

Supplementary Information

Linking Hydroclimate to Agricultural Input-Carbon and Irrigation-Water Productivity: A Source-Audited Province-Year Analysis across China

Aoqi Xu

Shan Miao

Zhiming Guo

Supplementary data architecture and validation

Table S1: Controlled data layers and validation status

Layer	Required contents	Validation status
Raw	Original files and source manifest	Thirteen source objects are hash-registered, including the archived NBS agricultural responses. The raw files remain unchanged, and each source block retains its identifiers, hashes, and processing-rule versions.
Intermediate	Separately validated hydrological and agricultural province-year files	The hydrological file contains 403 complete records. The agricultural file contains 403 unique keys, 403 IWP values, and 310 observed AICP values.
Processed	Hydro-agricultural panel, coverage report, and panel hash	The analytical panel contains 403 one-to-one matched records. AICP remains missing where an official component is unavailable; no primary value is statistically imputed.
Outputs	Script-generated validation, summary, lineage, feasibility, and association outputs	Hydrological, agricultural, merged-panel, carbon-market feasibility, and hydroclimatic association outputs are generated from archived code.

Table S2: Verification status of the data blocks used in the analysis

Data block	Inspected source	Intended analytical role	Status
Administrative boundaries	geoBoundaries CHN ADM1 object CHN-ADM1-43563684; representative year 2019; build 12 December 2023 [3, 6]	Fixed crosswalk; 31 mainland units retained; Hong Kong, Macao, and Taiwan excluded; Guangdong and Ningxia labels normalized	Aggregation validated
Irrigation-water use	Figshare HSWUD version 7; HSWUD_irr_v20260711.nc [9, 10]	Variable irr ; 696 monthly steps for 1965–2022; raw unit 10^8 m^3 per grid-month; 2010–2022 subset constructed	Constructed and validated
Precipitation and PET	Copernicus CDS ERA5 monthly averaged data on single levels; 2010–2022 subset [1, 4]	avg.tprate and avg.pevr ; $\text{kg m}^{-2} \text{ s}^{-1}$; exact-calendar rate conversion and PET sign rule applied	Constructed and validated
Drought	ERA5–Drought version 1.0; deterministic SPEI12, generalized-logistic fit, 1991–2020 reference period; 2008–2022 subset [2, 7]	2010–2022 annual mean, drought-month share, and cumulative severity after fill-value and spatial checks	Constructed and validated
Agricultural outcomes	Archived NBS provincial annual-data responses and request metadata; audited factor table [5, 8]	Real agricultural output, six input-carbon activities, AICP, and IWP	Constructed at observed coverage
Merged hydro-agricultural panel	Two validated intermediate files and the deterministic merge script	403 unique province–year records, observed completeness flags, source hashes, and rule versions	Constructed and validated
Carbon-market feasibility	Official national-market records and a contextual pilot-market review	Assessment of market operation, province–market–year assignment, coverage, and identifying variation	Audited and excluded
Hydroclimatic analysis outputs	Validated 403-row panel and archived analysis script	Linear, quadratic, and continuous context-interaction estimates with support records, clustered uncertainty, multiplicity adjustment, and prespecified sensitivity checks	Generated and hash-registered

Table S3: Validation summary for the reconstructed hydrological panel

Check	Result	Interpretation
Province-year completeness	403 of 403 rows complete	No hydrological value was statistically imputed.
Annual volume conservation	99.26–99.33%	Share of the raw HSWUD domain total assigned to the 31 retained mainland polygons.
Polygon-area reconstruction	Maximum absolute error 0.0052%	Equal-area polygon-grid weights reproduce the source boundary areas.
Cross-province overlap	0 cells above 1.001 combined fraction	No material double allocation across adjacent provinces.
Minimum spatial support	IWU 92.0%; climate 100.0%; SPEI12 99.4%	HSWUD volumes are summed over nonmissing product cells without rescaling; climate means require at least 95% coverage.
Official plausibility comparison	2010: 85.3%; 2020: 95.1%; 2022: 95.1%	HSWUD irrigation withdrawal as a percentage of broader official agricultural water use; the concepts are not identical.
Large annual IWU changes	4 province-year flags, all in 2014	Values exceeding 50% are retained and flagged as a common product discontinuity, not smoothed.
Physical-range checks	0 violations	Nonnegative IWU and precipitation, positive annual PET, and drought shares in $[0, 1]$.

Table S4: Agricultural outcome inputs, carbon factors, and observed coverage

Input or outcome	NBS source field	Conversion or factor	Observed n	Analytical status or limitation
Real agricultural output	Agricultural gross output value; agricultural gross-output index (previous year = 100)	Province-specific chain link; 2010 prices	403	No missing numerator; 2010 nominal value anchors each provincial real-output series.
Fertilizer	Nutrient-equivalent fertilizer use (10^4 t)	$0.89 \text{ kg C kg}^{-1}$	403	Material mass converted to kg before applying the factor.
Pesticide	Pesticide use (10^4 t)	$4.93 \text{ kg C kg}^{-1}$	372	Missing from the official source for all provinces in 2022.
Agricultural film	Agricultural plastic-film use (t)	$5.18 \text{ kg C kg}^{-1}$	372	Missing from the official source for all provinces in 2022.
Agricultural diesel	Agricultural diesel use (10^4 t)	$0.59 \text{ kg C kg}^{-1}$	310	Missing from the official source for all provinces in 2020–2022.
Ploughing activity	Total crop sown area (10^3 ha)	$3.126 \text{ kg C ha}^{-1}$	403	The factor is paired with the source activity; the published $312.60 \text{ kg C km}^{-2}$ coefficient is converted to hectares.
Irrigation activity	Effective irrigated area (10^3 ha)	$266.48 \text{ kg C ha}^{-1}$	403	Area-based input-carbon component, distinct from water volume.
AICP	Real output/input-related carbon emissions	Constant-2010 CNY t C^{-1}	310	Complete for 2010–2019; no factor or activity imputation.
IWP	Real output/reconstructed irrigation-water use	Constant-2010 m^{-3} CNY	403	Complete for 2010–2022; retains the four 2014 water-product flags.

Table S5: Feasibility assessment of candidate province-level carbon-market exposure measures

Feasibility criterion	Audited evidence	Status
Annual pilot-market price, volume, and amount	Contextual review only; annual underlying records not archived	FAIL
Province-market-year assignment	Pilot designation is known, but Shenzhen-Guangdong overlap requires sub-provincial allocation	FAIL
National-market operation, 2021-2022	Available only as a market-wide series common to all provinces	PASS AT MARKET LEVEL
Province-specific regulatory coverage	No complete 2021-2022 province-year share in the archived sources	FAIL
Identifying within-year exposure variation	National price, volume, and amount are absorbed by year fixed effects	FAIL
Analytical CET exposure	CET contract contains 0 source-verified rows	EXCLUDE

Note: A failed feasibility criterion is not repaired by interpolation, duplication of market-wide values, or pilot-status substitution.

Supplementary association and sensitivity results

Table S6: Formal 2014 discontinuity and association-stability analysis for the irrigation denominator

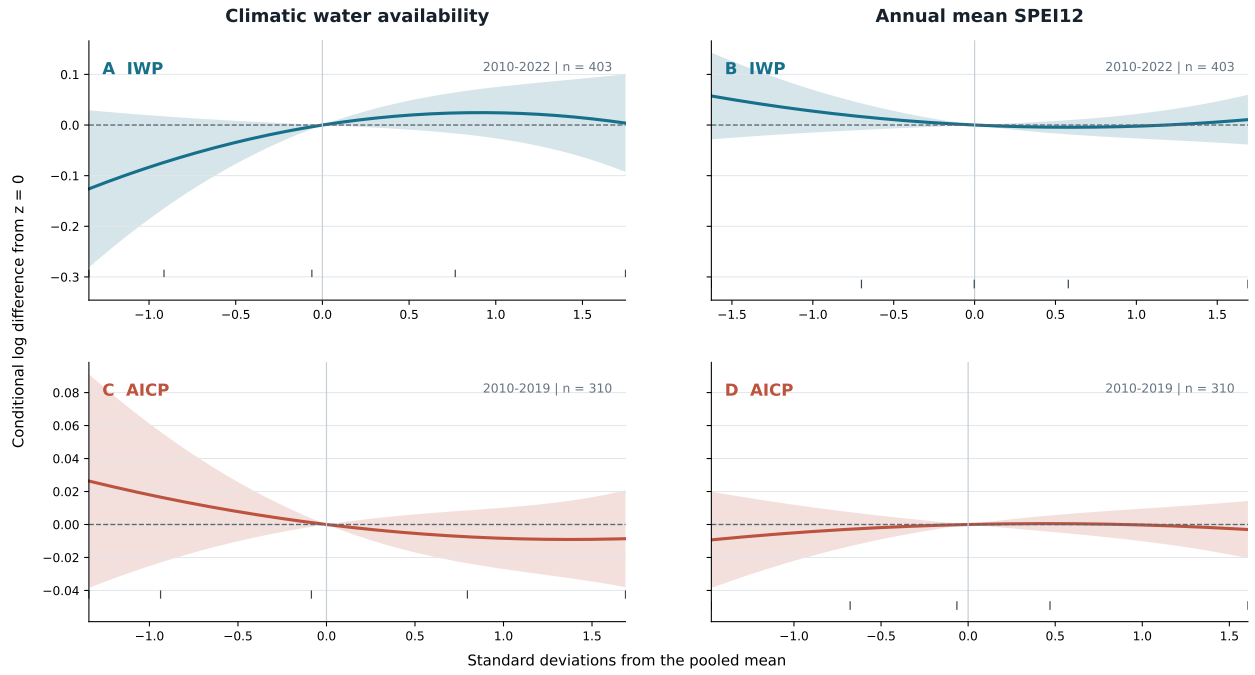
Diagnostic term	Estimate	Clustered SE	95% CI	<i>p</i>
IWU level shift at 2014	0.119	0.037	[0.043, 0.195]	0.003
IWU annual-slope change after 2014	-0.014	0.007	[-0.028, 0.001]	0.072
<i>AI</i> × post-2014 in the IWP model	-0.029	0.045	[-0.120, 0.062]	0.517
<i>SPEI12</i> × post-2014 in the IWP model	0.020	0.033	[-0.048, 0.088]	0.545

Notes: The product-break model regresses log reconstructed irrigation withdrawals on a linear trend, a 2014 level indicator, a post-2014 slope change, and province fixed effects. The association-stability model uses log IWP, both standardized hydroclimatic exposures, their post-2014 interactions, province and year fixed effects, and province-clustered inference. The joint interaction test is $F = 0.269$, $p = 0.766$. These tests diagnose product-version sensitivity and do not establish causality.

Table S7: Quadratic hydroclimatic association analysis over observed support

Outcome	Hydroclimatic measure	Linear term	Quadratic term	Curvature q
IWP	Climatic water availability	0.054 [-0.016, 0.124]	-0.030 [-0.071, 0.012]	0.382
	Annual mean SPEI12	-0.015 [-0.043, 0.014]	0.013 [-0.007, 0.032]	0.382
AICP	Climatic water availability	-0.013 [-0.042, 0.016]	0.005 [-0.011, 0.021]	0.548
	Annual mean SPEI12	0.002 [-0.010, 0.014]	-0.003 [-0.010, 0.004]	0.548

Notes: Entries are log-point coefficients with 95% province-clustered confidence intervals. Each model includes both standardized hydroclimatic measures, their squared terms, province fixed effects, and year fixed effects. Curvature q -values apply Benjamini–Hochberg adjustment to the four quadratic terms. The joint curvature tests are $p = 0.329$ for IWP and $p = 0.684$ for AICP; their adjusted q -values are 0.659 and 0.684. No breakpoint or optimum is estimated.



Panels show each outcome's 5th-95th percentile support; short marks indicate the 5th, 25th, 50th, 75th, and 95th percentiles.

Figure S1: Quadratic hydroclimatic associations over observed support. Curves show conditional log differences from the standardized exposure value $z = 0$, with the other hydroclimatic measure held at its standardized mean; shaded regions are 95% pointwise confidence intervals based on province-clustered covariance. Each panel is restricted to the outcome-specific 5th–95th percentile exposure range, and short marks indicate the 5th, 25th, 50th, 75th, and 95th percentiles. Models include linear and quadratic terms for both hydroclimatic measures and province and year fixed effects. The curves describe the specified quadratic form and do not identify thresholds or optimal exposure levels.

Table S8: Continuous dry-climate-context interaction analysis

Outcome	Interaction	Coefficient	Interaction q	N
IWP	$AI \times \text{dry-climate context}$	0.009 [-0.039, 0.058]	0.835	403
	$SPEI12 \times \text{dry-climate context}$	0.018 [-0.008, 0.043]	0.682	403
AICP	$AI \times \text{dry-climate context}$	-0.011 [-0.034, 0.012]	0.691	310
	$SPEI12 \times \text{dry-climate context}$	-0.001 [-0.012, 0.010]	0.835	310

Notes: Dry-climate context is the negative of each province's mean AI during the common 2010–2019 window, standardized across 31 provinces; higher values therefore denote a drier observed context. Coefficients measure the change in the within-province hydroclimatic slope per one-standard-deviation drier context. Models include both interactions, both main hydroclimatic terms, province and year fixed effects, and province-clustered inference. Interaction q -values adjust the four coefficients. The joint interaction tests are $p = 0.372$ for IWP and $p = 0.634$ for AICP; adjusted q -values are 0.634 and 0.634.

Table S9: Sensitivity of the quadratic and context analyses

Outcome	Specification	Curvature joint p	Curvature partial R^2	Context joint p	N
IWP	Primary extension	0.329	0.0164	0.372	403
	Common 2010–2019 support	0.290	0.0218	0.717	310
	Exclude four flagged 2014 rows	0.370	0.0149	0.426	399
	Exclude 2014	0.453	0.0137	0.444	372
	Province-specific linear trends	0.127	0.0124	0.947	403
	Drought-month-share alternative	0.357	0.0073	0.494	403
AICP	Primary extension	0.684	0.0026	0.634	310
	Common 2010–2019 support	0.684	0.0026	0.634	310
	Exclude four flagged 2014 rows	0.629	0.0035	0.638	306
	Exclude 2014	0.682	0.0031	0.610	279
	Province-specific linear trends	0.461	0.0064	0.339	310
	Drought-month-share alternative	0.903	0.0005	0.695	310

Notes: Joint tests assess the two added quadratic terms or the two dry-context interactions. Partial R^2 is the incremental in-sample explanatory share of the quadratic block relative to the corresponding linear fixed-effects model. Sensitivity tests are diagnostic and are not counted as independent confirmation. The AICP common-window row equals its primary 2010–2019 sample by construction.

Table S10: Hypothetical static carbon-price scenarios for the observed input-carbon inventory

Scenario Benchmark		Nominal rate	2010-price rate	Median burden		Aggregate burden
		CNY t ⁻¹ CO ₂		(Q1–Q3)		
				% of real agricultural output		
Low	2023 year-end close	79.42	60.14	0.526	(0.442–0.627)	0.511
Central	2024 value/volume average	95.84	72.43	0.634	(0.533–0.755)	0.615
High	2024 peak close	105.65	79.84	0.698	(0.587–0.833)	0.678

Notes: The accounting base is the observed six-source input-related carbon inventory (310 province-years, 2010–2019), converted from tonnes C to tonnes CO₂ by 44/12; it is not total agricultural greenhouse-gas emissions or an enacted tax base. Nominal benchmarks are national ETS prices observed in 2023 and 2024 and are deflated with the national CPI to constant-2010 CNY. Each scenario assumes full liability and no exemptions, thresholds, free allocation, pass-through, behavioral response, or equilibrium adjustment. The scenario span reflects rate choice only; it is not a confidence interval and does not quantify carbon-inventory uncertainty. Because the burden is algebraically proportional to 1/*AICP*, it is not used as a separate regression outcome.

References

- [1] Copernicus Climate Change Service. ERA5 monthly averaged data on single levels from 1940 to present, 2019. URL <https://doi.org/10.24381/cds.f17050d7>. Copernicus Climate Data Store dataset; accessed 26 August 2026; CC BY.
- [2] European Centre for Medium-Range Weather Forecasts. Monthly drought indices from 1940 to present derived from ERA5 reanalysis, 2025. URL <https://doi.org/10.24381/9bea5e16>. ERA5–Drought version 1.0; accessed 26 August 2026; CC BY.
- [3] geoBoundaries. China administrative level 1 boundaries, 2023. URL <https://www.geoboundaries.org/api/current/gbOpen/CHN/ADM1/>. Boundary ID CHN-ADM1-43563684; representative year 2019; build 12 December 2023; accessed 26 August 2026; public domain.
- [4] Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Dinand Schepers, Adrian Simmons, Cornel Soci, et al. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049, 2020. doi: 10.1002/qj.3803.
- [5] National Bureau of Statistics of China. Provincial annual data: Agriculture, 2010–2022, 2026. URL <https://data.stats.gov.cn/dg/website/page.html#/pc/national/fsYearData>. National Data portal; source indicator identifiers, units, annotations, request payloads, and responses archived; accessed 27 August 2026.
- [6] Daniel Runfola, Austin Anderson, Heather Baier, Matt Crittenden, Elizabeth Dowker, Sydney Fuhrig, Seth Goodman, Grace Grimsley, Rachel Layko, Graham Melville, Maddy Mulder, Rachel Oberman, Joshua Panganiban, Andrew Peck, Leigh Seitz, Sylvia Shea, Hannah Slevin, Rebecca Youngerman, and Lauren Hobbs. geoBoundaries: A global database of political administrative boundaries. *PLOS ONE*, 15(4):e0231866, 2020. doi: 10.1371/journal.pone.0231866.
- [7] Sergio M Vicente-Serrano, Santiago Beguería, and Juan I López-Moreno. A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *Journal of Climate*, 23(7):1696–1718, 2010. doi: 10.1175/2009JCLI2909.1.
- [8] Yuan Xue, Chunhua Li, Jingwen Li, Hui Lyu, Qingyun Lai, Zhilin Kang, Peng Yao, and Jiahui Li. Analysis of spatial and temporal characteristics and drivers of agricultural carbon emissions in china. *Chinese Journal of Eco-Agriculture*, 32(11):1805–1817, 2024. doi: 10.12357/cjea.20240038.
- [9] Yuqian Zhang and Yunhe Yin. HUSWD datasets (.nc), 2024. URL <https://doi.org/10.6084/m9.figshare.27610524.v7>. Figshare dataset, version 7, published 11 July 2026; CC BY 4.0.
- [10] Yuqian Zhang, Yunhe Yin, Mijia Yin, and Xufang Zhang. A high-resolution gridded dataset for china’s monthly sectoral water use. *Scientific Data*, 12:1157, 2025. doi: 10.1038/s41597-025-05400-2.