDVI_drop_anom_m × loss_dt	r = 0.127 ** ($p = 0.009$)	r = 0.327 *** ($p < 0.001$)
	NDVI_drop_anom_m × proxy_loss	r = 0.002 ($p = 0.971$)	r = 0.043 ($p = 0.503$)
	NDVI_drop_anom_m × proxy_loss_dt	r = 0.130 ** ($p = 0.007$)	r = 0.291 *** ($p < 0.001$)
Oilseeds (N = 671/511)	NDVI_drop_m × loss	r = −0.043 ($p = 0.268$)	r = −0.158 *** ($p < 0.001$)
	NDVI_drop_m × loss_dt	r = −0.146 *** ($p < 0.001$)	r = −0.276 *** ($p < 0.001$)
	NDVI_drop_anom_m × loss	r = −0.103 ** ($p = 0.008$)	r = −0.140 ** ($p = 0.001$)
	NDVI_drop_anom_m × loss_dt	r = −0.094 * ($p = 0.015$)	r = −0.161 *** ($p < 0.001$)

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Province-cluster bootstrap (999 replications): grain NDVI_drop_anom_m × loss_dt $p = 0.146$ (full) and 0.008 (growing season); grain NDVI_drop_m × loss_dt $p = 0.026$ (full); oilseed NDVI_drop_m × loss_dt $p = 0.034$ (growing season); all other cells $p > 0.05$. The negative oilseed correlations should not be read as monitoring failure: oilseeds are a secondary category in most counties, and their yields are driven by planting-area decisions and sub-county microclimate (Section 4.5).

4. Discussion

4.1. Phenological Artifacts: What Do Conventional Damage Metrics Actually Measure?

The core methodological finding of this study is that damage metrics based on absolute vegetation-index levels systematically mix damage signals with seasonal phase. The raw NDVI_drop correlates with the detrended, deseasonalized metric at only 0.07, suggesting that conclusions drawn from this construction may largely reflect phenology. Absolute-level constructions such as the raw baseline-minus-minimum difference embed the seasonal cycle by construction. We use this construction as an illustrative example of a broader class of absolute-level drop metrics; our argument does not depend on the prevalence of this particular formula. Seasonal-handling challenges in such analyses are documented in seasonal-change detection and productivity estimation studies [12,16,69]. For example, CASA-type productivity estimates that do not account explicitly for the seasonal cycle can be biased in the way our results predict [69]. The mechanism is intuitive: both the six-month pre-event baseline and the event-window NDVI are embedded in the seasonal curve, so their difference naturally carries phase information related to the onset month. NDVI "increases" during winter-onset events, and spring-onset events "recover" fastest; these are different expressions of the same artifact [12,16] (Figure 5). Deseasonalization must be combined with detrending: the +0.0032 NDVI/year greening trend over the study period would cause anomalies relative to a full-period climatology to retain low-frequency trend components, contaminating damage estimates and cross-period comparisons. Detrended deseasonalization removes both biases at minimal cost and should be incorporated as a routine step in agricultural drought impact assessment.

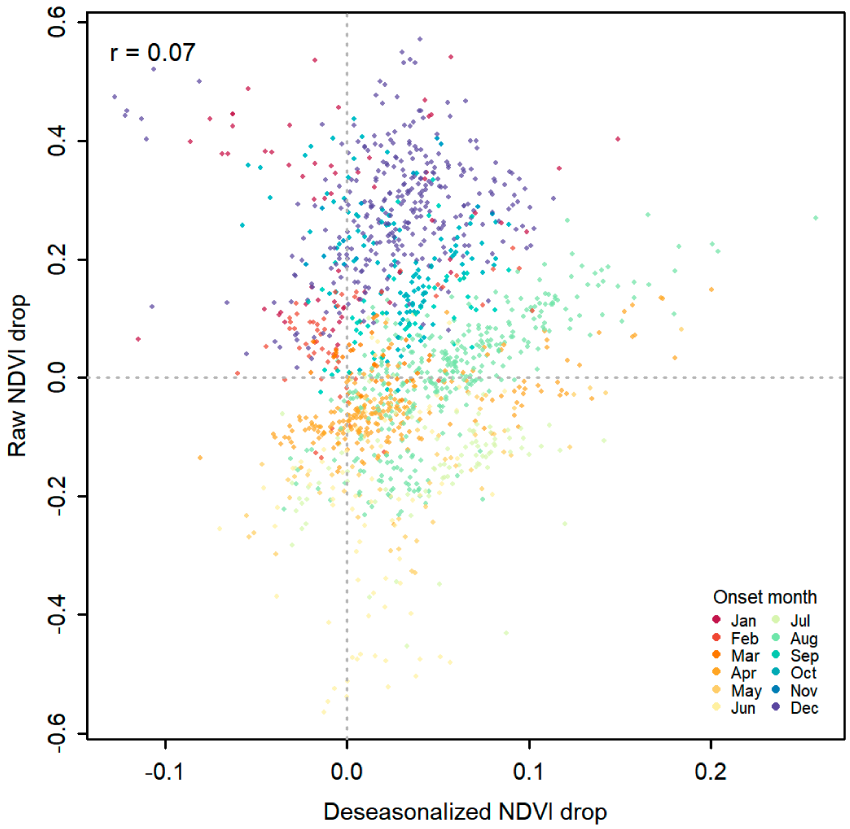

Figure 5. Raw damage versus deseasonalized damage for the 1433 events, colored by onset month (r = 0.07). Events cluster by onset month along divergent trajectories rather than along a single relationship, confirming that the raw metric is dominated by phenological phase rather than damage.

The artifact is not specific to drought monitoring. Any inference that reads ecological change from absolute vegetation-index levels—productivity and carbon sequestration estimation [69], or post-disturbance recovery assessment—inherits the same bias, and recovery trajectories in particular are shaped by nutrient availability and other processes that greenness alone does not capture [3,69]. We hope the diagnostics developed here (coupling tests, forward validation, censoring-aware recovery analysis, cropland masking) will be useful beyond the drought case.

4.2. What Drives Agricultural Vegetation Damage?

With the artifact removed, peak intensity is the most robust and interpretable predictor of deseasonalized damage (wild cluster bootstrap $p = 0.020$; also robust in the cropland-masked specification, Table 7): deeper single-point deficits are more likely to breach crop tolerance thresholds [13,14]. Pre-drought greenness is stably positive in sign across pooled, fixed-effects, coupling-robust, and masked specifications, but the coupling diagnostics (Section 3.3) show that a mechanical component cannot be excluded; we therefore present the "biomass-at-risk" interpretation [14,15] as suggestive rather than established. The drought-duration coefficient is positive in sign across all specifications and thresholds, clearly significant in the growing-season subsample ($p = 0.005$) and under the larger event sample of the −1.3 threshold ($p = 0.016$), but marginal in the full-sample few-cluster inference (wild bootstrap $p = 0.148$)—its biological mechanism (root-zone reserve depletion) is well defined, while its statistical testability is limited by the event sample size and the number of clusters. Duration and peak intensity are complementary in the full sample (bivariate correlation only 0.31; variance inflation factors below 1.6), so joint use outperforms any single indicator. Spring-onset events cause less damage (sine term $p < 0.001$), possibly because of better spring soil-moisture backgrounds and lower water consumption of crops still at the seedling stage [16,17]. We also note a measurement caveat in the direction of underestimation: the Savitzky–Golay filtering and monthly maximum value compositing applied in the NDVI product smooth the 16-day series before we see it, attenuating short dips; this biases damage estimates downward precisely for short events and is a plausible contributor to the weak full-sample duration effect, though it cannot manufacture the growing-season duration signal we do detect.

4.3. Nighttime Light: No Robust Moderating Role

The nighttime light baseline is not significant in the full sample ($p = 0.61$), not significant in the fixed-effects model, not significant under cropland masking, and not significant in the survival analysis of recovery; its apparent significance in the raw growing-season specification of the original manuscript (Table 2 of the previous version) did not survive the transition to the deseasonalized metric (Model C: 0.004, $p = 0.39$), and adding the county built-up fraction as a control flips the small coefficient from +0.003 to −0.004 (both insignificant; Section 3.3). The interpretation we offered previously—that light brightness proxies agricultural adaptive capacity—is therefore not supported (Figure 6): if anything, NTL tracks urban economic activity and agricultural intensity (exposure), and it is mechanically confounded with built-up land, which lowers county-mean NDVI. We now read the NTL results as hypothesis-generating at best; direct data such as irrigation coverage or insurance penetration would be needed for a cleaner test [21,22].

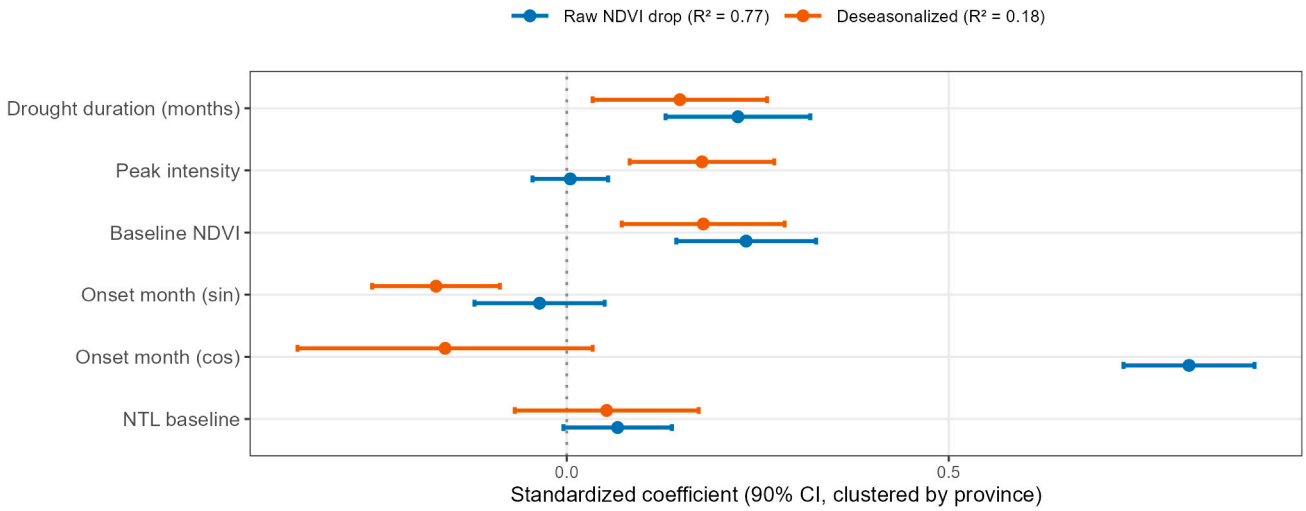

Figure 6. Standardized coefficients with 90% confidence intervals (clustered by province) for the raw ($R^2 = 0.77$) and deseasonalized ($R^2 = 0.18$) damage models. In the raw model, the onset-month cosine term dominates—the model predicts phenological phase rather than damage; after deseasonalization, peak intensity becomes the leading robust predictor of damage, and pre-drought baseline greenness the leading construction-sensitive (suggestive) predictor.

4.4. The Predictability Boundary and Its Monitoring Implications

Under county-grouped cross-validation, the out-of-sample R^2 for deseasonalized damage ranges from 0.18 (linear, retrospective) to 0.44 (random forest), and the real-time-feasible forward variant is, if anything, slightly more predictable (0.22) (Figure 7): standard remote sensing severity variables identify the direction and relative magnitude of damage, and nonlinear specifications (interactions, threshold effects) more than double predictability, yet even flexible models leave about 56% of the damage variance unpredictable (the random forest is stable across random seeds: mean 0.435 ± 0.006 over 10 seeds, range 0.426–0.445; Supplementary file, rf_seed_stability.csv)—prediction errors for individual events remain large. The random forest variable importance ranking (impurity-based; peak intensity > pre-drought baseline greenness > nighttime light > onset month > duration) agrees with the linear-model conclusions. Much of the unpredictable variance presumably comes from factors beyond severity variables—crop phenological stage, soil moisture, management interventions—precisely the gaps that the soil moisture products, crop distribution maps, and insurance records now accumulating in the remote sensing community can fill [37,38]. The crop-category yield validation (Section 3.6) provides the empirical boundary condition: county-level mixed greenness is blind to the yields of a secondary category (oilseeds), while cropland-masked deseasonalized anomalies do recover the dominant-category (grain) signal; crop classification maps (mature products such as ESA WorldCover 10 m and CLCD 30 m) are layers that can be introduced immediately, without waiting for new satellite missions; monitoring systems targeting yield impacts should not apply relative stress indicators off the shelf but explicitly incorporate production scale and crop composition information.

4.5. Limitations

First, statistical inference. The event threshold affects sample composition and power: sensitivity analyses (Supplementary Table S1) show that under the SPEI-3 ≤ −1.3 specification, duration ($p = 0.016$), peak intensity ($p = 0.037$), and baseline NDVI ($p = 0.019$) are all significant; the −1.8 specification (N = 173) is underpowered, but all coefficient signs are consistent. Identification of the damage signal relies substantially on inter-provincial and interannual cross-sectional variation: Conley spatial HAC (100, 200, and 500 km cutoffs)

and province × start-year two-way clustering leave the core coefficients and significance consistent with the baseline, but under the most stringent province × start-year interaction fixed effects, precision is limited (peak intensity $p = 0.20$). Second, measurement and scale. NDVI saturates at high biomass; the product's Savitzky–Golay filtering and monthly MVC attenuate short dips (Section 4.2); county-level aggregation masks within-field heterogeneity; and monthly resolution may miss rapid damage at the 10-day scale. EVI/NIRv are natural alternatives [7,70]. Third, validation boundary. Grain validation covers harvest years 2002–2016 only (county-level grain production is discontinued thereafter), and the CLCD cropland stock used for area normalization includes non-grain crops; oilseed validation, although covering all years 2000–2023, shows zero or negative correlations whose interpretation is limited by planting scale and phenology mismatch (Section 3.6). Multiple layers—crop composition, planting-area decisions, and statistical conventions—lie between county-level greenness and crop production [71]. The significantly negative correlation of the raw metric with grain losses is reported descriptively; candidate explanations are discussed in Section 3.6, but a full decomposition of the sign lies beyond the crop-composition data available to this study. Fourth, residual dynamics. The composite trajectories (Figure 4) show post-event oscillations of comparable amplitude to the in-event dip, indicating residual seasonality and possibly lagged or legacy responses that a within-window minimum does not capture; the survival analysis (Table 5) mitigates but does not eliminate this concern. Fifth, proxy indirectness. Nighttime light is an indirect and confounded proxy for socioeconomic conditions (Section 4.3); direct data such as irrigation coverage and insurance penetration would provide cleaner tests.

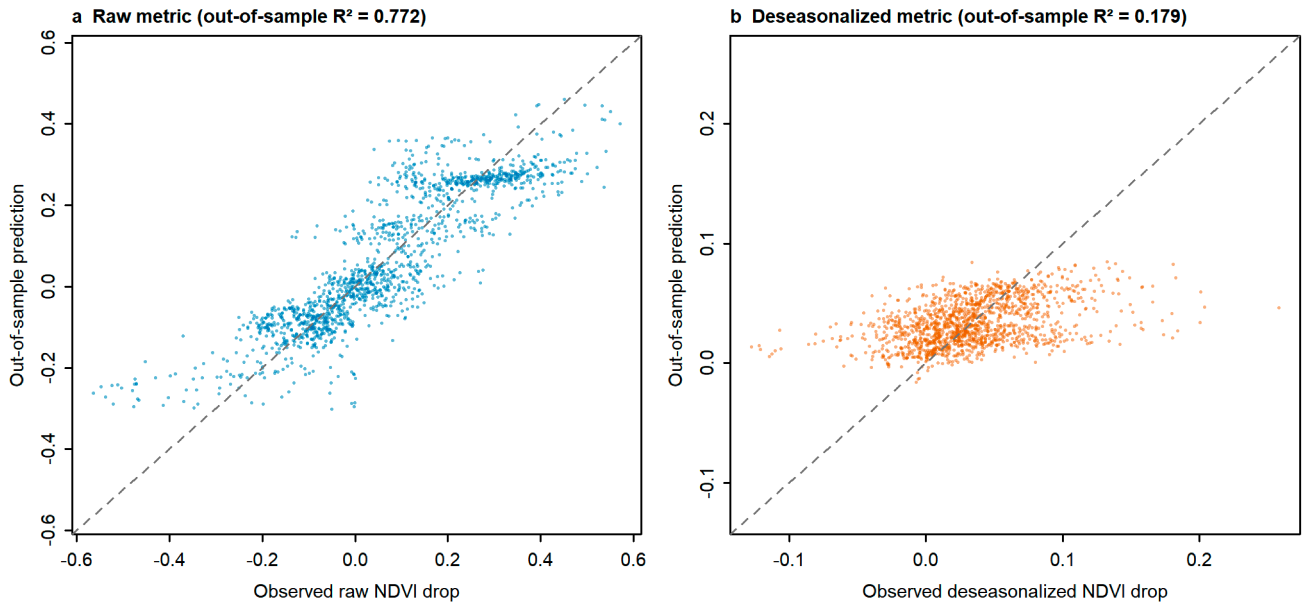

Figure 7. Out-of-fold predictions from county-grouped 5-fold cross-validation versus observed damage for the raw ((**a**), $R^2 = 0.77$) and deseasonalized ((**b**), linear $R^2 = 0.18$; random forest $R^2 = 0.44$) metrics, with the 1:1 line. The raw metric appears highly predictable, but its predictions collapse onto a narrow band driven by onset month; deseasonalized damage remains only partially predictable out of sample.

5. Conclusions

Using 1433 strictly identified severe drought events in China's 13 major grain-producing provinces from 2000 to 2023, this study shows that conventional NDVI damage metrics are dominated by the phenological cycle—their correlation with the detrended, deseasonalized metric is only 0.07 —and that this artifact propagates into recovery metrics, cross-validation

scores, and, as it turned out, into our own yield validation before correction. After detrending, deseasonalization, and cropland masking, agricultural vegetation damage is robustly predicted by peak intensity; pre-drought baseline greenness is a significant but construction-sensitive predictor (suggestive evidence), and duration is significant in the growing season and in larger-sample specifications; censoring-aware survival models (Weibull AFT) show that severity variables accelerate half-recovery attainment through the event-anchored recovery target rather than indicating slower ecosystem recovery, while nighttime light has no robust effect. County-grouped out-of-sample predictability remains limited ($R^2 = 0.18$ linear, 0.44 random forest), and the corrected validation now supports the deseasonalized metric where it matters: on cropland pixels, it is positively and significantly associated with detrended grain-production losses for growing-season events (Spearman r = 0.33, robust to province-clustered inference), whereas the raw metric carries no positive yield signal. For oilseeds—a secondary category in most counties—county-level greenness remains blind. For operational agricultural drought monitoring, our recommendations are: damage assessment must explicitly deseasonalize and detrend; validation against yield must use crop calendar-matched harvest years and detrended loss definitions that retain the observed harvest; early-warning indicators should combine peak intensity with duration; monitoring models should be calibrated by region; crop-specific monitoring in mixed cropping areas must overlay crop distribution maps; and soil moisture and crop physiological data are the key next step to break through the $R^2 \approx 0.44$ predictability boundary.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/rs18234118/s1, Figure S1, threshold-sensitivity analysis; Table S1, regression results under the SPEI-3 ≤ -1.3 and ≤ -1.8 specifications; Table S2, non-detrended climatology-anomaly sensitivity comparison; Table S3, drought-free window (CLEAR_M) sensitivity; Table S4, raw-metric growing-season specification; Table S5, Benjamini–Hochberg FDR correction details; Table S6, region-specific growing-season sensitivity; Table S7, metric-construction sensitivity (seasonally matched baseline and quadratic detrending); R_scripts.zip, the complete R code for event identification, metric construction, all analyses, and figure generation.

Author Contributions: Conceptualization, Y.W. and Z.Z.; methodology, Y.W.; software, Y.W.; validation, Y.W.; formal analysis, Y.W.; investigation, Y.W.; resources, Z.Z.; data curation, Y.W.; writing—original draft preparation, Y.W.; writing—review and editing, Y.W. and Z.Z.; visualization, Y.W.; supervision, Z.Z.; project administration, Z.Z.; funding acquisition, Z.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Social Science Fund of China (General Project), grant number 25BGL193: "Mechanisms and Paths for Enhancing Production Resilience in Major Grain-Producing Areas under Extreme Climate Shocks".

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The precipitation, PET, and NDVI datasets are publicly available from the National Tibetan Plateau Data Center [24,26,28]; the CLCD land cover data are available from the authors of reference [29]; the nighttime light data are available from the source in reference [30]; county-level production and socioeconomic data are from the China Statistical Yearbook (County Level) [31–54]. The event-level analysis dataset (county-mean and cropland-masked event metrics) and the complete R code used for event identification, metric construction, and statistical analysis are provided as supplementary files; the county-month cropland-masked NDVI panel can be regenerated with the provided Google Earth Engine script.

Acknowledgments: During the preparation of this manuscript, the authors used Kimi (Moonshot AI) for language translation and polishing. The authors reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

Abbreviation	Definition
NDVI	Normalized Difference Vegetation Index
SPEI	Standardized Precipitation Evapotranspiration Index
PET	Potential Evapotranspiration
NTL	Nighttime Light
MVC	Maximum Value Compositing
IQR	Interquartile Range
FE	Fixed Effects
HAC	Heteroskedasticity- and Autocorrelation-Consistent
FDR	False Discovery Rate
MODIS	Moderate Resolution Imaging Spectroradiometer
DMSP-OLS	Defense Meteorological Satellite Program–Operational Linescan System
VIIRS	Visible Infrared Imaging Radiometer Suite
CRU	Climate Research Unit
ESA	European Space Agency
CLCD	China Land Cover Dataset
GEE	Google Earth Engine
AFT	Accelerated Failure Time
HR	Hazard Ratio
TR	Time Ratio

References

1. Ding, Y.; Xu, J.; Wang, X.; Cai, H.; Zhou, Z.; Sun, Y.; Shi, H. Propagation of meteorological to hydrological drought for different climate regions in China. *J. Environ. Manag.* **2021**, *283*, 111980. [CrossRef] [PubMed]
2. Schwalm, C.R.; Anderegg, W.R.L.; Michalak, A.M.; Fisher, J.B.; Biondi, F.; Koch, G.; Litvak, M.; Ogle, K.; Shaw, J.D.; Wolf, A.; et al. Global patterns of drought recovery. *Nature* **2017**, *548*, 202–205. [CrossRef] [PubMed]
3. Anderegg, W.R.L.; Schwalm, C.; Biondi, F.; Camarero, J.J.; Koch, G.; Litvak, M.; Ogle, K.; Shaw, J.D.; Shevliakova, E.; Williams, A.P.; et al. Pervasive drought legacies in forest ecosystems and their implications for carbon cycle models. *Science* **2015**, *349*, 528–532. [CrossRef] [PubMed]
4. National Bureau of Statistics of China. *China Statistical Yearbook 2023*; China Statistics Press: Beijing, China, 2023. Available online: http://www.stats.gov.cn/sj/ndsj/2023/indexeh.htm (accessed on 20 November 2024).
5. Wang, A.; Lettenmaier, D.P.; Sheffield, J. Soil moisture drought in China, 1950–2006. *J. Clim.* **2011**, *24*, 3257–3271. [CrossRef]
6. Vicente-Serrano, S.M.; Beguería, S.; López-Moreno, J.I. A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index. *J. Clim.* **2010**, *23*, 1696–1718. [CrossRef]
7. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [CrossRef]
8. Huang, S.; Tang, L.; Hupy, J.P.; Wang, Y.; Shao, G. A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. *J. For. Res.* **2021**, *32*, 1–6. [CrossRef]
9. Isbell, F.; Craven, D.; Connolly, J.; Loreau, M.; Schmid, B.; Beierkuhnlein, C.; Bezemer, T.M.; Bonin, C.; Bruelheide, H.; de Luca, E.; et al. Biodiversity increases the resistance of ecosystem productivity to climate extremes. *Nature* **2015**, *526*, 574–577. [CrossRef] [PubMed]
10. De Keersmaecker, W.; Lhermitte, S.; Tits, L.; Honnay, O.; Somers, B.; Coppin, P. A model quantifying global vegetation resistance and resilience to short-term climate anomalies and their relationship with vegetation cover. *Glob. Ecol. Biogeogr.* **2015**, *24*, 539–548. [CrossRef]
11. Scheffer, M.; Carpenter, S.; Foley, J.A.; Folke, C.; Walker, B. Catastrophic shifts in ecosystems. *Nature* **2001**, *413*, 591–596. [CrossRef] [PubMed]
12. Verbesselt, J.; Hyndman, R.; Newnham, G.; Culvenor, D. Detecting trend and seasonal changes in satellite image time series. *Remote Sens. Environ.* **2010**, *114*, 106–115. [CrossRef]

13. Lesk, C.; Anderson, W.; Rigden, A.; Coast, O.; Jägermeyr, J.; McDermid, S.; Davis, K.F.; Konar, M. Stronger temperature–moisture couplings exacerbate the impact of climate warming on global crop yields. *Nat. Food* **2021**, *2*, 683–691. [CrossRef] [PubMed]
14. Daryanto, S.; Wang, L.; Jacinthe, P.A. Global synthesis of drought effects on maize and wheat production. *PLoS ONE* **2016**, *11*, e0156362. [CrossRef] [PubMed]
15. Fan, Y.; Miguez-Macho, G.; Jobbágy, E.G.; Jackson, R.B.; Otero-Casal, C. Hydrologic regulation of plant rooting depth. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 10572–10577. [CrossRef] [PubMed]
16. Kong, D.; Zhang, Q.; Singh, V.P.; Shi, P. Seasonal vegetation response to climate change in the Northern Hemisphere (1982–2013). *Glob. Planet. Change* **2017**, *148*, 1–8. [CrossRef]
17. Sloat, L.L.; Henderson, A.N.; Lamanna, C.; Enquist, B.J. The effect of the foresummer drought on carbon exchange in subalpine meadows. *Ecosystems* **2015**, *18*, 533–545. [CrossRef]
18. Elvidge, C.D.; Sutton, P.C.; Ghosh, T.; Tuttle, B.T.; Baugh, K.E.; Bhaduri, B.; Bright, E. A global poverty map derived from satellite data. *Comput. Geosci.* **2009**, *35*, 1652–1660. [CrossRef]
19. Jean, N.; Burke, M.; Xie, M.; Davis, W.M.; Lobell, D.B.; Ermon, S. Combining satellite imagery and machine learning to predict poverty. *Science* **2016**, *353*, 790–794. [CrossRef] [PubMed]
20. Henderson, J.V.; Storeygard, A.; Weil, D.N. Measuring Economic Growth from Outer Space. *Am. Econ. Rev.* **2012**, *102*, 994–1028. [CrossRef] [PubMed]
21. Bryan, E.; Deressa, T.T.; Gbetibouo, G.A.; Ringler, C. Adaptation to climate change in Ethiopia and South Africa: Options and constraints. *Environ. Sci. Policy* **2009**, *12*, 413–426. [CrossRef]
22. Kristjanson, P.; Harvey, B.; Van Epp, M.; Thornton, P.K. Social learning and sustainable development. *Nat. Clim. Chang.* **2014**, *4*, 5–7. [CrossRef]
23. Zhang, Z.; Huo, L.; Su, Y.; Shen, H.; Yang, G. Estimation of Corn Net Primary Productivity and Carbon Sequestration Based on the CASA Model: A Case Study of the Fen River Basin. *Sustainability* **2024**, *16*, 2942. [CrossRef]
24. Peng, S. *1-km Monthly Precipitation Dataset for China (1901–2023)*; National Tibetan Plateau Data Center: Beijing, China, 2019. [CrossRef]
25. Peng, S.; Ding, Y.; Liu, W.; Li, Z. 1 km monthly temperature and precipitation dataset for China from 1901 to 2017. *Earth Syst. Sci. Data* **2019**, *11*, 1931–1946. [CrossRef]
26. Peng, S. *1-km Monthly Potential Evapotranspiration Dataset for China (1901–2024)*; National Tibetan Plateau Data Center: Beijing, China, 2025. [CrossRef]
27. Peng, S.; Ding, Y.; Wen, Z.; Chen, Y.; Cao, Y.; Ren, J. Spatiotemporal change and trend analysis of potential evapotranspiration over the Loess Plateau of China during 2011–2100. *Agric. For. Meteorol.* **2017**, *233*, 183–194. [CrossRef]
28. Gao, J.; Shi, Y.; Zhang, H.; Chen, X.; Zhang, W.; Shen, W.; Xiao, T.; Zhang, Y. *China Regional 250 m Normalized Difference Vegetation Index Data Set (2000–2022)*; National Tibetan Plateau Data Center: Beijing, China, 2023. [CrossRef]
29. Yang, J.; Huang, X. The 30 m annual land cover dataset and its dynamics in China from 1990 to 2019. *Earth Syst. Sci. Data* **2021**, *13*, 3907–3925. [CrossRef]
30. Tian, Y.; Cheng, K.M.; Zhang, Z.; Zhang, T.; Feng, J.; Ren, Z.; Li, S.; Yan, D.; Xu, B. An Extended VIIRS-like Artificial Nighttime Light Data Reconstruction (1986–2024): A Novel Modeling Framework and Data Product. *arXiv* **2025**, arXiv:2508.00590.
31. National Bureau of Statistics of China. *China Statistical Yearbook (County-Level) 2001*; China Statistics Press: Beijing, China, 2001.
32. National Bureau of Statistics of China. *China Statistical Yearbook (County-Level) 2002*; China Statistics Press: Beijing, China, 2002.
33. National Bureau of Statistics of China. *China Statistical Yearbook (County-Level) 2003*; China Statistics Press: Beijing, China, 2003.
34. National Bureau of Statistics of China. *China Statistical Yearbook (County-Level) 2004*; China Statistics Press: Beijing, China, 2004.
35. National Bureau of Statistics of China. *China Statistical Yearbook (County-Level) 2005*; China Statistics Press: Beijing, China, 2005.
36. National Bureau of Statistics of China. *China Statistical Yearbook (County-Level) 2006*; China Statistics Press: Beijing, China, 2006.
37. National Bureau of Statistics of China. *China Statistical Yearbook (County-Level) 2007*; China Statistics Press: Beijing, China, 2007.
38. National Bureau of Statistics of China. *China Statistical Yearbook (County-Level) 2008*; China Statistics Press: Beijing, China, 2008.
39. National Bureau of Statistics of China. *China Statistical Yearbook (County-Level) 2009*; China Statistics Press: Beijing, China, 2009.
40. National Bureau of Statistics of China. *China Statistical Yearbook (County-Level) 2010*; China Statistics Press: Beijing, China, 2010.
41. National Bureau of Statistics of China. *China Statistical Yearbook (County-Level) 2011*; China Statistics Press: Beijing, China, 2011.
42. National Bureau of Statistics of China. *China Statistical Yearbook (County-Level) 2012*; China Statistics Press: Beijing, China, 2012.
43. National Bureau of Statistics of China. *China Statistical Yearbook (County-Level) 2013*; China Statistics Press: Beijing, China, 2013.
44. National Bureau of Statistics of China. *China Statistical Yearbook (County-Level) 2014*; China Statistics Press: Beijing, China, 2014.
45. National Bureau of Statistics of China. *China Statistical Yearbook (County-Level) 2015*; China Statistics Press: Beijing, China, 2015.
46. National Bureau of Statistics of China. *China Statistical Yearbook (County-Level) 2016*; China Statistics Press: Beijing, China, 2016.
47. National Bureau of Statistics of China. *China Statistical Yearbook (County-Level) 2017*; China Statistics Press: Beijing, China, 2017.
48. National Bureau of Statistics of China. *China Statistical Yearbook (County-Level) 2018*; China Statistics Press: Beijing, China, 2018.
49. National Bureau of Statistics of China. *China Statistical Yearbook (County-Level) 2019*; China Statistics Press: Beijing, China, 2019.

50. National Bureau of Statistics of China. *China Statistical Yearbook (County-Level) 2020*; China Statistics Press: Beijing, China, 2020.
51. National Bureau of Statistics of China. *China Statistical Yearbook (County-Level) 2021*; China Statistics Press: Beijing, China, 2021.
52. National Bureau of Statistics of China. *China Statistical Yearbook (County-Level) 2022*; China Statistics Press: Beijing, China, 2022.
53. National Bureau of Statistics of China. *China Statistical Yearbook (County-Level) 2023*; China Statistics Press: Beijing, China, 2023.
54. National Bureau of Statistics of China. *China Statistical Yearbook (County-Level) 2024*; China Statistics Press: Beijing, China, 2024.
55. Ministry of Natural Resources of the People's Republic of China. *Public Edition of the 1:1,000,000 Fundamental Geographic Information Data (2021)*; National Catalogue Service System for Geographic Information: Beijing, China. Available online: https://www.webmap.cn/commres.do?method=result100W (accessed on 15 October 2026).
56. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2025. Available online: https://www.R-project.org/ (accessed on 15 October 2026).
57. Barrett, T.; Dowle, M.; Srinivasan, A.; Gorecki, J.; Chirico, M.; Hocking, T.; Schwendinger, B.; Krylov, I. data.table: Extension of 'data.frame'. R Package Version 1.18.4, 2026. Available online: https://CRAN.R-project.org/package=data.table (accessed on 15 October 2026).
58. Bergé, L.; Butts, K.; McDermott, G. fixest: A fast and feature-rich framework for econometric estimations in R. *arXiv* **2026**, arXiv:2601.21749. [CrossRef]
59. Wickham, H.; Miller, E.; Smith, D. haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files. R Package Version 2.5.5, 2025. Available online: https://CRAN.R-project.org/package=haven (accessed on 15 October 2026).
60. Therneau, T.M. Survival: Survival Analysis. R Package Version 3.8.3, 2025. Available online: https://CRAN.R-project.org/package=survival (accessed on 15 October 2026).
61. Wright, M.N.; Ziegler, A. ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *J. Stat. Softw.* **2017**, *77*, 1–17. [CrossRef]
62. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [CrossRef]
63. Hosking, J.R.M. L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics. *J. R. Stat. Soc. Ser. B* **1990**, *52*, 105–124. [CrossRef]
64. Cameron, A.C.; Gelbach, J.B.; Miller, D.L. Bootstrap-based improvements for inference with clustered errors. *Rev. Econ. Stat.* **2008**, *90*, 414–427. [CrossRef]
65. Klein, J.P.; Moeschberger, M.L. *Survival Analysis: Techniques for Censored and Truncated Data*, 2nd ed.; Springer: New York, NY, USA, 2003. [CrossRef]
66. Cox, D.R. Regression Models and Life-Tables. *J. R. Stat. Soc. Ser. B* **1972**, *34*, 187–220. [CrossRef]
67. West, H.; Quinn, N.; Horswell, M. Remote sensing for drought monitoring & impact assessment: Progress, past challenges and future opportunities. *Remote Sens. Environ.* **2019**, *232*, 111291. [CrossRef]
68. Kerr, Y.H.; Waldteufel, P.; Wigneron, J.P.; Delwart, S.; Cabot, F.; Boutin, J.; Escorihuela, M.J.; Font, J.; Reul, N.; Gruhier, C.; et al. The SMOS Mission: New Tool for Monitoring Key Elements of the Global Water Cycle. *Proc. IEEE* **2010**, *98*, 666–687. [CrossRef]
69. Song, Z.; Tian, H.; Li, Z.; Luo, Y.; Liu, Y. Changes in plant nutrient utilization during ecosystem recovery after wildfire. *J. Environ. Manag.* **2021**, *295*, 112994. [CrossRef] [PubMed]
70. Huete, A.R.; Liu, H.Q.; Batchily, K.; van Leeuwen, W. A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sens. Environ.* **1997**, *59*, 440–451. [CrossRef]
71. Lobell, D.B.; Azzari, G.; Burke, M.; Gourlay, S.; Jin, Z.; Kilic, T.; Murray, S. Eyes in the Sky, Boots on the Ground: Assessing Satellite- and Ground-Based Approaches to Crop Yield Measurement and Analysis in Uganda. *Am. J. Agric. Econ.* **2020**, *102*, 466–486. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.