

Journal Name

Crossmark

PAPER

RECEIVED
dd Month yyyy
REVISED
dd Month yyyy

Divergent estimates of agricultural climate sensitivity from survey and satellite yield data

Caro Park^{1*}, Jonathan Proctor² and Peter Huybers³¹Center for Climate Resilience, University of Pennsylvania²Food and Resource Economics, University of British Columbia³Department of Earth and Planetary Sciences, Harvard University

*Author to whom any correspondence should be addressed.

E-mail: parkcaro@upenn.edu

Keywords: agriculture, crop yields, survey, satellite, climate sensitivity, food security**Abstract**

Accurate monitoring of crop yields is critical for understanding historical food production trends, identifying yield gaps, assessing food security risks, and evaluating climate impacts. Two primary means of estimating yields are through surveys and satellite remote sensing. We compare satellite-derived greenness indices with national-level survey data from the Food and Agriculture Organization of the United Nations (FAO) and find systematic discrepancies. Correspondence between satellite and survey yield estimates is substantially weaker in countries where there are reporting-year offsets, and correspondence is modestly stronger in countries with higher GDP per capita and cropping intensity. Across all countries and 19 crops, weather variables explain 25% of satellite-derived yield variability but only 3% of survey-derived yield variability. Accordingly, satellite-derived yields are found to be roughly three times more sensitive to changes in soil moisture than survey-based yields, though the interquartile ranges of these estimates overlap. Across space, countries with lower GDP per capita tend to have a larger discrepancy in weather-driven yield variability between satellite and survey data. Collectively, the patterns we document are more consistent with systematic limitations in survey-reported yields, though satellite observations also have non-negligible errors. These findings imply that inferences about climate impacts on agriculture – and therefore assessments of adaptation needs and food security risks – depend on the choice of yield data source. This dependence is particularly acute in low-latitude, low-GDP per capita regions where agricultural climate vulnerability is thought to be greatest due to high reliance on agriculture for livelihood and subsistence.

1 Introduction

Empirical inferences of how climate change will impact agricultural yields are generally made by inferring the sensitivity of yields to weather variability, and then applying these sensitivities in the context of projected future climate conditions [1, 2, 3]. Differences among results are often attributable to model specifications and emissions scenarios, as well as adaptation measures expected to alter the exposure and sensitivity to weather, over which there continues to be debate [4, 5, 6, 7, 8]. A more basic issue that has received comparatively less attention, however, is the degree to which weather controls interannual yield variability [9, 10]. To the extent that weather drives year-to-year yield fluctuations, climate change implies larger production risk, and adaptation that reduces weather sensitivity becomes more consequential. Furthermore, increases in the variance of weather-driven yield anomalies are projected to grow with warming [11], amplifying this concern.

Yield data reported by the Food and Agriculture Organization of the United Nations (FAO) have long been used to estimate the influence of climatic factors on agricultural productivity [2, 12, 13, 14]. These data are compiled by national statistical agencies in collaboration with the FAO, often relying on self-reported surveys, expert assessments, and limited field sampling [15, 16, 17]. There is a substantial literature discussing limitations of FAO-reported yield data [18], including under-reporting, inconsistent methods across countries, and political incentives to misreport [19, 20]. These issues may be particularly severe in poor countries, where

data infrastructure is weak, agricultural activity is often informal or subsistence-based, capacity for regular field data collection is limited [21, 22], and the impacts of climate change may be the largest [23]. Recent efforts to compile subnational crop statistics have revealed the extent of data scarcity, particularly in Sub-Saharan Africa, where chronic malnutrition affects nearly 282 million people and yet agricultural production statistics are among the most sparse in the world [24]. The implications of these errors for estimates of climatic influences on yields, however, remain poorly quantified. Despite these limitations, FAO-reported yields remain a valuable and commonly used data product due to their global coverage, their ease of use, and the lack of other survey-based options.

Satellite-based measures of vegetation activity are another imperfect source of information about crop yields. The Normalized Difference Vegetation Index (NDVI) and solar-induced fluorescence (SIF), the latter of which is the basis of the satellite-based product we use here (Methods), have both been shown to be predictive of yields across a range of crops and settings [25, 26, 27, 28, 29]. SIF has been shown to have higher accuracy than NDVI for crop yield prediction in some contexts because it is a more direct proxy for photosynthetic activity rather than vegetation greenness [30, 31]. A benefit of satellite measures is that they are measured relatively uniformly across global croplands, which facilitates comparisons of agricultural weather sensitivity across regions. A challenge, however, is that the relationship between satellite observations and true yields could differ across space and income levels due to differing cropping systems, canopy structures, and management practices [32, 33, 34]. Weather-attributable variations in satellite-derived yields are susceptible to biases associated with errors in crop masks, heterogeneous management practices, and cloud contamination [35]. Furthermore, NDVI and SIF indicators are known to respond heterogeneously to weather events like drought depending on biome as well as resolution, crop type, and local management conditions [36, 37]. SIF responds primarily to canopy-level photosynthetic activity, and therefore cannot fully detect yield losses that arise from stresses – such as heat damage at pollination or pathogen pressure – that reduce grain production without substantially altering canopy structure or photosynthesis. In our analysis, we use Contiguous SIF (CSIF), which is derived by training a neural network on sparse OCO-2 SIF retrievals and filling spatial and temporal gaps using MODIS surface reflectance. Over croplands, this procedure produces values that tend to be lower than direct OCO-2 SIF measurements [38]. Despite these limitations, CSIF provides information about crop productivity at global scale that is collected independently of survey-based estimates.

A growing literature documents how non-classical measurement error in both reported agricultural data [39, 40] and satellite measures [41] can have consequences for statistical inference. Satellite-derived greenness indices, however, tend to correlate more strongly with objective crop-cut estimates than with farmer-reported yields [42, 43, 44]. And these discrepancies are concentrated in low-income settings where data infrastructure is weakest [40]. At the national level, FAO yields aggregate across heterogeneous reporting systems and methods, and the implications of these errors for inferences about climate sensitivity remain poorly understood. A potential concern is that survey yields in low-income countries may contain substantial structured measurement error, such as reduced reported yield variance due to interpolation of missing values, which could lead to underestimates of climate-driven yield losses.

Here, we compare survey-reported yields to satellite-derived yield proxies across crops and countries. We find relatively strong correspondence between survey-derived and satellite-derived yields in high-income countries, but this relationship deteriorates in many lower-income settings. Correspondingly, we find that the magnitude and spatial pattern of weather-attributable yield variability differs substantially between these data sources, and that these differences can meaningfully affect inferences about agricultural climate sensitivity and risk. Inferences about yield responses to soil moisture, temperature, and other climatic factors underpin projections of severe yield losses under warming [1, 10, 45] and are only as reliable as the yield data on which they are estimated.

2 Data and Methods

Survey-derived yield measurements come from FAOSTAT [15]. Data are taken from 2000–2023 for all countries. The temporal range is selected to match that of the satellite data. The following 19 crops are studied: barley, cassava, groundnuts with shell, maize, millet, oats, potatoes, pulses, rapeseed, rice (paddy), rye, seed cotton, sorghum, soybeans, sugar beet, sunflower seed, sweet potatoes, wheat, and yams. Included in this dataset are FAOSTAT-determined flags that indicate whether a measurement is an official or estimated value. We use these flags as a proxy for FAOSTAT’s assessment of data reliability.

For satellite-based information regarding yield we use clear-sky Contiguous Solar-Induced Fluorescence (CSIF), as provided by Zhang et al. [38]. CSIF is derived by training a neural network on sparse OCO-2 SIF retrievals with the first four MODIS VNIR bands (blue, green, red, NIR) from the Terra and Aqua satellites as inputs, producing spatially contiguous estimates at 0.05 degree resolution and 4-day intervals. Because CSIF captures photosynthetic activity via canopy greenness and structure rather than harvested biomass, it serves as a proxy for, rather than a direct measure of, crop yield. Over croplands, the gap-filling procedure produces values that are on average 12% lower than direct OCO-2 SIF measurements [38], and the relationship between SIF and yield may be nonlinear at high canopy densities where SIF can saturate or decouple from biomass accumulation. Because both limitations apply broadly across croplands rather than selectively to particular regions or income levels, they are unlikely to explain the spatially structured survey–satellite divergence we document.

To develop the satellite-derived yield estimates (i.e., the CSIF crop yield proxy), we first obtained the globally gridded fraction of harvested area in the year 2000 for each crop from Ramankutty et al. (2008) [46]. The static year-2000 crop masks may not reflect subsequent changes in cropping patterns, potentially introducing a time-varying bias in regions where agricultural land use has shifted substantially. For each grid cell where crops are grown, we extracted the planting and harvesting dates from the Center for Sustainability and the Global Environment’s georeferenced crop calendar dataset and determined each crop’s growing season [47]. If more than one growing season exists for a given crop, both growing seasons were included. Then, for each year from 2000–2023, we extracted CSIF data from the growing-season months. Year is assigned by harvest month, not planting month, in accordance with the FAO data; in subsequent analyses we also allow for a ± 1 year reporting mismatch between FAO yields and the CSIF proxy. Finally, we average CSIF values for each country, crop, and year over the growing season, weighting spatially by the fraction of harvested area for each crop from Ramankutty et al. (2008)[46]. We weight to mitigate the potential issue that, with the native 0.05 degree resolution of CSIF, individual grid cells may contain multiple crop types; residual mixing may introduce noise, particularly in regions with intercropped or fragmented landscapes [35, 48]. Our full procedure outlined here results in a crop-specific, country-level annual CSIF proxy intended to capture interannual variation in photosynthetic activity over harvested areas.

We obtained globally gridded surface soil moisture data at 0.25 degree daily resolution from the European Space Agency Climate Change Initiative [49, 50]. We obtained globally gridded maximum daily temperature data at 0.5 degree resolution from the National Oceanic and Atmospheric Administration (NOAA) Physical Sciences Laboratory (<https://psl.noaa.gov>). We aggregated temperature and soil moisture to the country–crop–year level in the same way as CSIF.

2.1 Correspondence between survey (V) and satellite (S) yields

We use a two-stage approach to analyze differences between survey- and satellite-derived yields. In the first stage, we measure survey–satellite correspondence by calculating the Pearson correlation between survey (V) and satellite (S) yields for each country–crop pair,

$$\rho_{VS} = \frac{\text{Cov}(V, S)}{\sigma_V \sigma_S}. \quad (1)$$

In the second stage, we model cross-sectional variation in ρ_{VS}^2 using a linear mixed-effects model with crop-level random intercepts,

$$\rho_{VS}^2 = (\alpha_0 + \alpha_i) + \sum_{n=1}^p \alpha_n X_n + \epsilon_{ij}, \quad (2)$$

where α_0 is the fixed intercept, α_i is the random intercept for crop type i , X_n are the country–crop-level predictors described in Table 1, and j indexes country. We use random intercepts for crop type because of their uneven distribution across countries. As a sensitivity analysis, we also estimate the model using crop type as a fixed effect, and results are nearly identical (see Supplementary Table S1). All continuous covariates are standardized to zero mean and unit variance to facilitate comparison of effect magnitudes. We analyze the squared correlation in the second stage to focus on the strength rather than the direction of association; results are consistent when analyzing ρ_{VS} rather than ρ_{VS}^2 (see Supplementary Table S2).

Table 1: Predictors from covariate matrix X that are included in the mixed-effects model.

Predictor	Description	Source
Average GDP per capita (log scale)	Country-level mean GDP per capita over the study period (2000–2023), used as a proxy for overall economic status. Log-transformed to represent proportional income differences and reduce the influence of its right-skewed distribution.	World Bank [51]
Year offset	Indicates whether the relationship between the CSIF proxy and the FAO-reported yield for each country–crop pair is strongest when there is a year offset, designated as “Lead” or “Lag” or “None.”	Calculated (FAO-STAT) [15]
Flagged percentage	Percentage of FAO-reported yield measurements demarcated with an FAO flag indicating a non-official estimate. Represents FAOSTAT’s assessment of data reliability.	FAOSTAT [15]
Total harvested-area fraction	Country-level sum of grid-cell harvested-area fractions for the specific crop being analyzed (each grid cell contributes a value in $[0, 1]$ representing the fraction of its area harvested for that crop). The sum is dimensionless and serves as a proxy for total harvested area; it is not expressed in hectares because grid-cell areas vary with latitude.	Ramankutty et al. (2008) [46]
Mean cropland fraction	Country-wide temporally averaged fraction of land covered by any cropland over the study period. Represents overall agricultural intensity.	MODIS Land Cover Type product [52]

128 The year offset is computed as a diagnostic of reporting or phenology mismatches by
 129 determining, for each country–crop pair, whether the correlation between the CSIF proxy and
 130 FAO-reported yield is strongest when the FAO series is shifted by one year earlier, unshifted, or
 131 shifted by one year later. This procedure yields a categorical predictor that captures potential
 132 reporting-year inconsistencies; it is not used to define the first-stage correlation outcome, which is
 133 always computed without shifting. Year offsets may arise due to crop-specific harvest calendars and
 134 a country’s reporting protocols; for example, a country might report maize yields to FAO under the
 135 harvest-year label even though the crop’s peak growing season (captured by CSIF) occurs primarily
 136 in the previous calendar year. For the 55% of country–crop pairs assigned a lead or lag category
 137 through this method, approximately 8% exhibit a statistically significant difference in correlation
 138 relative to the unshifted series ($p < 0.1$), assessed using a two-tailed Fisher’s Z-transformation to
 139 compare correlation coefficients (see Supplementary Table S3).

140 2.2 Yield variability attributable to weather

141 To quantify the share of interannual yield variability attributable to weather, we estimate simple
 142 linear regressions relating yields to growing-season weather conditions. For each country–crop pair,
 143 yield is regressed on growing-season averages of soil moisture and daily maximum temperature,

$$\log(Y_{jt}) = \beta_0 + \beta_1 SM_{jt} + \beta_2 TMAX_{jt} + \varepsilon_{jt}, \quad (3)$$

144 where Y_{jt} = yield in country j and year t , SM_{jt} = mean soil moisture and $TMAX_{jt}$ = mean
 145 maximum temperature over the crop-specific growing season.

146 All variables are subset to the longest common interval they share and linearly detrended to
 147 remove long-run changes associated with technological progress and management improvements.
 148 The median length of the resulting time series is 21 years, with a range of 10–24 years. Yield is
 149 log-transformed to stabilize variance and facilitate comparison across crops and countries. Models
 150 are estimated separately for each country–crop pair using both survey-derived and satellite-derived
 151 yields.

152 Because this analysis is conducted at the country–crop level and with a limited number of
 153 years, we adopt a deliberately simple modeling framework focused on the linear relationship
 154 between weather variables and yield. This simplicity is itself informative: the degree to which even
 155 basic weather metrics explain yield variability differs markedly between data sources, and this
 156 contrast does not depend on the choice of a specific functional form. We note that this linear

specification does not fully capture well-documented nonlinear yield responses to temperatures [1, 45] and soil moisture [10]; because yield losses accelerate at extremes, non-weather-induced noise in survey data will attenuate estimated sensitivity more in a nonlinear framework than in a linear one, so our approach likely understates the true divergence between data sources. With country–crop models based on 10–24 years of data and 1,744 separate regressions, statistical power varies across models, and we have not applied corrections for multiple comparisons, which may affect the interpretation of individual model significance.

3 Results

3.1 Survey–satellite yield correspondence varies across crops and countries

Figure 1 illustrates how correspondence between the two yield measures varies across crops in the United States, where survey data are generally considered high quality. This provides a best-case benchmark for agreement between the two data sources. Survey–satellite correspondence in other high-income countries, including European countries with smaller average field sizes, is comparable to that observed in the United States (Supplementary Figure S1). Most grains show strong correlations with the CSIF proxy, consistent with previous findings [31]. Cereal crops such as maize are particularly well suited to satellite-based observation of photosynthesis because they are typically grown in large, dense fields and have physiological and biochemical traits that enhance detectability [30, 53, 54]. Similarly, C4 crops like maize and sorghum tend to have particularly strong SIF signals due to their photosynthetic pathway, which may partly explain the stronger correspondence observed for these crops. In contrast, smaller, low-lying crops such as groundnuts and sweet potatoes are planted at lower densities, which may make satellite-based measurement of yields more difficult.

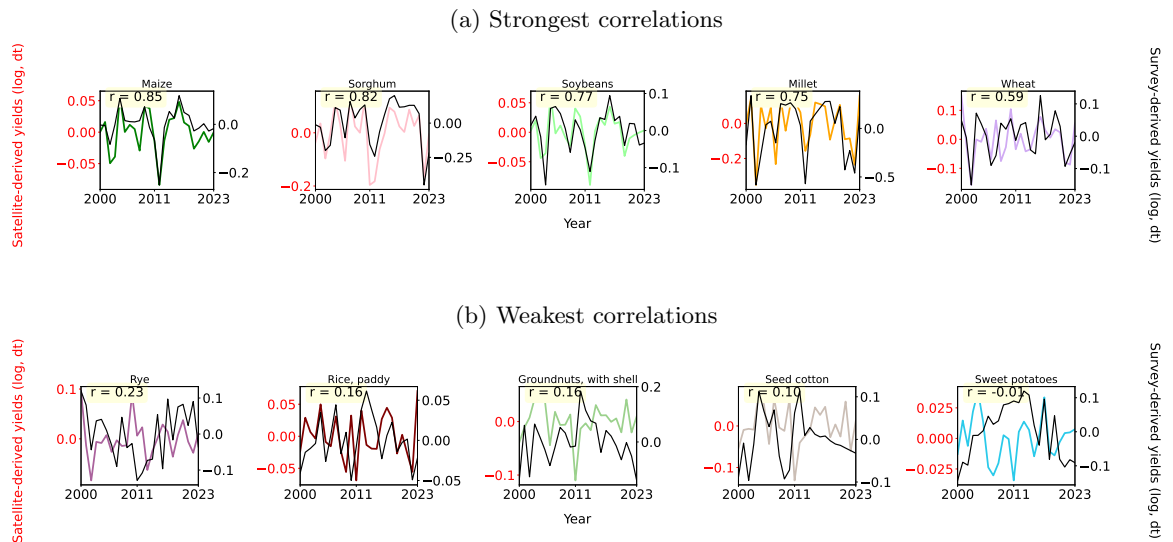


Figure 1: Time series of FAO-reported yields (black lines) and the CSIF proxy (colored lines) for various crops in the USA, ordered by strength of correlation. Pearson correlation coefficient is shown in the top left corner of each panel; when squared, this is equivalent to the squared correlation used in modeling survey–satellite correspondence (ρ_{VS}^2).

Although maize is generally regarded as an ideal crop for remote monitoring, correspondence between survey- and satellite-derived yields remains highly heterogeneous across countries. Figure 2 illustrates this variation by showing countries with the highest and lowest correlations for maize. In many cases, weak correlations coincide with FAO quality flags, though this pattern does not hold universally.

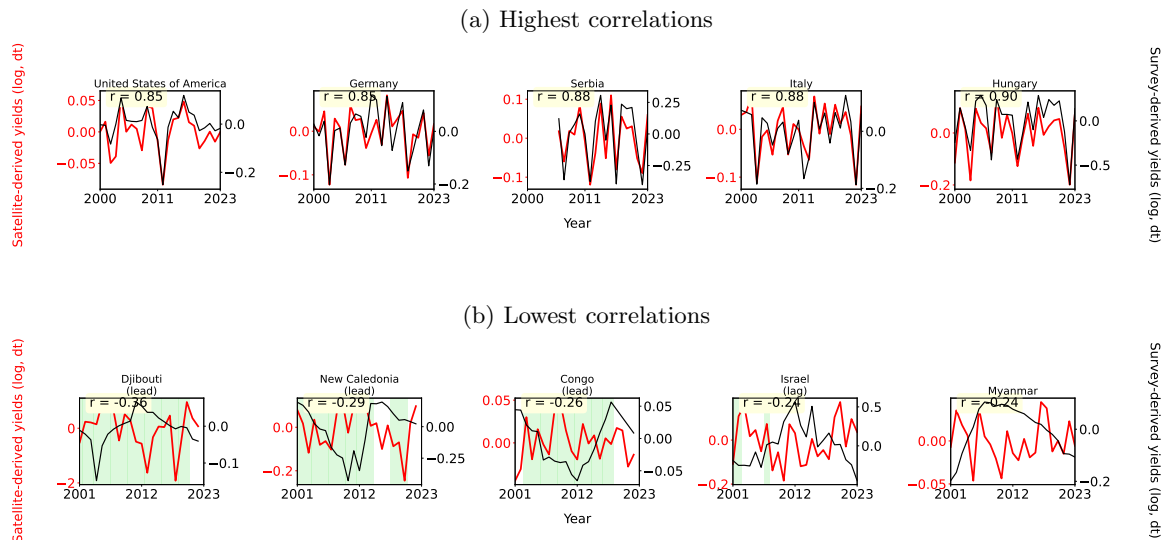


Figure 2: Time series of FAO-reported yields (black lines) and CSIF proxy (red lines) for maize yields of (a) the countries with the highest correlation and (b) the countries with the lowest correlation. Green highlighting indicates whether FAOSTAT flagged that year's yield measurement for a given country. Lead/lag indicates that the correlation between the CSIF proxy and the FAO yield series is strongest when the FAO series is shifted by one year relative to CSIF.

3.2 Correspondence between survey- and satellite-derived yields is associated with socioeconomic factors

To summarize correspondence across the full dataset, we compute the squared correlation between survey- and satellite-derived yields for each country–crop pair, ρ_{VS}^2 , and model cross-sectional variation in this index using a linear mixed-effects model with crop-level random intercepts. Across all country–crop pairs, ρ_{VS}^2 ranges from near zero to above 0.8 (Figure 3a). All continuous predictors are standardized to unit variance, so coefficients represent changes in ρ_{VS}^2 associated with a one-standard deviation increase in each predictor. The intercept of 0.16 is the predicted ρ_{VS}^2 for a country–crop pair with no reporting-year offset and all covariates at their sample means. The marginal mean (i.e., the average) across all observations is 0.11 (Supplementary Figure S2), providing a useful benchmark for interpreting effect magnitudes. The gap between the model intercept and mean reflects the lower correspondence at ± 1 -year offsets, which apply to the majority of observations.

Table 2 shows that survey–satellite correspondence is systematically higher in countries with higher GDP per capita and more intensively cropped agricultural systems. A one-standard deviation increase in both average GDP per capita and mean cropland fraction is associated with an increase of 0.02 in survey–satellite correspondence (95% CI: 0.01–0.03), translating to a roughly 18% increase relative to the mean. Total harvested-area fraction is also positively associated with correspondence (0.01; 95% CI: 0.01–0.02).

In contrast, reporting inconsistencies are associated with substantially weaker survey–satellite correspondence. Country–crop pairs exhibiting a reporting-year offset – whether a lead or a lag – show an average reduction in ρ_{VS}^2 of 0.09 (95% CI: –0.11 to –0.08), which exceeds half of the mean correspondence. The share of FAO-flagged observations is not significantly associated with correspondence after controlling for GDP and reporting-year offsets ($p = 0.70$).

ρ_{VS}^2 : correlation squared of survey- & satellite-derived yields			
Predictors	Estimates	CI	p
Intercept	0.16 ***	0.14 – 0.17	<0.001
Avg GDP/capita (log scale)	0.02 ***	0.01 – 0.03	<0.001
Survey flag pct	-0.00	-0.01 – 0.01	0.704
Offset (year $\pm$ 1)	-0.09 ***	-0.11 – -0.08	<0.001
Total harvested-area fraction	0.01 ***	0.01 – 0.02	<0.001
Mean cropland fraction	0.02 ***	0.01 – 0.03	<0.001
Random Effects			
σ^2	0.019		
τ_0 cropname	0.001		
ICC	0.037		
N cropname	19		
Observations	1744		
Marginal R2 / Conditional R2	0.175 / 0.206		
+ p < 0.1 * p < 0.05 ** p < 0.01 *** p < 0.001			

Table 2: Mixed-effects model estimates for predictors of survey–satellite yield correspondence. The dependent variable is the squared correlation between survey-derived and satellite-derived yields, ρ_{VS}^2 , computed for each country–crop pair. The model includes crop-level random intercepts and is estimated on $N = 1,744$ country–crop observations across 19 crops. The intraclass correlation coefficient (ICC) of 0.037 indicates that crop type accounts for a small share of residual variance, with most variation occurring within crops across country-year observations. The random effect variance (τ_0) represents crop-to-crop variability in baseline correspondence.

208 These patterns are reflected spatially in Figure 3: countries with low correspondence in panel
209 (a) tend to coincide with lower GDP in (b) and higher prevalence of reporting-year offsets in (c),
210 while (d) highlights that flagged reporting is concentrated in specific regions.

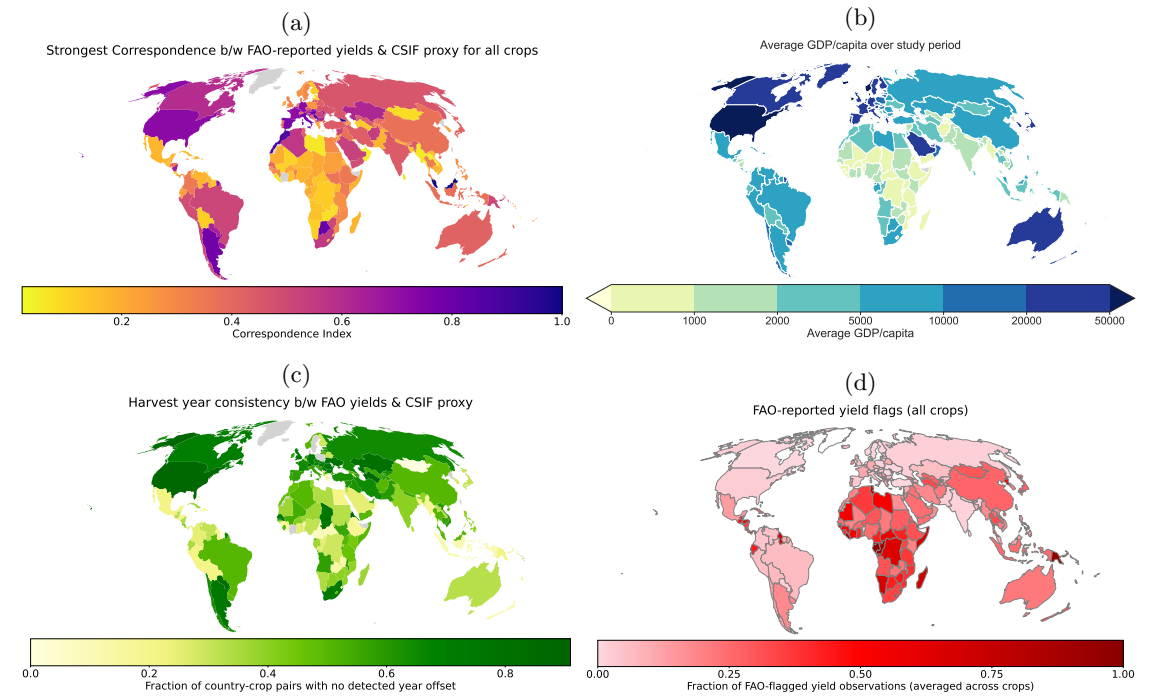


Figure 3: **a)** Global map of the maximum squared correlation between FAO-reported yields and the CSIF proxy across all crops grown in a given country. **b)** Global map of average GDP per capita over the study period (2000–2023). **c)** For each country, the fraction of its crops with no detected reporting-year offset between FAO yields and the CSIF proxy. Higher values indicate more consistent harvest-year reporting. **d)** Global map of the fraction of FAO-reported yield observations flagged as non-official estimates, averaged across crops over the study period.

Because weather sensitivity is inferred from interannual crop yield variability, systematic differences in yield variability between data sources should translate into different estimates of climate–yield relationships.

3.3 Yield variability attributable to weather and yield sensitivity to weather are higher with satellite-derived yields than survey-derived yields

We next examine whether survey–satellite discrepancies translate into different inferences about the role of weather in driving yields. For each country–crop pair, we estimate a simple weather–yield regression using detrended log yields as the dependent variable and detrended growing-season soil moisture (SM) and maximum temperature (TMAX) as predictors. We summarize model performance using adjusted R^2 , which measures the fraction of interannual yield variability explained by weather after accounting for degrees of freedom.

Table 3 reports median R^2 , median adjusted R^2 , the fraction of country–crop models passing an overall F-test ($p < 0.05$), the median estimated soil moisture coefficient with its interquartile range, and the fraction of models for which the SM coefficient is statistically distinguishable from zero. Across all crops, weather explains substantially more interannual variability in satellite-derived yields than in survey-derived yields (adjusted $R^2 = 0.25$ vs. 0.03, an ~ 8 -fold difference). This contrast is also reflected in model significance: 60% of satellite-based regressions pass an overall F-test, compared to only 20% of survey-based regressions.

		R2	Adj R2	F-test pass	SM coefficient	SM p.v. pass
All crops (pooled)	Satellite	0.32	0.25	60%	2.66 (0.89, 5.51)	49%
	Survey	0.12	0.03	20%	0.81 (−1.24, 4.59)	15%
Maize	Satellite	0.38	0.32	68%	2.54 (1.36, 5.27)	52%
	Survey	0.16	0.07	27%	1.65 (−0.83, 7.11)	19%
Sorghum	Satellite	0.36	0.30	68%	2.96 (1.65, 6.42)	61%
	Survey	0.13	0.05	19%	1.66 (−1.57, 5.65)	17%
Wheat	Satellite	0.30	0.23	55%	3.26 (1.25, 7.01)	52%
	Survey	0.14	0.05	27%	0.92 (−1.7, 6.0)	21%
Cassava	Satellite	0.21	0.13	39%	1.04 (0.35, 2.68)	34%
	Survey	0.11	0.01	16%	0.24 (−0.51, 2.11)	11%
Potatoes	Satellite	0.26	0.18	50%	3.13 (1.25, 6.0)	45%
	Survey	0.09	−0.01	10%	0.63 (−0.86, 2.93)	11%

Table 3: Performance of simple weather–yield models estimated separately for each country–crop time series. Each row gives the median across country-level models for that crop; the “All crops” row pools all country–crop models. Five representative crops are shown here; all 19 crops are shown individually in Supplementary Table S4.

The gap between data sources is pronounced for major cereals. For maize, weather explains roughly 32% of interannual yield variability in satellite-based estimates, compared to 7% in survey-based estimates; for wheat, the corresponding figures are 23% and 5%. These crops are widely cultivated and relatively well-suited for remote sensing, which theoretically should result in more consistent measurement of year-to-year crop productivity and, by extension, weather-driven fluctuations.

In contrast, crops such as cassava and potatoes are more difficult to predict using simple weather metrics, especially when relying on survey data. These crops are often produced at smaller spatial scales, frequently by smallholder farmers, and are grown under more heterogeneous management practices, which may weaken the apparent relationship between reported yields and aggregate weather conditions. Even so, satellite-derived yields for these crops still exhibit substantially higher weather-driven yield variability than their survey-based counterparts.

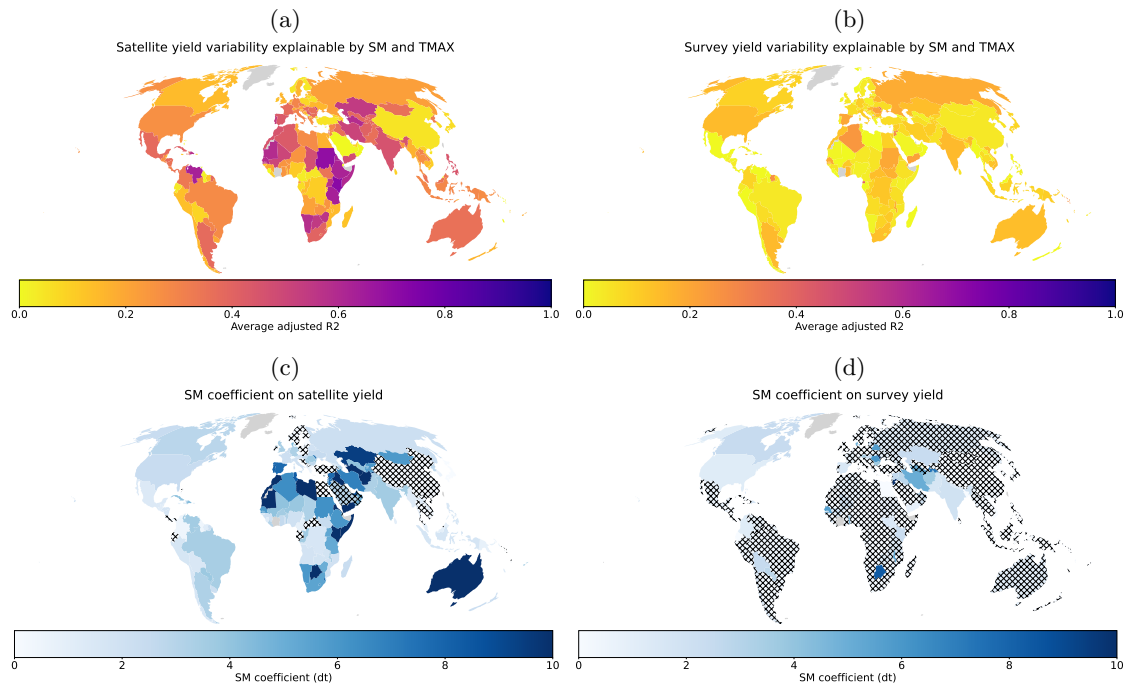


Figure 4: **(a,b)** Interannual yield variability explained by growing-season soil moisture and maximum temperature, averaged across crops within each country, for yields derived from (a) the CSIF proxy and (b) FAO-reported data. **(c,d)** Soil moisture sensitivity, averaged across all crops, derived from the weather models for the CSIF proxy and FAO-reported data. Hatching indicates statistical insignificance (p -value > 0.05), obtained via inverse normal method for combining p -values across all crops [55].

241 The survey–satellite discrepancies are unevenly distributed across space. As shown in Figure 4,
 242 the fraction of yield variability explained by weather is highest in low-latitude regions when
 243 satellite-derived yields are used, with particularly strong model performance across Sub-Saharan
 244 Africa, parts of the Middle East, and South America (the same spatial pattern holds for each of the
 245 five representative crops shown in Supplementary Figure S3, indicating that the cross-crop average
 246 is not driven by any single crop). This may be driven in part by higher rates of water limitation in
 247 these generally low-latitude and semi-arid regions. In contrast, models based on survey yields tend
 248 to perform better in mid- to high-latitude regions, with generally weak explanatory power in the
 249 tropics. The gap in weather-explained yield variability is systematically larger in poorer countries
 250 than richer countries ($p < 0.05$). As visualized in Figure 5, a 10-fold increase in GDP per capita
 251 reduces this gap by approximately 6 percentage points, or roughly a third of the average
 252 country-level gap (0.18). The pattern is similar in both tropical and non-tropical countries, though
 253 somewhat more pronounced in tropical countries (see Supplementary Tables S5–S6).

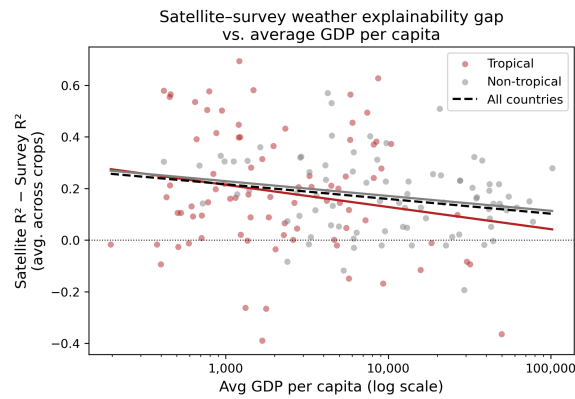


Figure 5: Difference between satellite-derived and survey-based weather explainability (adjusted R^2 values, averaged across crops as in Figure 4) as a function of GDP per capita (USD, log scale). Each data point represents an individual country, categorized into Tropical (red) and Non-tropical (gray) regions.

Adjusted R^2 summarizes overall weather explainability, but it is not a direct measure of how sensitive yields in an area are to weather shocks. We therefore examine the estimated soil moisture coefficient as a more direct diagnostic of yield sensitivity to water limitation. Soil moisture captures growing-season water availability and is a key determinant of yields in many rainfed systems [10]. Averaged across all crops, the estimated soil moisture response is substantially larger for satellite-derived yields (2.66; IQR: 0.89–5.51) than for survey yields (0.81; IQR: –1.24–4.59), and the share of models with detectable soil moisture effects is higher for satellite yields (49%) than for survey yields (15%). The spatial pattern of satellite-based yield sensitivity to soil moisture (Figure 4c) mirrors that of satellite-based yield variability explained by weather (Figure 4a), with higher sensitivity in low-latitude and semi-arid regions that also tend to have populations with lower income. Plot-level studies find that satellite-derived greenness indices correlate more strongly with objective crop-cut measurements than with farmer-reported yields [42, 43, 44], consistent with survey data underestimating true weather-driven yield variability. While the magnitude of true yield losses remains uncertain, the weight of evidence suggests that inferred weather sensitivity is attenuated in survey-based estimates, and that the choice of yield data source materially affects climate impact assessments.

4 Discussion

Correspondence between survey and remotely sensed yields is stronger in countries with extensive, well-documented agricultural systems, which tend to be those with higher GDP per capita and larger total harvested-area fractions. Where correspondence is weaker, the presence of flagged FAO data and frequent reporting-year offsets points to challenges in survey reporting, although satellite-based proxies are also affected by imperfect crop masking and retrieval artifacts. The GDP effect is statistically significant but modest in absolute terms: a one-standard deviation increase in GDP per capita is associated with a 0.02 increase in ρ_{VS}^2 , compared to a reduction of 0.09 associated with reporting-year offsets. Reporting practices and phenological mismatches, rather than economic development per se, are the strongest predictors of survey–satellite disagreement in these data. However, because GDP per capita covaries with other predictors in the model – richer countries tend to have more extensive cropland, fewer flagged observations, and fewer reporting-year offsets – the partial effect of 0.02 understates the total association between economic development and survey–satellite correspondence (see Supplementary Table S7). Further, survey yields appear to underestimate weather-driven variability relative to satellite-derived yields. We note that country-level weather variables may not capture the weather that actually drives yield variability, but may instead proxy for finer-scale processes such as water limitation operating below the country scale [10].

The systematic association between survey data quality indicators and weaker weather explainability documented here, together with the plot-level evidence that satellites accord better with crop-cut measurements than with farmer reports [42, 43, 44], points toward attenuation in survey-reported yields. Measurement error in survey yields can arise from interpolation of missing data, statistical processing by national agencies, and under-reporting of poor harvest years [39, 40]. Interpreting the socioeconomic gradient in the survey–satellite gap is complicated by the fact that the same factors predicting better data quality also likely predict greater buffering capacity

through irrigation, mechanization, and management. Disentangling measurement error from genuine adaptation requires independent ground-truth data not currently available at global scale. Country-level aggregation may also mask subnational heterogeneity in both data quality and climate sensitivity; analyses at finer spatial scales, where data permit, would be a valuable complement.

Important results in the climate–yield literature rely heavily on survey-reported yields from well-instrumented countries [1], and estimates of the cumulative impact of anthropogenic climate change on global agricultural productivity draw on FAO-derived data [56]. If survey yields in low-income countries contain substantial non-weather noise that attenuates estimated climate sensitivity, as our results suggest, then global studies likely underestimate true climate-driven yield losses. Satellite-based estimates could overstate sensitivity in the opposite direction if CSIF responds to canopy stress that does not proportionally reduce harvested yields, so true sensitivity probably lies between the two. This uncertainty is most consequential in Sub-Saharan Africa, South Asia, and other tropical regions where food insecurity and data scarcity coincide.

Neither data source is currently sufficient on its own. Satellite estimates require ground-based calibration to be expressed in yield units [57], and combined survey–satellite products [48, 58, 59, 60, 61] inherit survey biases if those biases are unquantified. The most direct path forward is improved uncertainty quantification for survey yields and systematic validation of satellite proxies in regions and crops where correspondence is weakest. Subnational compilations such as HarvestStat [24, 62], higher-resolution SIF products from TROPOMI [63] and the forthcoming FLEX mission, and scalable satellite-based yield estimation frameworks [64] all contribute to this agenda. Accurate determination of how much global agriculture is susceptible to weather and climate, and how much adaptation is already manifest, depends on resolving these measurement questions.

Acknowledgments

We thank Angela Rigden for database assistance, and Sam Myers and Matt Smith for advising the early stages of this project.

Data availability

https://github.com/carospark/FA0_ag_check_final

References

- [1] W. Schlenker and M. J. Roberts, “Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change,” *Proceedings of the National Academy of Sciences*, vol. 106, no. 37, pp. 15594–15598, 2009.
- [2] D. B. Lobell, W. Schlenker, and J. Costa-Roberts, “Climate trends and global crop production since 1980,” *Science*, vol. 333, no. 6042, pp. 616–620, 2011.
- [3] D. B. Lobell and S. M. Gourdji, “The influence of climate change on global crop productivity,” *Plant physiology*, vol. 160, no. 4, pp. 1686–1697, 2012.
- [4] IPCC, “Managing the risks of extreme events and disasters to advance climate change adaptation.” Intergovernmental Panel on Climate Change, 2012. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change.
- [5] E. E. Butler and P. Huybers, “Adaptation of US maize to temperature variations,” *Nature Climate Change*, vol. 3, no. 1, pp. 68–72, 2013.
- [6] E. Grigorieva, A. Livenets, and E. Stelmakh, “Adaptation of agriculture to climate change: A scoping review,” *Climate*, vol. 11, no. 10, p. 202, 2023.
- [7] A. Hultgren, T. Carleton, M. Delgado, D. R. Gergel, M. Greenstone, T. Houser, S. Hsiang, A. Jina, R. E. Kopp, S. B. Malevich, *et al.*, “Impacts of climate change on global agriculture accounting for adaptation,” *Nature*, vol. 642, no. 8068, pp. 644–652, 2025.
- [8] L. R. Vargas Zeppetello, J. Proctor, and P. Huybers, “Is water stress the root cause of the observed nonlinear relationship between yield losses and temperature?,” *AGU Advances*, vol. 6, no. 4, p. e2025AV001704, 2025.
- [9] D. K. Ray, J. S. Gerber, G. K. MacDonald, and P. C. West, “Climate variation explains a third of global crop yield variability,” *Nature communications*, vol. 6, no. 1, p. 5989, 2015.

- [10] J. Proctor, A. Rigden, D. Chan, and P. Huybers, “More accurate specification of water supply shows its importance for global crop production,” *Nature Food*, vol. 3, no. 9, pp. 753–763, 2022.
- [11] J. Proctor, L. Vargas Zeppetello, D. Chan, and P. Huybers, “Climate change increases the interannual variance of summer crop yields globally through changes in temperature and water supply,” *Science Advances*, vol. 11, no. 36, p. eady3575, 2025.
- [12] C. Lesk, P. Rowhani, and N. Ramankutty, “Influence of extreme weather disasters on global crop production,” *Nature*, vol. 529, no. 7584, pp. 84–87, 2016.
- [13] D. Neupane, P. Adhikari, D. Bhattarai, B. Rana, Z. Ahmed, U. Sharma, and D. Adhikari, “Does climate change affect the yield of the top three cereals and food security in the world?,” *Earth*, vol. 3, no. 1, pp. 45–71, 2022.
- [14] P. Agnolucci, C. Rapti, P. Alexander, V. De Lipsis, R. A. Holland, F. Eigenbrod, and P. Ekins, “Impacts of rising temperatures and farm management practices on global yields of 18 crops,” *Nature Food*, vol. 1, no. 9, pp. 562–571, 2020.
- [15] Food and Agriculture Organization of the United Nations, “FAOSTAT statistical database.” <https://www.fao.org/faostat/en/>, 2024. Accessed 2025-05-05.
- [16] Food and Agriculture Organization of the United Nations, “Master sampling frames for agricultural statistics,” tech. rep., FAO Statistics Division, Rome, Italy, 2015. Accessed 2025-05-05.
- [17] Food and Agriculture Organization of the United Nations, “Technical report on the integrated survey framework,” tech. rep., FAO Statistics Division, Rome, Italy, 2017. Accessed 2025-05-05.
- [18] P. Grassini, L. G. van Bussel, J. Van Wart, J. Wolf, L. Claessens, H. Yang, H. Boogaard, H. de Groot, M. K. van Ittersum, and K. G. Cassman, “How good is good enough? data requirements for reliable crop yield simulations and yield-gap analysis,” *Field Crops Research*, vol. 177, pp. 49–63, 2015.
- [19] M. Jerven, *Poor numbers: how we are misled by African development statistics and what to do about it*. Cornell University Press, 2013.
- [20] S. Desiere, L. Staelens, and M. D’Haese, “When the data source writes the conclusion: evaluating agricultural policies,” *The Journal of Development Studies*, vol. 52, no. 9, pp. 1372–1387, 2016.
- [21] C. Carletto, A. Dillon, and A. Zezza, “Agricultural data collection to minimize measurement error and maximize coverage,” in *Handbook of agricultural economics*, vol. 5, pp. 4407–4480, Elsevier, 2021.
- [22] H. Ayalew, J. Chamberlin, C. Newman, K. A. Abay, F. Kosmowski, and T. Sida, “Revisiting the size–productivity relationship with imperfect measures of production and plot size,” *American Journal of Agricultural Economics*, vol. 106, no. 2, pp. 595–619, 2024.
- [23] N. S. Diffenbaugh and M. Burke, “Global warming has increased global economic inequality,” *Proceedings of the National Academy of Sciences*, vol. 116, no. 20, pp. 9808–9813, 2019.
- [24] D. Lee, W. Anderson, X. Chen, F. Davenport, S. Shukla, R. Sahajpal, M. Budde, J. Rowland, J. Verdin, L. You, M. Ahouangbenon, K. F. Davis, E. Kebede, S. Ehrmann, C. Justice, and C. Meyer, “HarvestStat Africa – harmonized subnational crop statistics for Sub-Saharan Africa,” *Scientific Data*, vol. 12, p. 690, 2025.
- [25] M. Mkhabela, P. Bullock, S. Raj, S. Wang, and Y. Yang, “Crop yield forecasting on the Canadian Prairies using MODIS NDVI data,” *Agricultural and Forest Meteorology*, vol. 151, no. 3, pp. 385–393, 2011.
- [26] M. Roznik, M. Boyd, and L. Porth, “Improving crop yield estimation by applying higher resolution satellite NDVI imagery and high-resolution cropland masks,” *Remote Sensing Applications: Society and Environment*, vol. 25, p. 100693, 2022.

- [27] I. Becker-Reshef, C. Justice, B. Barker, M. Humber, F. Rembold, R. Bonifacio, M. Zappacosta, M. Budde, T. Magadzire, C. Shitote, *et al.*, “Strengthening agricultural decisions in countries at risk of food insecurity: The GEOGLAM crop monitor for early warning,” *Remote Sensing of Environment*, vol. 237, p. 111553, 2020.
- [28] B. Peng, K. Guan, W. Zhou, C. Jiang, C. Frankenberg, Y. Sun, L. He, and P. Köhler, “Assessing the benefit of satellite-based solar-induced chlorophyll fluorescence in crop yield prediction,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 90, p. 102126, 2020.
- [29] G. H. Mohammed, R. Colombo, E. M. Middleton, U. Rascher, C. Van Der Tol, L. Nedbal, Y. Goulas, O. Pérez-Priego, A. Damm, M. Meroni, *et al.*, “Remote sensing of solar-induced chlorophyll fluorescence (SIF) in vegetation: 50 years of progress,” *Remote sensing of environment*, vol. 231, p. 111177, 2019.
- [30] L. He, T. Magney, D. Dutta, Y. Yin, P. Köhler, K. Grossmann, J. Stutz, C. Dold, J. Hatfield, K. Guan, *et al.*, “From the ground to space: Using solar-induced chlorophyll fluorescence to estimate crop productivity,” *Geophysical Research Letters*, vol. 47, no. 7, p. e2020GL087474, 2020.
- [31] Z. Li, K. Guan, and B. Peng, “Estimating crop yield using solar-induced chlorophyll fluorescence: A comparison with NDVI and EVI across U.S. corn and soybean fields,” *Remote Sensing of Environment*, vol. 268, p. 112760, 2022.
- [32] K. Guan, J. Wu, J. S. Kimball, M. C. Anderson, S. Froking, B. Li, C. R. Hain, and D. B. Lobell, “The shared and unique values of optical, fluorescence, thermal and microwave satellite data for estimating large-scale crop yields,” *Remote sensing of environment*, vol. 199, pp. 333–349, 2017.
- [33] K. Guan, J. A. Berry, Y. Zhang, J. Joiner, L. Guanter, G. Badgley, and D. B. Lobell, “Improving the monitoring of crop productivity using spaceborne solar-induced fluorescence,” *Global change biology*, vol. 22, no. 2, pp. 716–726, 2016.
- [34] T. S. Magney, D. R. Bowling, B. A. Logan, K. Grossmann, J. Stutz, P. D. Blanken, S. P. Burns, R. Cheng, M. A. Garcia, P. Köhler, *et al.*, “Mechanistic evidence for tracking the seasonality of photosynthesis with solar-induced fluorescence,” *Proceedings of the National Academy of Sciences*, vol. 116, no. 24, pp. 11640–11645, 2019.
- [35] D. B. Lobell, “The use of satellite data for crop yield gap analysis,” *Field Crops Research*, vol. 143, pp. 56–64, 2013.
- [36] M. Zhao and S. W. Running, “Drought-induced reduction in global terrestrial net primary production from 2000 through 2009,” *Science*, vol. 329, no. 5994, pp. 940–943, 2010.
- [37] Z. Jin, G. Azzari, and D. B. Lobell, “Improving the accuracy of satellite-based high-resolution yield estimation: A test of multiple scalable approaches,” *Remote Sensing of Environment*, vol. 202, pp. 400–412, 2017.
- [38] Y. Zhang, J. Joiner, S. H. Alemohammad, S. Zhou, and P. Gentine, “A global spatially contiguous solar-induced fluorescence (CSIF) dataset using neural networks,” *Biogeosciences*, vol. 15, no. 19, pp. 5779–5800, 2018.
- [39] C. Carletto, D. Jolliffe, and R. Banerjee, “From tragedy to renaissance: Improving agricultural data for better policies,” *Journal of Development Studies*, vol. 51, no. 2, pp. 133–148, 2015.
- [40] K. A. Abay, T. Wossen, G. T. Abate, J. R. Stevenson, H. Michelson, and C. B. Barrett, “Inferential and behavioral implications of measurement error in agricultural data,” *Annual Review of Resource Economics*, vol. 15, pp. 63–83, 2023.
- [41] J. Proctor, T. Carleton, and S. Sum, “Parameter recovery using remotely sensed variables,” tech. rep., National Bureau of Economic Research, 2023.
- [42] S. Gourlay, T. Kilic, and D. B. Lobell, “A new spin on an old debate: Errors in farmer-reported production and their implications for inverse scale-productivity relationship in Uganda,” *Journal of Development Economics*, vol. 141, p. 102376, 2019.

- [43] D. B. Lobell, G. Azzari, M. Burke, S. Gourlay, Z. Jin, T. Kilic, and S. Murray, “Eyes in the sky, boots on the ground: Assessing satellite- and ground-based approaches to crop yield measurement and analysis,” *American Journal of Agricultural Economics*, vol. 102, no. 1, pp. 202–219, 2020.
- [44] D. B. Lobell, S. Di Tommaso, C. You, I. Yacoubou Djima, M. Burke, and T. Kilic, “Sight for sorghums: Comparisons of satellite- and ground-based sorghum yield estimates in Mali,” *Remote Sensing*, vol. 12, no. 1, p. 100, 2020.
- [45] B. Schauburger, S. Archontoulis, A. Arneth, J. Balkovic, P. Ciais, D. Deryng, J. Elliott, C. Folberth, N. Khabarov, C. Müller, T. A. M. Pugh, S. Rolinski, S. Schaphoff, E. Schmid, X. Wang, W. Schlenker, and K. Frieler, “Consistent negative response of US crops to high temperatures in observations and crop models,” *Nature Communications*, vol. 8, p. 13931, 2017.
- [46] N. Ramankutty, A. T. Evan, C. Monfreda, and J. A. Foley, “Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000,” *Global biogeochemical cycles*, vol. 22, no. 1, 2008.
- [47] W. J. Sacks, D. Deryng, J. A. Foley, and N. Ramankutty, “Crop planting dates: an analysis of global patterns,” *Global ecology and biogeography*, vol. 19, no. 5, pp. 607–620, 2010.
- [48] M. Burke and D. B. Lobell, “Satellite-based assessment of yield variation and its determinants in smallholder African systems,” *Proceedings of the National Academy of Sciences*, vol. 114, no. 9, pp. 2189–2194, 2017.
- [49] A. Gruber, T. Scanlon, R. Van Der Schalie, W. Wagner, and W. Dorigo, “Evolution of the ESA CCI soil moisture climate data records and their underlying merging methodology,” *Earth System Science Data*, vol. 11, no. 2, pp. 717–739, 2019.
- [50] W. Dorigo, W. Wagner, C. Albergel, F. Albrecht, G. Balsamo, L. Brocca, D. Chung, M. Ertl, M. Forkel, A. Gruber, *et al.*, “ESA CCI soil moisture for improved Earth system understanding: State-of-the art and future directions,” *Remote Sensing of Environment*, vol. 203, pp. 185–215, 2017.
- [51] The World Bank, “World Bank country and lending groups,” 2023. Accessed 4 Jan 2024.
- [52] M. Friedl and D. Sulla-Menashe, *MCD12C1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 0.05Deg CMG V006*, 2015. distributed by NASA EOSDIS Land Processes DAAC, <https://doi.org/10.5067/MODIS/MCD12C1.006>. Accessed 2022-07-27.
- [53] G. Miao, K. Guan, A. E. Suyker, X. Yang, T. J. Arkebauer, E. A. Walter-Shea, H. Kimm, G. Y. Hmimina, J. A. Gamon, T. E. Franz, *et al.*, “Varying contributions of drivers to the relationship between canopy photosynthesis and far-red sun-induced fluorescence for two maize sites at different temporal scales,” *Journal of Geophysical Research: Biogeosciences*, vol. 125, no. 2, p. e2019JG005051, 2020.
- [54] P. Yang, C. van der Tol, W. Verhoef, A. Damm, A. Schickling, T. Kraska, O. Muller, and U. Rascher, “Using reflectance to explain vegetation biochemical and structural effects on sun-induced chlorophyll fluorescence,” *Remote sensing of environment*, vol. 231, p. 110996, 2019.
- [55] M. C. Whitlock, “Combining probability from independent tests: the weighted Z-method is superior to Fisher’s approach,” *Journal of Evolutionary Biology*, vol. 18, no. 5, pp. 1368–1373, 2005.
- [56] A. Ortiz-Bobea, T. R. Ault, C. M. Carrillo, R. G. Chambers, and D. B. Lobell, “Anthropogenic climate change has slowed global agricultural productivity growth,” *Nature Climate Change*, vol. 11, pp. 306–312, 2021.
- [57] R. P. Sishodia, R. L. Ray, and S. K. Singh, “Applications of remote sensing in precision agriculture: A review,” *Remote sensing*, vol. 12, no. 19, p. 3136, 2020.
- [58] J. Gogoi, N. K. Newlands, Z. Mehrabi, N. C. Coops, and N. Ramankutty, “Assessing the performance of satellite-based models for crop yield estimation in the Canadian Prairies,” *Canadian Journal of Remote Sensing*, vol. 49, no. 1, p. 2252926, 2023.

- 496 [59] J. Jägermeyr, C. Müller, A. C. Ruane, J. Elliott, J. Balkovic, O. Castillo, B. Faye, I. Foster,
497 C. Folberth, J. A. Franke, *et al.*, “Climate impacts on global agriculture emerge earlier in new
498 generation of climate and crop models,” *Nature Food*, vol. 2, no. 11, pp. 873–885, 2021.
- 499 [60] S. Di Tommaso, S. Wang, V. Vajipey, N. Gorelick, R. Strey, and D. B. Lobell, “Annual
500 field-scale maps of tall and short crops at the global scale using GEDI and Sentinel-2,” *Remote
501 Sensing*, vol. 15, no. 17, p. 4123, 2023.
- 502 [61] J. L. Soler, T. Friedel, and S. Wang, “Combining deep learning and street view imagery to
503 map smallholder crop types,” *arXiv preprint arXiv:2309.05930*, 2023.
- 504 [62] K. F. Davis, W. Anderson, S. Ehrmann, R. Flach, C. Meyer, J. Proctor, D. K. Ray, L. You,
505 M. Foley, H. Kerdiles, A. Hultgren, P. Huybers, E. Kebede, M. Meroni, C. Park, and D. Lee,
506 “HarvestStat: A global effort towards open and standardized sub-national agricultural data,”
507 *Environmental Research Letters*, vol. 20, no. 5, p. 052001, 2025.
- 508 [63] L. Guanter, C. Bacour, A. Schneider, I. Aben, T. A. van Kempen, F. Maignan, C. Retscher,
509 P. Köhler, C. Frankenberg, J. Joiner, and Y. Zhang, “The TROPOSIF global sun-induced
510 fluorescence dataset from the Sentinel-5P TROPOMI mission,” *Earth System Science Data*,
511 vol. 13, no. 11, pp. 5423–5440, 2021.
- 512 [64] O. Kira, J. Wen, J. Han, A. J. McDonald, C. B. Barrett, A. Ortiz-Bobea, Y. Liu, L. You,
513 N. D. Mueller, and Y. Sun, “A scalable crop yield estimation framework based on remote
514 sensing of solar-induced chlorophyll fluorescence (SIF),” *Environmental Research Letters*,
515 vol. 19, p. 044071, 2024.

516 **5 Supplementary**

ρ_{VS}^2 (survey–satellite corr. ²) – crop as <i>fixed</i> effect			
Predictors	Estimates	CI	p
Intercept	0.15 ***	0.12 – 0.18	<0.001
Avg GDP per capita (log scale)	0.02 ***	0.01 – 0.03	<0.001
Survey flag pct	-0.00	-0.01 – 0.01	0.75
Offset (year $\pm$ 1)	-0.09 ***	-0.11 – -0.08	<0.001
Total harvested-area fraction	0.01 ***	0.01 – 0.02	<0.001
Mean cropland fraction	0.02 ***	0.01 – 0.03	<0.001
Crop fixed effects			
	reference: Barley		
Cassava	-0.00	-0.04 – 0.04	0.973
Groundnuts, with shell	0.01	-0.03 – 0.04	0.755
Maize	0.07 ***	0.03 – 0.10	<0.001
Millet	0.01	-0.04 – 0.05	0.809
Oats	-0.02	-0.06 – 0.02	0.305
Potatoes	-0.02	-0.06 – 0.02	0.283
Pulses	0.02	-0.02 – 0.06	0.356
Rapeseed	-0.03	-0.08 – 0.01	0.124
Rice, paddy	-0.01	-0.05 – 0.03	0.667
Rye	-0.04 +	-0.09 – 0.00	0.065
Seed cotton	0.01	-0.04 – 0.05	0.785
Sorghum	0.02	-0.02 – 0.05	0.425
Soybeans	0.07 ***	0.03 – 0.10	<0.001
Sugar beet	0.06 **	0.02 – 0.11	0.006
Sunflower seed	0.03	-0.02 – 0.07	0.235
Sweet potatoes	-0.01	-0.05 – 0.03	0.787
Wheat	0.02	-0.02 – 0.06	0.327
Yams	0.00	-0.05 – 0.05	0.887
Model fit			
R ² / adj. R ²	0.221 / 0.211		
N cropname	19		
Observations	1744		

+ p < 0.1 * p < 0.05 ** p < 0.01 *** p < 0.001

Table S1: Fixed-effects model estimates for predictors of survey–satellite yield correspondence. The dependent variable is the squared correlation between survey-derived and satellite-derived yields, ρ_{VS}^2 , computed for each country–crop pair. The model includes crop-level fixed intercepts.

ρ_{VS} : correlation of survey & satellite (supplementary)			
Predictors	Estimates	CI	p
Intercept	0.30 ***	0.27 – 0.32	<0.001
Avg GDP per capita (log scale)	0.01 +	-0.00 – 0.02	0.073
Survey flag pct	-0.03 ***	-0.04 – -0.01	<0.001
Offset (year $\pm$ 1)	-0.34 ***	-0.36 – -0.32	<0.001
Total harvested-area fraction	0.02 ***	0.01 – 0.03	<0.001
Mean cropland fraction	0.03 ***	0.02 – 0.04	<0.001
Random Effects			
σ^2	0.055		
τ_0 cropname	0.001		
ICC	0.026		
N cropname	19		
Observations	1744		
Marginal R2 / Conditional R2	0.390 / 0.406		
+ p < 0.1 * p < 0.05 ** p < 0.01 *** p < 0.001			

Table S2: Mixed-effects model estimates for predictors of survey–satellite yield correlation. The dependent variable is the unsquared Pearson correlation ρ_{VS} between survey-derived and satellite-derived yields, computed for each country–crop pair. Specification is otherwise identical to the main model in Table 2, which uses ρ_{VS}^2 as the outcome. The model includes crop-level random intercepts and is estimated on $N = 1,744$ country–crop observations across 19 crops. Results are broadly consistent with the main ρ_{VS}^2 specification: survey–satellite correspondence is reduced for country–crop pairs with reporting-year offsets, and increased with greater total harvested-area fraction and mean cropland fraction. The share of FAO-flagged observations is negatively associated with correspondence in this specification ($p < 0.001$), whereas it is not independently significant in the ρ_{VS}^2 model. The intraclass correlation coefficient (ICC) of 0.026 indicates that crop type accounts for a small share of residual variance.

Harvest-year offset diagnostic (N = 1,744 country–crop pairs)			
Assigned alignment	N	% of pairs	Sig. > unshifted ($p < 0.1$)
Unshifted (no offset)	789	45.2	–
Lag (–1 yr)	434	24.9	39
Lead (+1 yr)	521	29.9	34
Any offset (lead or lag)	955	54.8	73 (7.6% of offset)
Total	1,744	100.0	

Table S3: Harvest-year offset diagnostic. For each country–crop pair the CSIF proxy is correlated against the unshifted detrended log yield and its ± 1 -year shifts; the highest-correlation alignment is assigned. The final column counts offset pairs whose best-shifted correlation is significantly greater than the unshifted correlation (two-tailed Fisher’s Z, $p < 0.1$; the test requires $n > 3$, excluding one no-offset pair with $n = 3$ from that column only).

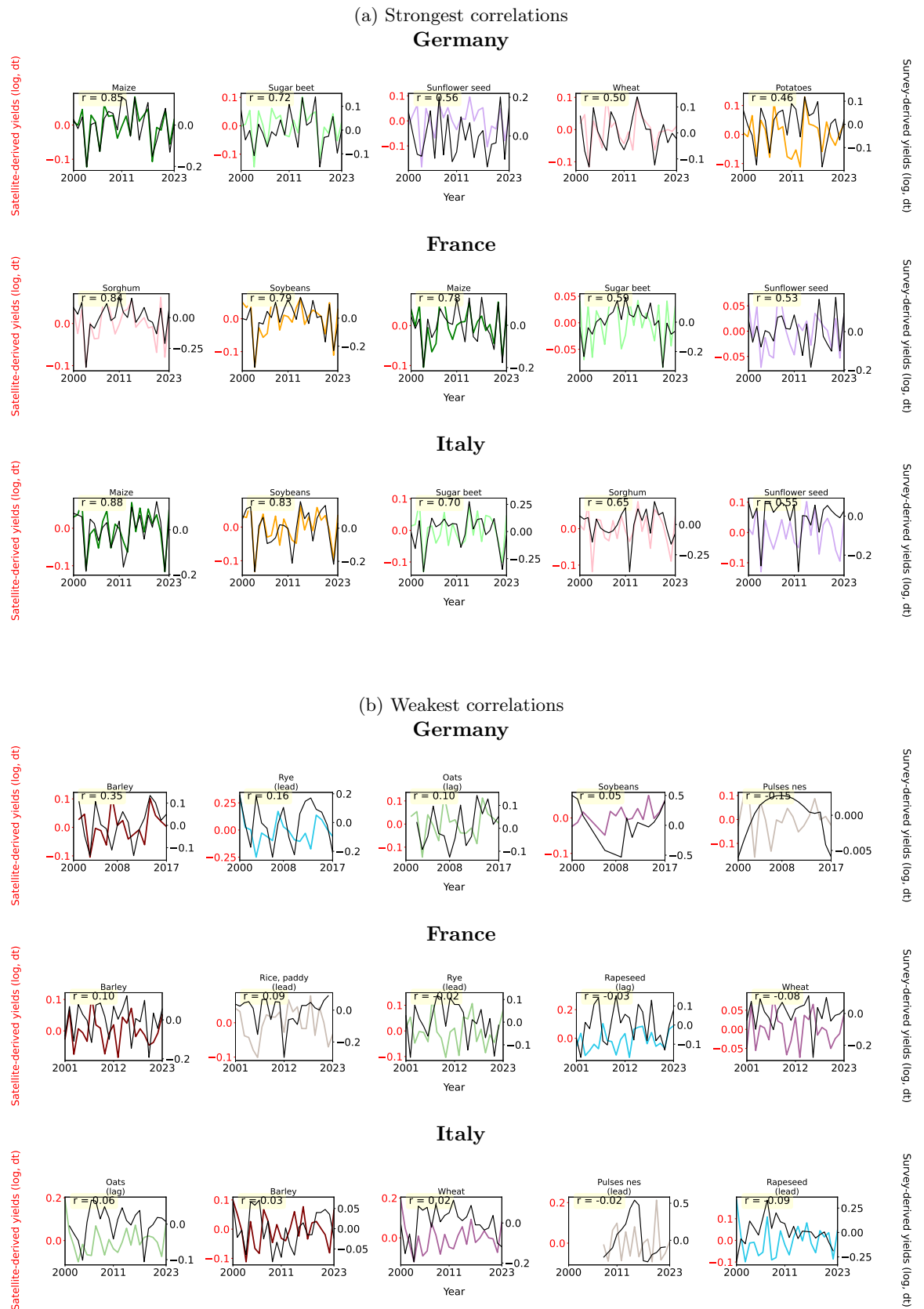


Figure S1: Time series of FAO-reported yields (black lines) and the CSIF proxy (colored lines) for various crops in Germany, France, and Italy, ordered by strength of correlation. (a) Strongest and (b) weakest survey–satellite correlations; within each panel, rows show Germany, France, and Italy (top to bottom). Pearson correlation coefficient is shown in the top-left corner of each subpanel.

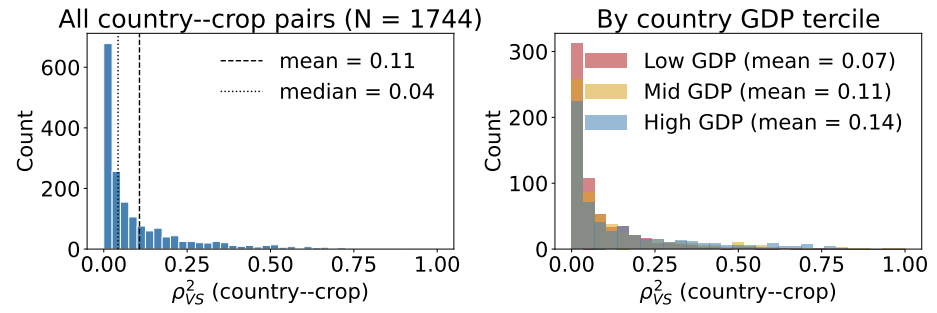


Figure S2: Distribution of country–crop survey–satellite squared correlations (ρ_{VS}^2). (a) Histogram across all $N = 1,744$ country–crop pairs. The marginal mean (0.11, dashed line) and median (dotted line) values are indicated. The distribution is right-skewed, with the bulk of country–crop pairs falling below $\rho_{VS}^2 = 0.2$ and a long tail extending toward strong correspondence ($\rho_{VS}^2 > 0.6$). The mean (0.11) is lower than the Table 2 model intercept (0.16) because the intercept conditions on no reporting-year offset, whereas the majority of country–crop pairs in the sample exhibit a ± 1 -year offset, which is associated with substantially lower correspondence. (b) The same distribution split by country-level GDP per capita tertile (Low / Mid / High); per-tertile means are shown in the legend. Survey–satellite correspondence shifts noticeably toward higher values in higher-GDP countries, consistent with the mixed-effects results in Table 2.

		R2	Adj R2	F-test pass	SM coefficient	SM p.v. pass
All crops	Satellite	0.32	0.25	60%	2.66 (0.89, 5.51)	49%
	Survey	0.12	0.03	20%	0.81 (−1.24, 4.59)	15%
Barley	Satellite	0.31	0.25	56%	4.02 (1.18, 6.5)	53%
	Survey	0.14	0.05	24%	1.49 (−1.2, 5.71)	17%
Cassava	Satellite	0.21	0.13	39%	1.04 (0.35, 2.68)	34%
	Survey	0.11	0.01	16%	0.24 (−0.51, 2.11)	11%
Groundnuts, with shell	Satellite	0.41	0.35	72%	3.7 (1.33, 6.44)	59%
	Survey	0.11	0.01	11%	0.64 (−0.67, 4.01)	8%
Maize	Satellite	0.38	0.32	68%	2.54 (1.36, 5.27)	52%
	Survey	0.16	0.07	27%	1.65 (−0.83, 7.11)	19%
Millet	Satellite	0.38	0.32	62%	2.84 (1.35, 6.84)	45%
	Survey	0.18	0.09	26%	3.05 (−1.09, 7.15)	20%
Oats	Satellite	0.27	0.2	58%	4.45 (1.31, 5.96)	55%
	Survey	0.16	0.08	33%	1.18 (−0.18, 5.82)	18%
Potatoes	Satellite	0.26	0.18	50%	3.13 (1.25, 6.0)	45%
	Survey	0.09	−0.01	10%	0.63 (−0.86, 2.93)	11%
Pulses	Satellite	0.27	0.2	52%	1.86 (0.13, 4.77)	36%
	Survey	0.1	0.01	19%	−0.07 (−1.65, 2.04)	16%
Rapeseed	Satellite	0.26	0.19	52%	1.38 (−0.77, 4.22)	28%
	Survey	0.1	0.01	11%	1.49 (−1.39, 5.21)	10%
Rice, paddy	Satellite	0.3	0.24	58%	2.08 (0.64, 4.77)	42%
	Survey	0.12	0.03	18%	0.77 (−1.1, 4.33)	15%
Rye	Satellite	0.27	0.2	60%	1.86 (−0.65, 3.98)	37%
	Survey	0.1	0.01	13%	0.41 (−2.22, 4.62)	13%
Seed cotton	Satellite	0.38	0.32	67%	3.04 (1.19, 6.16)	59%
	Survey	0.12	0.03	20%	1.11 (−2.05, 5.59)	13%
Sorghum	Satellite	0.36	0.3	68%	2.96 (1.65, 6.42)	61%
	Survey	0.13	0.05	19%	1.66 (−1.57, 5.65)	17%
Soybeans	Satellite	0.35	0.29	63%	2.32 (0.7, 4.64)	44%
	Survey	0.14	0.06	27%	2.34 (−1.54, 5.17)	16%
Sugar beet	Satellite	0.36	0.3	64%	2.59 (0.58, 5.17)	42%
	Survey	0.12	0.03	24%	1.1 (−1.03, 5.35)	11%
Sunflower seed	Satellite	0.42	0.36	74%	3.7 (1.96, 6.58)	61%
	Survey	0.14	0.03	21%	0.99 (−2.64, 4.66)	14%
Sweet potatoes	Satellite	0.36	0.3	59%	2.01 (0.7, 4.68)	55%
	Survey	0.1	0.01	9%	−0.14 (−2.09, 2.0)	12%
Wheat	Satellite	0.3	0.23	55%	3.26 (1.25, 7.01)	52%
	Survey	0.14	0.05	27%	0.92 (−1.7, 6.0)	21%
Yams	Satellite	0.33	0.26	61%	2.82 (1.4, 5.25)	57%
	Survey	0.14	0.04	20%	0.17 (−1.32, 2.14)	11%

Table S4: Soil-moisture (SM) and maximum-temperature explainability of survey vs. satellite yield variability, by crop. R2 and Adj R2 are medians across country-level regressions; the SM coefficient is the median (Q25, Q75); “pass” columns give the share of country regressions significant at $p < 0.05$.

Per-crop satellite yield variability explained by SM and TMAX

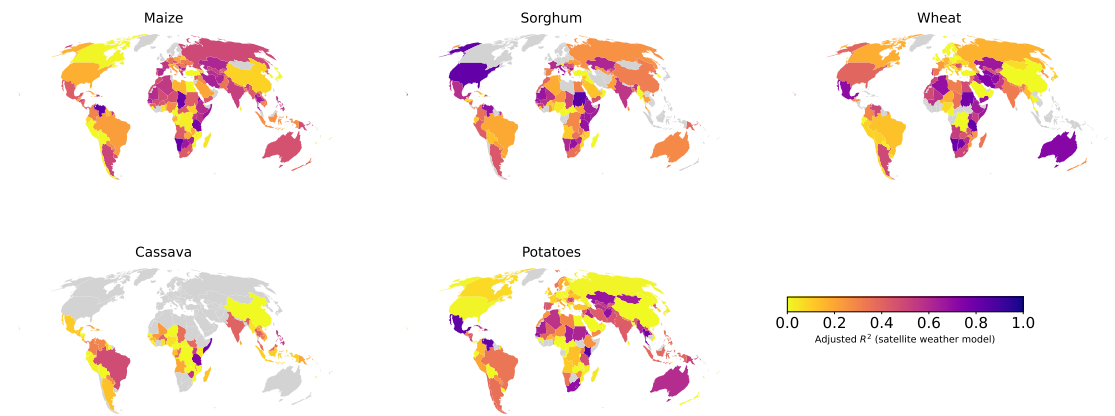


Figure S3: Per-crop satellite-derived yield variability explained by growing-season soil moisture and maximum temperature. Each panel shows the country-level adjusted R^2 from a linear regression of detrended log CSIF on detrended growing-season soil moisture and maximum temperature, fit separately for each country–crop pair with at least 10 years of data. The five crops shown (Maize, Sorghum, Wheat, Cassava, Potatoes) are those reported in Table 3. The same broad spatial pattern observed in Figure 4 – higher weather-explained variability across Sub-Saharan Africa, parts of the Middle East, and South America, and lower values across mid- to high-latitude regions – is visible across crops, indicating that the cross-crop average in Figure 4a is not driven by any single crop. Countries lacking sufficient data for a given crop are shown in grey.

Univariate OLS: (Satellite R^2 – Survey R^2) $\sim$ avg GDP per capita					
Sample	$\beta_{\log GDP}$	SE	CI	p	N
Overall	-0.025 *	0.010	-0.045 – -0.005	0.016	168
Tropical	-0.037 +	0.022	-0.080 – 0.006	0.087	83
Non-tropical	-0.025 *	0.012	-0.048 – -0.002	0.037	85

+ p < 0.1 * p < 0.05 ** p < 0.01 *** p < 0.001

Table S5: Univariate OLS regressions of the country-level weather-explainability gap (Satellite R^2 – Survey R^2 , averaged across crops) on log GDP per capita. Estimates are reported overall and separately for tropical (absolute latitude $\leq 23.5^\circ$) and non-tropical countries. A one-unit increase in log GDP per capita corresponds to multiplying GDP by $e \approx 2.72$; equivalently, a 10-fold increase in GDP per capita is associated with a $\beta \times \ln(10) \approx 2.30 \beta$ change in the gap. Results indicate that the survey–satellite gap shrinks as GDP per capita rises, with a similar slope in tropical and non-tropical samples.

Joint OLS: (Satellite R^2 – Survey R^2) $\sim$ log GDP + Tropical				
Predictor	Estimate	SE	CI	p
Intercept	0.452 ***	0.111	0.233 – 0.671	<0.001
log GDP/capita	-0.031 *	0.012	-0.054 – -0.007	0.011
Tropical (1/0)	-0.032	0.035	-0.101 – 0.036	0.356
R^2 / Adj R^2	0.040 / 0.028			
F-statistic (2, 165)	3.41			p = 0.035
Observations	168			
+ p < 0.1 * p < 0.05 ** p < 0.01 *** p < 0.001				

Table S6: Joint OLS regression of the country-level weather-explainability gap (Satellite R^2 – Survey R^2 , averaged across crops) on log GDP per capita and a tropical-region indicator (1 if absolute latitude $\leq 23.5^\circ$, 0 otherwise). The log GDP coefficient remains negative and significant after controlling for tropical status, indicating that the income gradient in the survey–satellite gap is not simply a tropical/non-tropical artifact. The tropical indicator is not independently significant.

Partial vs. total association of GDP per capita with ρ_{VS}^2			
Model	Estimate	95% CI	p
Full model (all predictors)	0.021 ***	0.014 – 0.027	<0.001
GDP only (crop intercept retained)	0.028 ***	0.021 – 0.035	<0.001
+ p < 0.1 * p < 0.05 ** p < 0.01 *** p < 0.001			

Table S7: Partial versus total association between GDP per capita (standardized, log scale) and survey–satellite correspondence (ρ_{VS}^2). Both models are linear mixed-effects models with crop random intercepts; the “full” model includes all covariates from Table 2, while “GDP only” drops the other fixed-effect predictors. The GDP coefficient grows by 33% (0.021 to 0.028) when the covarying predictors are removed, consistent with the partial effect understating the total association between economic development and correspondence.