Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Drought does not raise water stress in the smallest countries and small island states

Muhammad Afnan Arif

muhdafnanarif@gmail.com

Asia-Europe Institute, Universiti Malaya

Research Article

Keywords: water stress, drought index, nonparametric panel, land area, small island states

Posted Date: September 15th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10671179/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Drought does not raise water stress in the smallest countries and small island states

Muhammad Afnan Arif [1]

Abstract

Sustainable Development Goal (SDG) indicator 6.4.2 reports water stress for almost every country yet compares countries sitting at very different points of its range. This study asks whether a standard drought index moves the indicator the same way in all of them. Two nested country panels are used, and estimation proceeds nonparametrically, with parametric fixed effects kept only as a reference point. The relationship is found to depend on the scale of a country. In larger countries it is negative and significant, consistent with drought raising water stress. In the smallest countries, and particularly small island states, it is weak, inconsistently signed, or absent. Land area and total population cannot be separated as the operative dimension. The moderation survives an attempt to explain it away as a measurement artefact, since rebuilding the index with area weighted sampling for smaller countries leaves the pattern unchanged. The mechanism concerns the physical scale at which a national average climate index corresponds to the hydrological system that supplies a country. That scale exists where large river basins integrate rainfall across a wide area, but not for small island economies reliant on groundwater, rainwater harvesting, or desalination. A single global model of the relationship is therefore not supported, and country scale is a testable dimension along which to condition future work on this indicator.

Keywords: water stress; drought index; nonparametric panel; land area; small island states

JEL codes: Q25, Q54, C14, C23

1. Introduction

SDG 6.4.2 reports the level of water stress for almost every country in the world, expressed as total freshwater withdrawal against available renewable resources. It carries substantial policy weight, anchoring national reporting under Agenda 2030 and increasingly appearing as a covariate in applied climate research. The indicator is also distributed very unevenly across the countries that report it, as Fig. 1 shows: a small number withdraw more than their renewable resources can sustain, while most sit far below that threshold. The map reports the most recent observation for each country between 2018 and 2022. Any single global relationship between climate and this indicator is therefore fitted across countries at very different points of its range, which raises a natural question: does drought move the indicator the same way in every country that reports it.

Prior work shows the indicator is sensitive both to how it is constructed and to what it omits. Hoekstra et al. (2012) find that computing scarcity monthly rather than annually reveals far more countries under severe stress. Rosa et al. (2020) separate scarcity driven by inadequate infrastructure from scarcity driven by absolute physical shortage. Briguglio (1995) argues that conventional aggregate measures understate the exposure of small island states to external and environmental shocks. Burke et al. (2015) document that climate variables enter economic relationships non-linearly rather than proportionally. What none of this work tests is whether the physical size of a country changes what a national climate average measures in the first place, and that is the question taken up here.

Drought does not affect water stress the same way everywhere, and the reason is not that some countries are simply more sensitive to climate variability than others. It is that a precipitation-based drought index computed as a national average carries different physical meaning depending on the size of the country it describes. In large countries with extensive river basins and irrigation networks, a national average captures broad scale hydrological

[1] Corresponding author. Asia-Europe Institute, Universiti Malaya, Kuala Lumpur, Malaysia. Email: muhdafnanarif@gmail.com

conditions bearing on national withdrawal, a point Fig. 2 makes visible at grid resolution. The gridded product shown there is compiled independently of the index used in estimation. In small countries, and particularly small island states, no such system exists, and supply depends instead on groundwater lenses, rainwater harvesting, or desalination, none of which respond to rainfall the way a river basin does.

This study estimates the relationship between drought and SDG 6.4.2 across a panel of 165 countries, using methods that leave the functional form unrestricted rather than imposing one. The main finding is that the drought effect is absent in the smallest countries, and particularly in small island states, while it is negative and significant elsewhere. Land area and total population describe the same scale and cannot be separated in this panel. It survives an explicit attempt to rule it out as a measurement artefact, being estimated on two independently constructed versions of the drought index and materially unchanged between them.

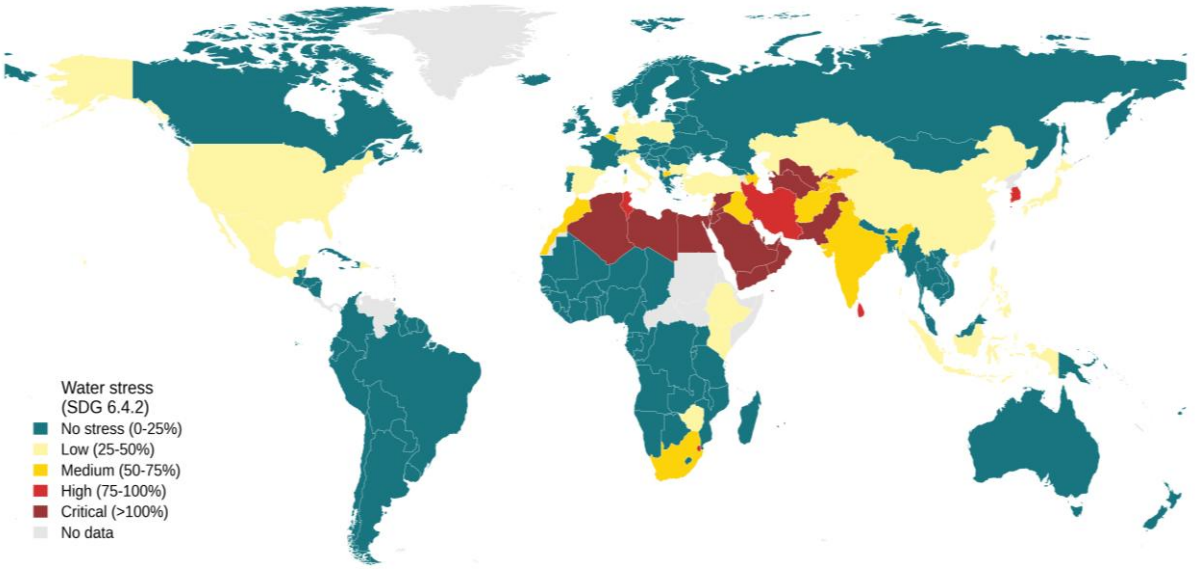


Fig. 1 Water stress by country

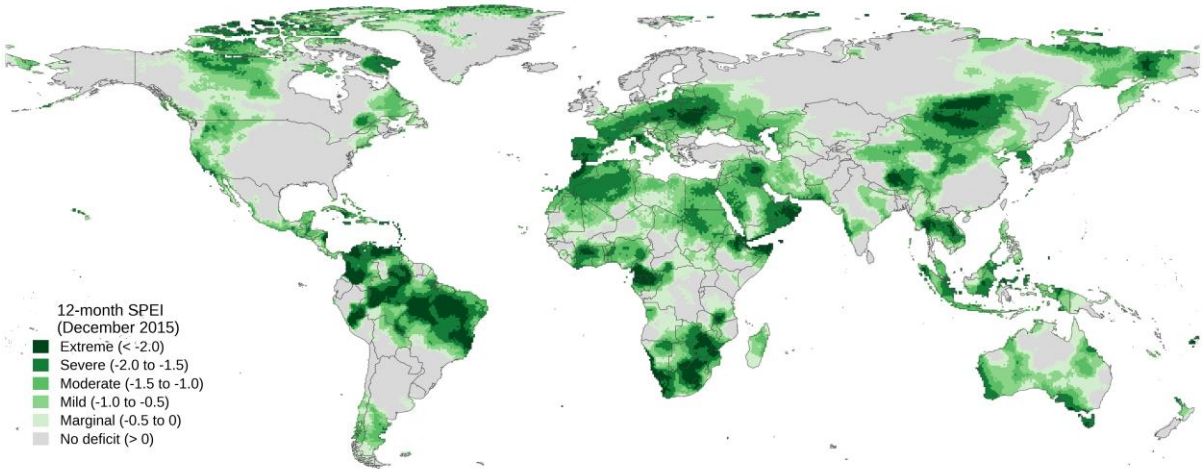


Fig. 2 Gridded drought index, December 2015

Two further contributions follow. Much of the empirical literature on water and environmental panels reports distributional heterogeneity using the location and scale quantile estimator of Machado and Santos Silva (2019). Diagnostic work carried out here found that its scale

equation produces negative variance predictions once countries and years are demeaned, and it is set aside for that reason. The result also bears on how the indicator is used, since comparing SDG 6.4.2 across countries of very different physical scale treats a climate signal as equally informative everywhere when it is not. The main limitation is that the available series are short, so the analysis cannot fully separate a measurement artefact from attrition of identifying variation. The study reviews the relevant literature, presents the data and empirical framework, reports results, and concludes with implications.

2. Literature review

2.1 Measurement and determinants of water stress

SDG 6.4.2 is one of several indicators used to summarise water scarcity at country level, and the choice of aggregation matters for what the indicator can and cannot show. Kummu et al. (2016) document the rise in the share of the world population living under water stress over the twentieth century. Hoekstra et al. (2012) show that such trends are sensitive to temporal resolution, since blue water footprint assessments computed at monthly resolution identify substantially more country months under severe stress than the same calculation performed on annual totals. Liu et al. (2017) review the broader family of water scarcity indicators and note persistent disagreement across methods regarding both the level and the geographic distribution of stress. Rosa et al. (2020) distinguish economic from physical water scarcity, since a substantial share of irrigated land in low income economies faces scarcity driven by inadequate infrastructure rather than by an absolute shortage of the resource itself. That distinction is directly relevant to why a single global relationship between climate and a withdrawal based indicator should not be assumed.

Trade and basin structure introduce further complications. Dalin et al. (2017) find that a substantial and growing share of the groundwater consumed in irrigation is embedded in food that crosses a national border, so that withdrawal recorded against one country partly reflects demand realised elsewhere. Veldkamp et al. (2017) show that interventions addressing scarcity in one part of a river basin frequently relocate rather than resolve it. Wisser et al. (2008) and Deryng et al. (2014) both find that a quantity aggregated to the country level, whether irrigation demand or crop yield response to heat, behaves differently from the same quantity measured locally. That difference is a further source of the country level heterogeneity motivating the nonparametric approach taken here. Hoekstra and Hung (2005) and Mekonnen and Hoekstra (2010) develop the virtual water and water footprint accounting framework underlying much of this trade related evidence. Sachs et al. (2019, 2021) illustrate how widely a single indicator architecture is now used to compare countries of very different physical scale, without addressing that comparability question directly.

The hydrological literature on groundwater depletion provides a further, largely separate body of evidence. Alley et al. (2002) establish that persistent groundwater abstraction in excess of recharge is a general rather than a localised condition. Taylor et al. (2013) show that recharge sensitivity to climate depends on the timing rather than the total of precipitation, a channel not captured by an annual drought index. Perrone and Jasechko (2019) find that the common response to a falling water table is to drill wells deeper rather than to reduce abstraction, a further reason a purely climate-based indicator may understate the severity of stress in some settings.

2.2 Climate variability and drought

The relationship between climate variability and water availability has been examined extensively at the global and regional scale, generally using indices constructed from gridded climate products rather than the point-based construction used here. Dai (2013) reports a rising share of global land area under drought over recent decades using a physically based measure of evaporative demand, while the trend implied by simpler formulations of the same concept is comparatively muted. That sensitivity to construction choice motivates the reconstruction exercise reported below. Cook et al. (2018) note that consensus on drought

trend and attribution weakens from the global to the regional scale, and Naumann et al. (2018) project that warming lengthens drought in several regions while changing little elsewhere. Pokhrel et al. (2021) project that exposure to extreme terrestrial water storage drought more than doubles under continued warming, indicating that a storage based view of drought can diverge from one built on precipitation alone. Vicente-Serrano et al. (2011) document that drought responds to large scale circulation with a lag varying by region and accumulation period. Both findings are consistent with a climate signal whose national relevance depends on the physical scale of the system it describes.

Index construction itself carries evidence on this point. Sippel et al. (2020) show that a climate change signal is detectable from short windows of daily weather data at global scale, motivating the high frequency, directly constructed index used here rather than a pre-existing gridded product. A separate but closely related applied economics literature uses panel variation in weather realisations to identify economic effects of climate. Dell et al. (2014) note that this identification strategy relies on within unit variation over time, shared by the two way fixed effects baseline used here. Burke et al. (2015) document a strongly non-linear relationship between temperature and output, peaking at a moderate temperature and declining sharply on either side, motivating the flexible specifications pursued nonparametrically in this study. Auffhammer (2018) reviews the sensitivity of estimated climate damages to functional form, directly relevant to the decision to estimate the drought and water stress relationship nonparametrically rather than through a single global linear or quadratic specification.

2.3 Institutional and economic mediators of water stress

A parallel literature examines how economic and institutional conditions mediate the translation of physical water availability into measured stress. Grafton et al. (2018) show that subsidised irrigation efficiency improvements frequently raise consumptive use at basin scale, so institutional response to scarcity is neither automatic nor uniform. Garrick et al. (2019) find that reallocation of water from agricultural to urban use is concentrated in a small number of countries and often contested where compensation arrangements are weak.

Microeconometric evidence on groundwater access reinforces the same theme. Sekhri (2014) documents measurable consequences for rural poverty where the water table falls below shallow wells. Hornbeck and Keskin (2014) and Blakeslee et al. (2020) find that adaptation to aquifer depletion in the United States and in India alike takes the form of switching crops rather than reducing cultivation outright. Ryan and Sudarshan (2022) find that rationing pump electricity changes groundwater withdrawal substantially, evidence that the marginal cost of abstraction, not only the resource endowment, governs observed use. Institutional and economic context therefore shapes how a physical constraint appears in withdrawal based indicators.

2.4 Country size and the economics of small states

The idea that national scale is itself an economically meaningful characteristic has a long history predating the water security literature above. Alesina and Spolaore (1997) model nation size as a trade off between the benefits of scale and the costs of governing a heterogeneous population. Alesina (2003) finds that smaller countries can perform well economically if sufficiently open to trade, a qualification that does not obviously extend to physical resource systems such as water the way it does to markets for goods. Alesina and Spolaore (2005) show that the benefits of scale are partly substitutable through economic integration, which has no clear analogue for a river basin since integrating with a neighbour does not import its watershed. Briguglio (1995) argues that conventional income measures understate the exposure of small island states to external and environmental shocks. That is consistent with the proposition tested here, that the relationship of a small state to a climate indicator cannot be assumed to mirror that of a large one.

2.5 Heterogeneous and nonparametric panel methods

Panel econometrics has moved steadily away from the assumption that a single slope coefficient describes every cross sectional unit. De Hoyos and Sarafidis (2006) develop a formal test for that assumption in the closely related context of cross sectional dependence, generalised by Pesaran (2021) to an arbitrary and potentially unknown pattern of dependence across units. Those tests, together with the poolability test applied below, provide the diagnostic basis for treating slope heterogeneity as a property of the data to be modelled rather than assumed away. Distributional heterogeneity has often been estimated through the location and scale quantile regression of Machado and Santos Silva (2019). The scale equation of that method was found in the course of preparing this study to be misspecified once applied to a two way within transformed panel of the kind used here, the finding discussed in the introduction.

The nonparametric and semiparametric estimators used below draw on kernel foundations set out by Härdle (1990) and Pagan and Ullah (1999), extended to mixed continuous and categorical panel data by Li and Racine (2007). They also draw on the differencing based alternative of Yatchew (2003), retained here only where kernel smoothing is not required. The linear quantile benchmark against which the nonparametric quantile estimator used here is compared originates with Koenker and Bassett (1978) and its panel extension by Koenker (2004). Standard inference throughout follows White (1980) and Breusch and Pagan (1980) for heteroskedasticity and the wild cluster bootstrap of Cameron et al. (2008) for the country clustered setting. The small sample bias in dynamic panel estimation documented by Judson and Owen (1999) motivates the decision, discussed in the introduction, not to include a lagged dependent variable. None of these estimators has previously been applied to SDG 6.4.2. The specific hypothesis that their flexibility reveals, namely that a climate index moderates its own relevance according to the physical scale of the country it describes, has not to the knowledge of the present author been tested directly. Table 1 summarises the empirical studies most directly relevant to this measurement and mechanism; the estimators themselves are set out formally in the methods.

Table 1. The empirical studies related to the measurement, mechanism, and drought sensitivity of SDG6.4.2.

Study	Focus	Sample	Key finding
Kummu et al. (2016)	Water stress measurement	Global, twentieth century	Population under water stress rose sharply across the twentieth century
Hoekstra et al. (2012)	Water stress measurement	Global blue water footprint	Monthly resolution identifies far more country months under severe stress than annual
Rosa et al. (2020)	Water stress measurement	Global irrigated land	Much scarcity in low income economies is economic rather than physical
Dalin et al. (2017)	Trade and aggregation	Global groundwater and trade	A growing share of irrigation groundwater is embedded in traded food
Veldkamp et al. (2017)	Basin structure	Global river basins	Interventions often relocate scarcity within a basin rather than resolving it
Dai (2013)	Drought measurement	Global land, recent decades	Share of land under drought rises, and the trend depends on index construction
Naumann et al. (2018)	Drought projection	Global warming scenarios	Warming lengthens drought in some regions and changes little elsewhere
Pokhrel et al. (2021)	Drought detection	Terrestrial water storage	Exposure to extreme storage drought is projected to more than double
Vicente-Serrano et al. (2011)	Drought dynamics	Global circulation modes	Drought responds to circulation with lags varying by region and accumulation period
Burke et al. (2015)	Climate and output	Country panel	Output responds non-linearly to temperature, peaking at moderate values
Grafton et al. (2018)	Water policy	Basin scale irrigation efficiency	Efficiency subsidies can raise consumptive use by cutting return flow
Garrick et al. (2019)	Water reallocation	Cross country	Agricultural to urban reallocation is concentrated and often contested

Sekhri (2014)	Groundwater access	Rural India	Rural poverty rises where the water table falls below shallow wells
Ryan and Sudarshan (2022)	Abstraction cost	Indian electricity rationing	Rationing pump electricity substantially changes groundwater withdrawal
Briguglio (1995)	Small states	Small island economies	Income measures understate exposure to external and environmental shocks

Notes: studies are ranked by relevance to the measurement of water stress, the construction of drought indices, and the small state mechanism central to this study, every row is cited in the preceding prose, method and diagnostic test studys are cited in the text but are not summarised here since Table 1 is restricted to empirical findings.

2.6 Theoretical framework

SDG 6.4.2 measures water stress against the true hydrological state of a country, which is unobserved; what is observed is a national average drought index. The gap between the two widens as land area falls, since a climate average maps onto actual supply only where a river basin is large enough to integrate rainfall across a wide area. Small states relying on groundwater, rainwater harvesting, or desalination do not meet that condition. Estimating on the observed index therefore recovers an attenuated version of the true sensitivity, attenuated in proportion to the share of the variance of the index attributable to this gap rather than genuine hydrological signal, a share decreasing in land area. This is the formal basis, derived in Appendix B.0, for allowing the drought coefficient to vary smoothly with log land area: near zero or unpredictably signed where the gap dominates, moving toward the true, correctly signed sensitivity as land area grows and the gap closes.

Brazil and Argentina, organised around the Amazon-Paraná and Paraná-Uruguay basins, exemplify the case where climate average and hydrological outcome are mechanically linked. Small island states are the opposite case. Parker et al. (2023), Cashman (2013), Gungoa and Kebede (2024), Gheuens et al. (2019), and Elimelech and Phillip (2011) each find water security there depends on distribution infrastructure, groundwater quality, and desalination rather than basin response to a rainfall anomaly. Each control enters for a mechanism-specific reason. Income proxies capacity to invest in infrastructure converting a hydrological endowment into usable supply, population density and urban share proxy withdrawal scale, and agricultural share proxies demand composition since irrigation is typically its largest claim. None is expected to explain the land area moderation itself.

Three hypotheses follow, tested formally in the results. H1: the drought coefficient, as a smooth function of log land area, increases in land area, from near zero or incorrectly signed in the smallest countries to negative and significant in the largest. H2: this should survive residualising land area against income, since the gap reflects physical hydrological scale rather than economic development. H3: an area weighted index for smaller countries should not remove the moderation, since averaging within the borders of a small country does not create a river basin it lacks. The mechanism does not extend cleanly to countries sharing a basin with a neighbour, where land area no longer indexes the scale relevant to supply. Fig. 3 traces that path from land area through the representativeness gap to the estimated coefficient, and sets out the three hypotheses together.

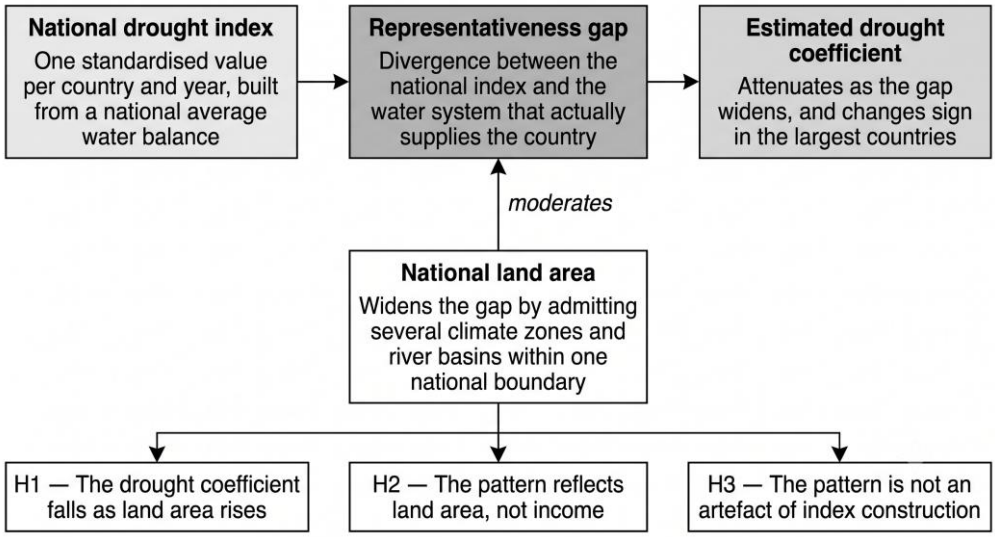


Fig. 3 Research framework

3. Data and methods

3.1 Data

The dependent variable is SDG 6.4.2, drawn from World Bank World Development Indicators, expressed as a percentage and log transformed. Two country panels are used. The first comprises 59 countries with continuous reporting sufficient for a minimum of 15 annual observations. The second extends this to 165 countries meeting the same completeness criterion, drawing covariates from the same World Bank series. Table A1 lists the countries in each panel. All 59 countries in the first panel are contained within the second, so the two samples are nested rather than independent, and the wider panel is best read as an extension to a more heterogeneous set of countries rather than an independent replication sample. Covariates are gross domestic product per capita, population density, and urban population share, all log transformed except the latter, and the share of agriculture in value added. Table 2 lists every variable.

The withdrawal data underlying the indicator are compiled at irregular intervals and the reported annual series are filled between compiled values, so the number of annual observations overstates the number of independent measurements. The extent of that filling is quantified with the results, together with its consequence for the estimates.

Drought is measured by the 12-month Standardised Precipitation Evapotranspiration Index, computed from daily precipitation and Penman-Monteith reference evapotranspiration obtained from the ERA5-Land reanalysis archive. Two constructions of this index are used throughout as a robustness pair. The first computes precipitation and reference evapotranspiration at a single geographic centroid per country. The second repeats this construction with an additional two interior sampling points for every country below the median land area in the sample, drawn from a corrected national bounding box, with the three points averaged at the daily water balance stage before standardisation. Standardisation follows the Gringorten plotting position transformed to the standard normal distribution, computed within calendar month.

Land area, expressed in square kilometres and log transformed, is drawn from World Bank series, fixed over the sample period, and used as the moderating variable central to the analysis. The 165 country panel spans a wide range of national scale, from small island states of a few hundred square kilometres to continental economies exceeding ten million. The composition of the two panels differs systematically along this dimension. Of the 106 countries present only in the wider panel, 58 have a median water stress value below 10%, against 21 of the 59 in the original panel. The additional countries are therefore disproportionately drawn

from settings where renewable water resources are abundant relative to withdrawal. They differ in outcome as well as in scale. Fig. 4 plots eight countries in each of six world regions, selected by length of continuous reporting, on a logarithmic scale, and neither the level nor the trend of water stress is uniform within a region, let alone across regions. This is the country level heterogeneity the estimation strategy is designed to accommodate rather than assume away. The matching drought series appear in Fig. 5. Because the index is standardised within country, those panels are informative about the timing and co-movement of drought episodes rather than about differences in drought level between countries, which standardisation removes by construction.

The nuisance specification is deliberately richer than the covariate list. Beyond the four covariates in levels, the conditioning set includes their first lags, their squares, and all pairwise products, giving 18 conditioning variables in total. This expansion is required because every estimator applied below places the conditioning set inside an unknown function rather than a linear index, and a conditioning set of four variables would not exercise that capability.

Table 2. List of data sources.

Variable	Acronym	Description	Source
Water stress	WS	Freshwater withdrawal as a share of total renewable freshwater resources	WDI
Income	GDP	Gross domestic product per capita	WDI
Population density	POPD	Population per square kilometre of land area	WDI
Urbanisation	URB	Share of total population living in urban areas	WDI
Agricultural share	AGVA	Agriculture value added as a share of total value added	WDI
Land area	LA	Total national land area, log transformed in estimation	WDI
Drought	DR	Twelve month standardised precipitation evapotranspiration index	ERA5-Land

Notes: WS, GDP, POPD, URB, AGVA, and LA are drawn from the World Bank World Development Indicators, DR is constructed by the author from the Copernicus Climate Change Service reanalysis, descriptive statistics show in Table A2.

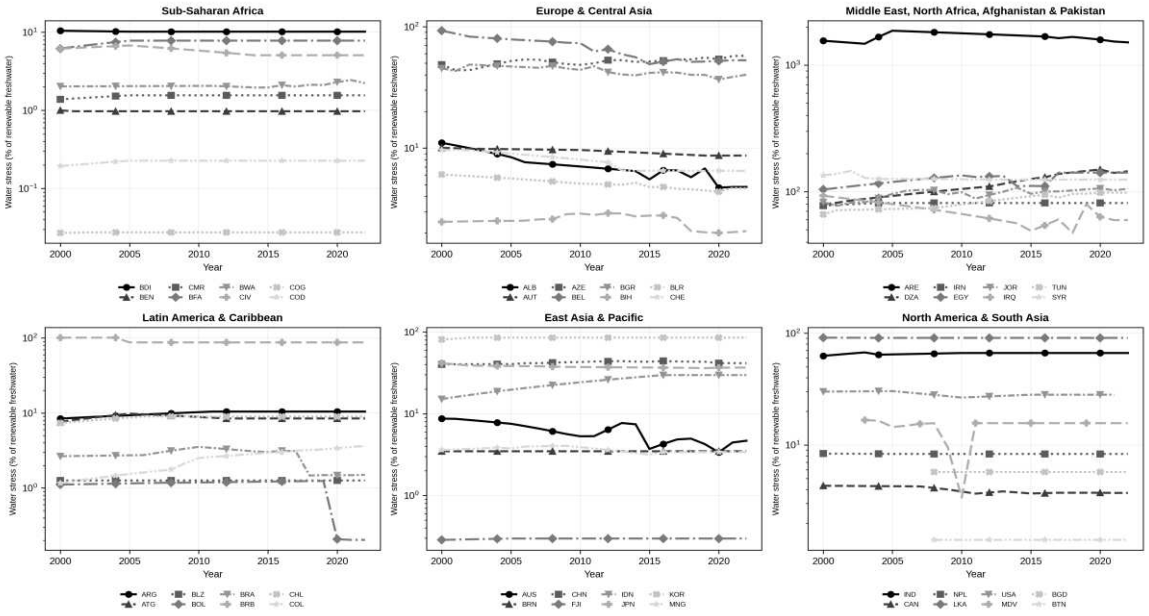


Fig. 4 Water stress by region

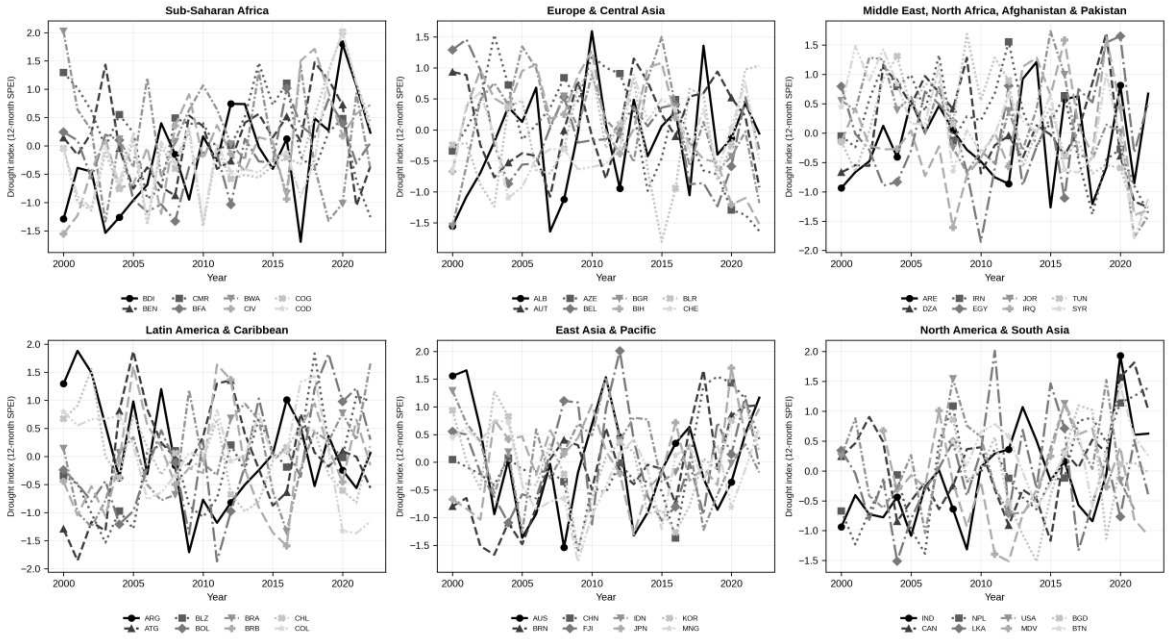


Fig. 5 Drought index by region

3.2 Estimation strategy

The estimation strategy is organised around four objectives. The first three ask whether drought affects water stress on average once the relationship between the covariates and both outcome and treatment is left unspecified. They then ask whether the drought coefficient varies with land area, and whether that variation is attributable to land area rather than income or common shocks. The fourth applies the double machine learning varying coefficient estimator developed in this study, denoted DML-VC, as a robustness check on the preceding three. All estimators share a common first stage. Country and year effects are removed by the two way within transformation, and the conditioning set is partialled out of both outcome and treatment by machine learning rather than a linear projection. The resulting residuals enter a Neyman orthogonal moment condition, so first order conditioning errors do not transmit to the parameter of interest. Cross fitting uses five folds formed by country, so the entire time series of a country falls inside a single fold and serial dependence cannot leak between estimation and evaluation.

Three estimators address the first objective. The partially linear panel specification with machine learning nuisance estimation is applied in both its within group form (DML-W) and its first difference form (DML-FD), following Clarke and Polselli (2026). Writing $\ddot{y}_{it}, \ddot{d}_{it}, \ddot{\mathbf{x}}_{it}$ for the two way within transformed outcome, treatment, and controls as in Appendix B, and Δ for the first difference operator, the two variants solve the analogous orthogonal moment on levels and on first differences:

$$\begin{aligned} \hat{\beta}_{\text{within}} &= \frac{\sum_i \sum_t \tilde{d}_{it}\tilde{y}_{it}}{\sum_i \sum_t \tilde{d}_{it}^2}, & \tilde{y}_{it} &= \ddot{y}_{it} - \hat{\ell}(\ddot{\mathbf{x}}_{it}), & \tilde{d}_{it} &= \ddot{d}_{it} - \hat{m}(\ddot{\mathbf{x}}_{it}), \\ \hat{\beta}_{\text{FD}} &= \frac{\sum_i \sum_t \Delta\tilde{d}_{it}\,\Delta\tilde{y}_{it}}{\sum_i \sum_t \Delta\tilde{d}_{it}^2}, & \Delta\tilde{y}_{it} &= \Delta y_{it} - \hat{\ell}_{\Delta}(\Delta\mathbf{x}_{it}), & \Delta\tilde{d}_{it} &= \Delta d_{it} - \hat{m}_{\Delta}(\Delta\mathbf{x}_{it}), \end{aligned} \tag{1}$$

where $\hat{\ell}, \hat{m}$ and their first differenced analogues $\hat{\ell}_{\Delta}, \hat{m}_{\Delta}$ are learned by cross fitted machine learning on folds formed by country, the two forms differing only in the assumption imposed on the distribution of the unobserved country effects. The average partial derivative of the outcome with respect to DR is estimated by automatic debiased machine learning (ADML), following the continuous treatment panel construction of Klosin and Vilgalys (2022) and the automatic Riesz representer of Chernozhukov et al. (2022):

$$\begin{aligned}\hat{\theta}_{\text{ADML}} &= \tfrac{1}{NT}\textstyle\sum_i \sum_t \left\{\partial_d \hat{\gamma}\left(\ddot{d}_{it}, \ddot{\mathbf{x}}_{it}\right) + \hat{\alpha}\left(\ddot{d}_{it}, \ddot{\mathbf{x}}_{it}\right)\left[\ddot{y}_{it} - \hat{\gamma}\left(\ddot{d}_{it}, \ddot{\mathbf{x}}_{it}\right)\right]\right\}, \\ \hat{\alpha} &= \underset{\alpha}{\operatorname{argmin}}\ \tfrac{1}{NT}\textstyle\sum_i \sum_t \left\{\alpha\left(\ddot{d}_{it}, \ddot{\mathbf{x}}_{it}\right)^2 - 2\ \partial_d \alpha\left(\ddot{d}_{it}, \ddot{\mathbf{x}}_{it}\right)\right\},\end{aligned} \tag{2}$$

where $\hat{\gamma}$ is a cross fitted machine learning regression of the outcome on drought and the controls, and $\hat{\alpha}$ is the automatic Riesz representer, obtained without specifying a conditional density by solving the quadratic objective above over a linear dictionary in d_{it} and $\mathbf{x}_{it}$. $\hat{\gamma}$ and $\hat{\alpha}$ are cross fitted on the same country blocked folds as the estimator above.

The second objective is addressed by four estimators. The two step estimator of Zimmert and Lechner (2026), denoted MLK, recovers the coefficient function by local constant kernel smoothing of the same country level orthogonal moments used in Appendix B.3 over the moderator:

$$\hat{\beta}_{\text{lc}}(c) = \frac{\sum_{i=1}^{N} K_h(c_i - c)\, B_i}{\sum_{i=1}^{N} K_h(c_i - c)\, A_i}, \tag{3}$$

where A_i, B_i are as defined in Appendix B.3. The DML-VC estimator of the fourth stage replaces this local constant step with the local linear refinement derived in Appendix B.4, the two forms differing only in the bandwidth rate and the boundary behaviour of the smoother. The generalized random forest of Athey et al. (2019), denoted GRF, and the local linear forest of Friedberg et al. (2021), denoted LLF, recover the same coefficient function without a bandwidth, replacing the fixed kernel with a data driven adaptive weight:

$$\begin{aligned}\hat{\theta}(c) &= \underset{\theta}{\operatorname{argmin}}\ \textstyle\sum_{i=1}^{N} \alpha_i(c)\, \tilde{d}_i\left(\tilde{y}_i - \theta\, \tilde{d}_i\right), \\ \alpha_i(c) &= \tfrac{1}{B}\textstyle\sum_{b=1}^{B} \frac{\mathbb{1}\{c_i \in L_b(c)\}}{|L_b(c)|},\end{aligned} \tag{4}$$

where $L_b(c)$ is the leaf of tree b containing the evaluation point c and $\tilde{d}_i, \tilde{y}_i$ are the country level orthogonalised residuals. LLF augments the same weights with a local linear correction in c_i rather than a local constant fit, which matters when the coefficient function is smooth.

The causal forest with fixed effects of Kattenberg et al. (2023), denoted CFFE, grows its tree using the same splitting rule as the two forests above, so no additional partialling of country and year effects occurs while the tree is grown. The distinguishing feature instead lies in the final coefficient step, where the grown forest supplies similarity weights $\alpha_i(c)$ exactly as in the forest weighted estimator defined above, and these weights recentre the within transformed outcome and treatment locally at each evaluation point before the coefficient is computed:

$$\begin{aligned}\hat{\tau}_{\text{CFFE}}(c) &= \frac{\sum_{i=1}^{N} \alpha_i(c)\, d_i^{\dagger}(c)\, y_i^{\dagger}(c)}{\sum_{i=1}^{N} \alpha_i(c)\, d_i^{\dagger}(c)^2}, \\ d_i^{\dagger}(c) = \ddot{d}_i - \textstyle\sum_{j=1}^{N} \alpha_j(c)\, \ddot{d}_j, &\quad y_i^{\dagger}(c) = \ddot{y}_i - \textstyle\sum_{j=1}^{N} \alpha_j(c)\, \ddot{y}_j,\end{aligned} \tag{5}$$

Here $\ddot{y}_i$ and $\ddot{d}_i$ denote the two way within transformed outcome and treatment, and the local recentring by $\alpha_j(c)$ substitutes for the machine learning stage used elsewhere, so country and year effects are removed jointly with the moderator rather than in a separate first stage.

Three further estimators target the third objective. The balanced moderation estimator of Bearth and Lechner (2025), denoted BGATE, reweights each country by the ratio of the marginal to the conditional density of GDP given its LA group before the coefficient function is re-estimated:

$$\begin{aligned}w_i &= \frac{\hat{p}(\text{income}_i)}{\hat{p}\left(\text{income}_i \mid g(c_i)\right)}, \\ \hat{\beta}_{\text{bal}}(c) &= \frac{\sum_{i=1}^{N} w_i\, K_h(c_i - c)\, B_i}{\sum_{i=1}^{N} w_i\, K_h(c_i - c)\, A_i},\end{aligned} \tag{6}$$

where $g(c_i)$ denotes the LA quartile of country i, so that an LA pattern surviving this reweighting is not attributable to the correlation between LA and GDP. An orthogonal spline

conditional average treatment effect (OSC) is estimated over a quadratic spline basis in LA with the basis terms selected by penalised regression, following Semenova et al. (2023):

$$\begin{aligned} \tau(c) &= \mathbf{b}(c)'\boldsymbol{\theta}, \qquad \mathbf{b}(c) = (1,\ c,\ c^2,\ (c-\kappa_1)_+^2,\ (c-\kappa_2)_+^2,\ (c-\kappa_3)_+^2)', \\ \widehat{\boldsymbol{\theta}} &= \underset{\boldsymbol{\theta}}{\operatorname{argmin}}\ \textstyle\sum_{i=1}^{N}(B_i - \mathbf{b}(c_i)'\boldsymbol{\theta}\, A_i)^2 + \lambda \parallel \boldsymbol{\theta} \parallel_1, \end{aligned} \tag{7}$$

where $\kappa_1, \kappa_2, \kappa_3$ are the sample quartiles of c_i and λ is chosen by cross validation, terms with $\hat{\theta}_j = 0$ dropped before the contrast at the extremes of the LA distribution is computed. Common correlated effects estimation with heterogeneous loadings addresses cross sectional dependence, estimating for each country separately a regression of the outcome on DR, the controls, and the raw cross sectional averages of each:

$$\begin{aligned} y_{it} - \bar{y}_i &= \beta_i\big(d_{it} - \bar{d}_i\big) + \mathbf{x}_{it}'\boldsymbol{\gamma}_i + \boldsymbol{\delta}_i'\,\bar{\mathbf{z}}_t + u_{it}, \\ \hat{\beta}_{\mathrm{MG}} &= \tfrac{1}{N}\textstyle\sum_{i=1}^{N}\hat{\beta}_i\,, \end{aligned} \tag{8}$$

where $\bar{y}_i, \bar{d}_i$ denote country specific time averages, $\bar{\mathbf{z}}_t$ collects the raw cross sectional averages of the outcome, the treatment, and the controls in year t, and $\hat{\beta}_{\mathrm{MG}}$ is the mean group estimator of Pesaran (2006), denoted CCEMG. The defactored variant, denoted CCE-DML, instead projects $\bar{\mathbf{z}}_t$ out of $y_{it}, d_{it}, \mathbf{x}_{it}$ country by country before the DML-W first stage is applied to the residuals.

A pooled common correlated effects correction is not used. Cross sectional averages are constant within a year and are absorbed exactly by the year effects once a two way within transformation has been applied, so adding them leaves the estimate unchanged, a point verified directly in the results below.

The fourth objective applies DML-VC. The conditioning functions are estimated by cross fitted machine learning as above, and the resulting orthogonal moments are aggregated to the country level. The coefficient function is then recovered by local linear smoothing over LA, with a bandwidth selected by leave one country out cross validation. Because the moderator is time invariant, the second stage has as many distinct design points as there are countries rather than country year observations, and both the bandwidth and the variance are constructed at that rate. The derivation, conditions, and asymptotic distribution appear in Appendix B.

Inference throughout is by country level block bootstrap for the coefficient function and by country clustered sandwich estimation for the scalar parameters. The clustered form is used because the alternative that treats country year observations as independent omits the within country cross products of the orthogonalised scores, and therefore understates the standard error whenever those scores are serially correlated. Cross sectional dependence is tested following Pesaran (2021), and where dependence is found the specification is re-estimated with heterogeneous loadings on the cross sectional averages, with the sensitivity of the results to the richness of that projection reported.

4. Results and discussion

4.1 Results

The average effect and its heterogeneity

Table 3 reports the preliminary diagnostics. Table 4 reports every estimator on the full panel. Read across the first row, the average effect of drought on water stress is not distinguishable from zero under any estimator that targets it. The three partially linear and average derivative estimators agree closely with one another and with the forest average, and none approaches conventional significance. Read down the remaining rows, the picture reverses. Every estimator that permits the coefficient to vary with land area finds that it does. The coefficient is positive and significant for the smallest countries, indistinguishable from zero at the median, and negative and significant for the largest. The contrast between the extremes of the LA distribution is significant at the 1% level under both MLK and DML-VC. These two findings are

the same finding. A coefficient that is positive for small countries and negative for large countries averages to approximately zero across the panel, so an insignificant average effect is what the theoretical framework predicts rather than evidence against it. The average effect is therefore reported as supporting evidence for the heterogeneity rather than as an independent result. Balancing the GDP distribution across LA groups under BGATE leaves the gradient essentially unchanged, so the LA pattern is not a restatement of a relationship between drought sensitivity and economic development; this is the direct test of the second hypothesis, and it is passed.

Table 3. Preliminary diagnostics.

Test	Panel A	Panel B
Poolability, common slopes across countries	Rejected	Rejected
Cross sectional dependence, within residuals	5.175***	8.124***
Cross sectional dependence, machine learning residuals	-	2.848***
Cross sectional dependence, common loading correction	-	6.599***
Countries	59	165
Country year observations	1,261	3,512
Mean series length	21.4	21.3

Notes: *** p < 0.01, Panel A comprises 59 countries, Panel B comprises 165 countries, the cross sectional dependence statistic follows Pesaran (2021) and is asymptotically standard normal under the null of independence, the cross sectional dependence statistics are computed on the full panel before lag based rows are dropped whereas the country year observation counts refer to the smaller estimation sample, the machine learning residual row is computed after the cross fitted conditioning step described above.

Table 4. Machine learning estimates, full panel, single centroid drought index.

Panel A: Average effect of DR (Stage 1)

Quantity	DML-W	DML-FD	ADML
Average effect of DR	-0.0022	-0.0013	-0.0008
Countries	165	165	165
Observations	3,512	3,347	3,512

Panel B: DR coefficient across LA (Stage 2)

Quantity	MLK	GRF	LLF	CFFE
Coefficient, 5th pct LA	0.0368***	-	0.0074	-
Coefficient, median LA	0.0032	-	-	-
Coefficient, 95th pct LA	-0.0255**	-	-0.0194	-
Contrast, 95th minus 5th pct	-0.0623***	-0.0268*	-0.0213	-0.0109
Average effect of DR	-	-0.0047	-	-0.0023
Gradient of DR on LA	-	-	-	-0.0015***
Countries	165	165	165	165
Observations	3,512	3,512	3,512	3,512

Panel C: Robustness to income and cross sectional dependence (Stage 3)

Quantity	BGATE	OSC	CCEMG	CCE-DML
Average effect of DR	-	-	-0.0005	-0.0048**
Coefficient on GDP	-	-	0.0916	-
Coefficient on POPD	-	-	0.7289**	-
Coefficient on URB	-	-	-0.0086	-
Coefficient on AGVA	-	-	-0.0116**	-
Contrast, 95th minus 5th pct	-	-0.0561**	-	0.0008
Gradient of DR on LA	-0.0057***	-	0.0000	-
Countries	165	165	165	165
Observations	3,512	3,512	3,677	3,677

Panel D: DML-VC (Stage 4)

Quantity	DML-VC
Coefficient, 5th pct LA	0.0310***
Coefficient, median LA	0.0020
Coefficient, 95th pct LA	-0.0279**
Contrast, 95th minus 5th pct	-0.0588***
Gradient of DR on LA	-0.0060**
Countries	165
Observations	3,512

Notes: this note applies to Panels A to D, * p < 0.10, ** p < 0.05, *** p < 0.01, DR is drought, LA is log land area, a dash indicates the estimator does not target that quantity or, for the four controls in Panel C, that the estimator places the control inside a nonparametric conditioning set rather than estimating it as a linear coefficient, estimator acronyms are defined at first use in the methods, BGATE equalises the GDP distribution across LA groups before estimation, CFFE reports the estimator at the depth selected by held out cross validation, the gradient row reports the change in the DR coefficient for a one unit increase in LA, the CCE mean group control coefficients are trimmed at the 2nd and 98th percentile across countries before averaging, standard errors are country clustered for scalar parameters and obtained by country level block bootstrap for the coefficient function, the forest and spline contrasts are bootstrapped over 60 and 499 country resamples respectively.

Sensitivity to the construction of the drought index

Table 5 repeats the analysis on the area weighted reconstruction, built to test whether the moderation reflects measurement imprecision in smaller countries, and on the narrower panel; the result is more equivocal than the single centroid index. Under the area weighted construction the contrast weakens materially, and for MLK the bootstrap confidence interval no longer excludes zero; DML-VC retains an interval that excludes zero on the same data, with the linear gradient significant at the 10% level. The moderation therefore survives in sign and approximate magnitude but is no longer precisely estimated, so a component of the gradient is attributable to how the index aggregates across grid cells. The third hypothesis is only partially supported, and this is the clearest empirical case for DML-VC. On the narrower panel the gradient is larger in magnitude and less precisely estimated, consistent with its smaller sample, and the linear gradient is significant at the 1% level.

Table 5. Sensitivity to the drought index construction and to the sample.

Quantity	Area weighted, MLK	Area weighted, DML-VC	Panel A, MLK	Panel A, DML-VC
Average effect of DR	-0.0027	-	-0.0207**	-
Coefficient, 5th pct LA	0.0135	0.0151	0.0096	0.0093
Coefficient, median LA	-0.0047	-0.0013	-0.0060	-0.0149*
Coefficient, 95th pct LA	-0.0233**	-0.0244*	-0.0613	-0.0569
Contrast, 95th minus 5th pct	-0.0368	-0.0395*	-0.0709	-0.0662
Confidence interval, contrast	[-0.0847, 0.0137]	[-0.0824, -0.0019]	[-0.1555, 0.0083]	[-0.1463, 0.0032]
Gradient of DR on LA	-	-0.0044*	-	-0.0135***
Countries	165	165	59	59

Notes: * p < 0.10, ** p < 0.05, *** p < 0.01, DR is drought and LA is log land area, a dash indicates the estimator does not target that quantity, the area weighted columns use the drought index constructed from three interior sampling points for every country below the median land area, Panel A comprises the 59 country sample, confidence intervals are country level block bootstrap intervals at the 95% level, the gradient row reports the coefficient on DR interacted with centred LA.

Cross sectional dependence

Cross sectional dependence is present in the two way within residuals of both panels, reported in Table 3 above. Two forms of common correlated effects correction are distinguished, because they give different answers. A correction that adds cross sectional averages to a

specification from which year effects are already removed is degenerate. Such averages vary only across years and are reduced to the order of 10 to the power minus 12 by the two way within transformation, leaving the dependence statistic unchanged to three decimal places. A correction that instead removes only country means, entering the cross sectional averages at raw level with freely estimated loadings, is not degenerate. Under that specification the dependence statistic falls from 8.124 to 6.599 on the wider panel, and the land area interaction remains negative and significant at the 1% level under both index constructions: the gradient survives a correction in which every country carries the same loading.

It does not survive a correction in which every country carries its own loading. Estimating the loadings country by country, and projecting the cross sectional averages out of each country series before the machine learning stage, removes the gradient entirely. Table 6 shows why that finding must be interpreted with care. As the projection is enriched from a single cross sectional average to the full set, the contrast decays smoothly and monotonically, and the decay tracks the share of the drought variation that the projection has absorbed. By the richest projection the country level regressions have consumed almost half of the treatment variation and retain roughly 15 degrees of freedom per country. A gradient driven by common shocks would collapse discontinuously once the dominant factor was removed rather than bleeding away in proportion to lost variation. The evidence is therefore consistent with attrition of the identifying variation rather than with confounding, but the two cannot be separated at the series length available here, a limitation of the data rather than a refutation of the result.

The same pattern is visible across every specification in Fig. 6, which collects the main estimation results and the two sensitivity analyses in a single display. Panel A plots the estimated DR coefficient across LA under the single centroid index and panel B under the area weighted index. Panel C shows the sensitivity of the LA contrast to the richness of the factor projection, and panel D compares the contrast across estimators, with open markers denoting estimators for which no interval is available.

Table 6. Sensitivity of the land area contrast to the richness of the factor projection.

Projection	Degrees of freedom per country	DR variation retained	Contrast, 95th minus 5th	Confidence interval
No factor projection	20.3	1.000	-0.0708***	[-0.1050, -0.0374]
Outcome average only	20.3	0.887	-0.0512**	[-0.1231, -0.0034]
Outcome and DR averages	19.3	0.801	-0.0364	[-0.0932, 0.0048]
Adding the income average	18.3	0.703	-0.0156	[-0.0481, 0.0122]
All six averages	15.3	0.526	-0.0014	[-0.0132, 0.0127]

Notes: * p < 0.10, ** p < 0.05, *** p < 0.01, DR is drought, each row projects the listed cross sectional averages out of the outcome, the treatment and the covariates country by country before the machine learning stage, the DR variation retained column reports the variance of the projected DR index relative to its variance under the two way within transformation alone, confidence intervals are country level block bootstrap intervals at the 95% level.

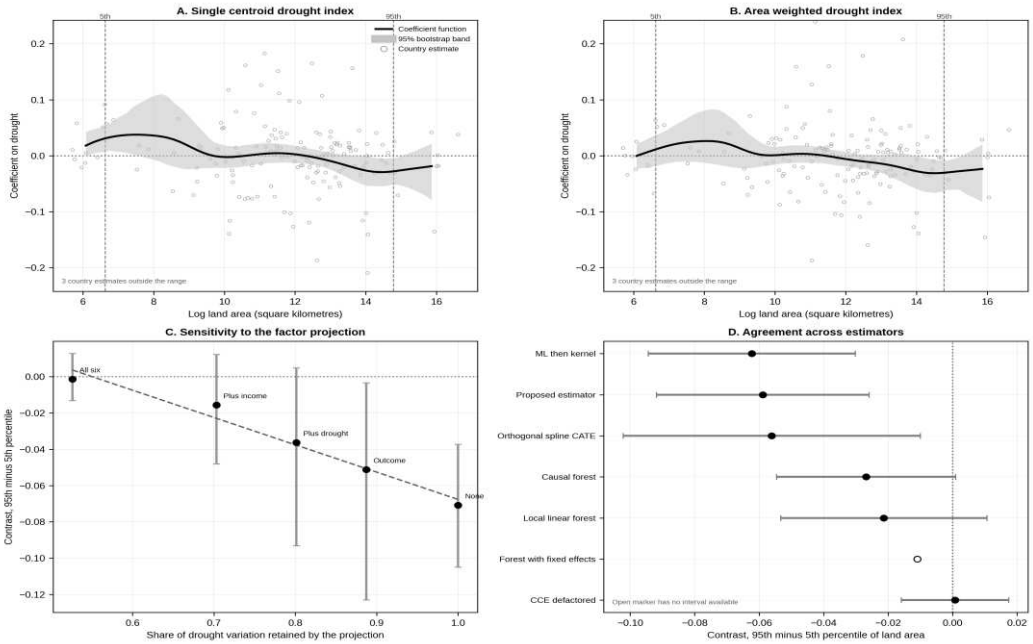


Fig. 6 Drought coefficient across land area and sensitivity analyses

Selection of the forest depth

Table 7 reports the held-out loss used to select the depth of the causal forest with fixed effects. Measured heterogeneity rises with depth over the range examined, so a depth chosen by inspection would be arbitrary. The held-out loss is minimised at the shallowest depth, which identifies the rise as overfitting rather than as additional structure recovered from the data.

Filling in the reported series can be measured without external metadata. A value carried forward produces a first difference of zero, and a value placed on a straight line between two compiled points produces a second difference of zero, so the annual values at which the second difference departs from zero are the values at which new information enters. Counting them gives a lower bound on the number of independent measurements for each country.

Filling is extensive. Averaged across countries, 36.7% of annual values carry information beyond a filled value, and the water stress series never changes at all in 16 countries. Any estimate drawn from this indicator rests on considerably less information than the count of country year observations suggests, and that is a property of the indicator rather than of the sample assembled here.

Filling is nonetheless unrelated to the moderator. The correlation between the observed share and log land area is 0.046, which is not distinguishable from zero, and Panel A of Table 8 shows no ordering across land area quintiles. Countries whose series never changes are spread across the distribution and are absent only from the largest quintile. Filling cannot therefore generate a relationship between the drought coefficient and country scale, because it is not distributed along the dimension that relationship is measured over.

Panel B of Table 8 confirms this directly. Restricting the sample to the 92 countries whose series carries genuine variation removes 44% of the observations and leaves the gradient intact, with a contrast slightly larger in magnitude and an interval that continues to exclude zero. The interval widens, as expected with fewer countries contributing.

Two limitations attach to this exercise. The count of informative values is a lower bound, since a filled value that happens to coincide with the path the series would have followed cannot be detected. Restricting the sample also removes the countries with the least information rather than the filling present within the countries retained.

Table 7. Depth selection for the causal forest with fixed effects.

Maximum depth	Held out loss	Standard deviation across folds	Dispersion of the estimated coefficient function
2	0.033937	0.012329	0.0046
3	0.034020	0.012283	0.0112
4	0.035662	0.010876	0.0140
6	0.035327	0.011166	0.0164
8	0.035319	0.011185	-

Notes: the held out loss is computed on countries withheld from estimation, with folds formed by country, the dispersion column reports the standard deviation of the estimated coefficient function across observations, the selected depth is the one minimising the held out loss.

Table 8. Extent of filling in the reported water stress series and its consequence for the gradient.

Panel A: Observed share by land area quintile

Quintile	Countries	Mean log area	Observed share	Constant series
1, smallest	33	8.21	0.324	5
2	33	10.84	0.481	4
3	33	12.03	0.283	4
4	33	13.02	0.370	3
5, largest	33	14.50	0.378	0

Panel B: Gradient by sample

Quantity	Full sample	Well observed sample
Countries	165	92
Observations	3,512	1,981
Contrast, 95th minus 5th percentile	-0.0779***	-0.0904***
Lower bound of the 95% interval	-0.1179	-0.1809
Upper bound of the 95% interval	-0.0451	-0.0287

Notes: *** $p < 0.01$, the observed share is the proportion of annual values at which the second difference of the reported water stress series departs from zero and which therefore carry information beyond a carried forward or linearly filled value, a constant series is one in which the reported value never changes over the observation window, the well observed sample retains countries whose observed share exceeds 0.20 and whose series takes more than one value, intervals are country block bootstrap intervals from 399 replications.

Land area is one of several correlated descriptions of country scale, and the data can separate some of them but not all. Table 9 reports the coefficient function estimated over three alternatives, with the moderator held fixed at the country mean in every case. Because the range of each moderator differs, the table also reports the slope for each unit of the moderator, which is the quantity that can be compared across columns.

Population density carries almost no gradient. The slope for each unit of density is a sixth of the slope for each unit of land area, and the interval covers zero comfortably, so the pattern is not a density effect and does not arise because populous countries differ from empty ones in how water is used.

Total population behaves almost identically to land area. The slope for each unit of the two moderators differs only in the fourth decimal place. Land area and population are strongly correlated across countries, and this panel cannot establish which of the two is the operative dimension. The reading offered here is that both describe the scale of a country, and that the estimates identify scale rather than land area specifically.

The gradient is also concentrated in the smallest countries. Excluding the 24 small island developing states halves the slope for each unit of land area and leaves an interval that includes zero. Part of the fall in the contrast follows mechanically from the narrower range that remains once the smallest countries are removed, but the slope falls by half as well, which does not. The relationship is therefore not a smooth gradient running across the whole size

distribution. It is better described as an absence of any drought response in very small countries, and particularly in small island states, against a negative response elsewhere.

That description sits closer to the mechanism than a gradient in land area does. Small island economies draw on groundwater, rainwater harvesting and desalination, and a national average climate index has no large basin to integrate over, so the index carries little information about the water actually available. The claim supported by these estimates is accordingly narrower than a monotone relationship with land area.

Table 9. Alternative moderators and the exclusion of small island developing states.

Quantity	Land area	Land area excluding island states	Total population	Population density
Countries	165	141	165	165
Moderator range, 95th minus 5th percentile	8.15	5.59	6.03	4.90
Contrast, 95th minus 5th percentile	-0.0779***	-0.0292	-0.0572	-0.0080
Lower bound of the 95% interval	-0.1179	-0.0592	-0.1266	-0.0583
Upper bound of the 95% interval	-0.0451	0.0013	0.0009	0.0389
Slope for each unit of the moderator	-0.0096	-0.0052	-0.0095	-0.0016

Notes: *** $p < 0.01$, every moderator is measured in logarithms and held fixed at the country mean, the slope for each unit of the moderator is the contrast divided by the moderator range and is reported because the range differs across columns, small island developing states follow the standard United Nations membership list and 24 of them appear in the estimation sample, intervals are country block bootstrap intervals from 399 replications.

A common factor can generate an apparent relationship between the drought coefficient and country scale, but only under a specific condition. The spurious component of the country coefficient is governed by the product of the loading of drought on the factor and the loading of water stress on the same factor. That product varies with land area only when both loadings vary with land area. If either loading is unrelated to land area, the factor structure cannot manufacture a gradient, whatever the strength of the dependence it induces.

Table 10 reports the relevant correlations. The loadings of water stress are unrelated to land area on every factor, with correlations of 0.022, 0.007, -0.056, and 0.114, none of which is distinguishable from zero. The loadings of drought are also close to zero, and only the second factor reaches the 5% level, which is the number of rejections eight independent tests would be expected to deliver by chance. The condition under which the factor structure could produce the observed gradient is therefore not satisfied in this panel.

The same estimates explain what the projection ladder removes. The four retained factors account for 54.2% of the variance of within transformed drought, leaving 45.8% idiosyncratic. The richest projection retains 52.6% of drought variation, which is close to that idiosyncratic share, so the ladder removes most of the common component of drought and little else. Drought is a spatially correlated phenomenon, since regional circulation produces simultaneous deficits across many countries at once, and the common component is therefore the substantive part of the signal rather than a nuisance. A correction that removes it removes most of what the drought index measures.

This does not establish that the gradient is causal. It establishes that the collapse of the contrast under the richest projection is uninformative about whether the gradient is real, because the projection removes the variation in which drought is chiefly expressed. The analysis cannot exclude the possibility that the gradient operates through the common component in a manner that a full projection legitimately removes, and the interpretation offered here rests on the absence of any association between the loadings and land area.

Table 10. Correlation between country factor loadings and log land area.

Factor	Variance share	Drought loading	Water stress loading
1	0.241	-0.071	0.022
2	0.136	-0.177**	0.007
3	0.091	-0.129	-0.056
4	0.074	0.034	0.114

Notes: ** $p < 0.05$, entries are correlations between the country loading on each principal component of within transformed drought and log land area, loadings are estimated over a balanced rectangle of 22 years and 152 countries, four principal components are retained because four are required to account for at least 50% of the variance of within transformed drought.

4.2 Discussion

The central finding is that the relationship between a precipitation based drought index and SDG 6.4.2 depends on the scale of a country, and is absent in the smallest countries and particularly in small island states. This dependence survives estimation strategies that place no functional form restriction on the role of the covariates. 10 estimators drawn from three distinct methodological families agree on the direction and, where they target the same quantity, on the approximate magnitude. That agreement is the strongest evidence the study offers, because the families differ in where they place the flexibility. The debiased machine learning family keeps the target parameter low dimensional and puts the flexibility in the conditioning functions. The forest family puts the flexibility in the target and discovers the pattern from the data, and the two step family does both.

The pattern points to a mismatch between the spatial scale at which a climate index is constructed and the scale at which the water system it proxies for actually operates. This is consistent with the small island water security literature. Parker et al. (2023) and Gheuens et al. (2019) describe water systems in small states governed by infrastructure and local hydrogeology rather than a national rainfall anomaly, and Gungoa and Kebede (2024) provide direct empirical support for this disconnection in a groundwater quality setting. None of these studies frames its findings in terms of land area specifically, and this study does not claim that land area is the causal variable rather than a correlate of the hydrological scale mismatch it proxies for.

Three further limitations qualify these conclusions. Land area is an observable proxy for the underlying hydrological scale mechanism rather than the mechanism itself. Income, urbanisation and the other covariates are treated as exogenous throughout, with no instrument proposed for the plausible reverse channel from water stress to economic structure. And the moderation weakens under the area weighted index and does not survive a rich factor projection, both qualifications already discussed above.

5. Conclusion

This study finds that a standard precipitation-based drought index does not bear the same relationship to SDG 6.4.2 across countries of different geographic scale. The relationship is negative and significant in larger countries and absent in the smallest, and particularly in small island states, a pattern on which 10 estimators from three methodological families agree. The two qualifications noted earlier, the sensitivity to index construction and the loss of the gradient under a rich factor projection, are not independent findings but two views of the same limitation. The series length available cannot fully separate a genuine measurement artefact

from attrition of identifying variation, and a longer or subnational panel is the natural way to resolve that ambiguity.

Subject to those qualifications, the result suggests that cross country panels using national average climate indices as covariates for water related outcomes should treat country scale as a relevant source of heterogeneity. SDG 6.4.2 in particular may require different explanatory frameworks for large, hydrologically integrated economies than for small states whose water security depends on sources other than surface hydrology. Two extensions follow directly. An estimator that models group structure and factor structure jointly would address the dependence problem without consuming the within country variation that a country-by-country projection removes. A longer panel, or one at subnational resolution, would allow the measurement and confounding explanations of the weakened gradient to be distinguished.

Appendix A. Additional tables

Table A1. Country samples.

Afghanistan (AFG)	Angola (AGO)	Albania (ALB)	United Arab Emirates (ARE)*
Argentina (ARG)*	Antigua and Barbuda (ATG)	Australia (AUS)	Austria (AUT)
Azerbaijan (AZE)	Burundi (BDI)	Belgium (BEL)*	Benin (BEN)
Burkina Faso (BFA)	Bangladesh (BGD)*	Bulgaria (BGR)	Bahrain (BHR)
Bosnia and Herzegovina (BIH)	Belarus (BLR)	Belize (BLZ)	Bolivia (BOL)*
Brazil (BRA)*	Barbados (BRB)	Brunei Darussalam (BRN)	Bhutan (BTN)
Botswana (BWA)*	Canada (CAN)*	Switzerland (CHE)*	Chile (CHL)*
China (CHN)*	Côte d'Ivoire (CIV)	Cameroon (CMR)	Democratic Republic of the Congo (COD)
Republic of the Congo (COG)	Colombia (COL)*	Comoros (COM)	Republic of Cabo Verde (CPV)
Cuba (CUB)	Cyprus (CYP)	Czech Republic (CZE)	Germany (DEU)*
Dominica (DMA)	Denmark (DNK)	Dominican Republic (DOM)	Algeria (DZA)
Ecuador (ECU)*	Egypt (EGY)*	Spain (ESP)*	Estonia (EST)
Ethiopia (ETH)*	Finland (FIN)	Fiji (FJI)	France (FRA)*
Gabon (GAB)	United Kingdom (GBR)*	Georgia (GEO)	Ghana (GHA)*
Guinea (GIN)	The Gambia (GMB)	Guinea-Bissau (GNB)	Equatorial Guinea (GNQ)
Greece (GRC)*	Grenada (GRD)	Guatemala (GTM)	Guyana (GUY)
Honduras (HND)	Croatia (HRV)	Haiti (HTI)	Hungary (HUN)*
Indonesia (IDN)*	India (IND)*	Ireland (IRL)	Iran (IRN)
Iraq (IRQ)	Iceland (ISL)	Israel (ISR)*	Italy (ITA)*
Jamaica (JAM)	Jordan (JOR)*	Japan (JPN)*	Kazakhstan (KAZ)
Kenya (KEN)*	Kyrgyzstan (KGZ)	Cambodia (KHM)	Republic of Korea (KOR)*
Kuwait (KWT)*	Lao PDR (LAO)	Lebanon (LBN)	Liberia (LBR)
Libya (LBY)	Saint Lucia (LCA)	Sri Lanka (LKA)*	Lesotho (LSO)
Lithuania (LTU)	Luxembourg (LUX)	Latvia (LVA)	Morocco (MAR)*
Moldova (MDA)	Madagascar (MDG)	Maldives (MDV)	Mexico (MEX)*
North Macedonia (MKD)	Mali (MLI)	Malta (MLT)	Myanmar (MMR)*
Mongolia (MNG)*	Mozambique (MOZ)	Mauritania (MRT)	Mauritius (MUS)
Malawi (MWI)	Malaysia (MYS)	Namibia (NAM)	Niger (NER)
Nigeria (NGA)*	Nicaragua (NIC)	Netherlands (NLD)*	Norway (NOR)*
Nepal (NPL)*	New Zealand (NZL)	Oman (OMN)	Pakistan (PAK)*
Peru (PER)*	Philippines (PHL)*	Papua New Guinea (PNG)	Poland (POL)*
Portugal (PRT)*	Paraguay (PRY)*	PSE (PSE)	Qatar (QAT)*
Romania (ROU)*	Russian Federation (RUS)	Rwanda (RWA)	Saudi Arabia (SAU)*
Senegal (SEN)	Singapore (SGP)	Sierra Leone (SLE)	El Salvador (SLV)
Serbia (SRB)	São Tomé and Principe (STP)	Suriname (SUR)	Slovakia (SVK)
Slovenia (SVN)	Sweden (SWE)*	Kingdom of eSwatini (SWZ)	Syria (SYR)
Chad (TCD)	Togo (TGO)	Thailand (THA)*	Tajikistan (TJK)
Turkmenistan (TKM)	Timor-Leste (TLS)	Trinidad and Tobago (TTO)	Tunisia (TUN)*
Turkey (TUR)	Tanzania (TZA)*	Uganda (UGA)	Ukraine (UKR)
Uruguay (URY)*	United States (USA)*	Uzbekistan (UZB)	Saint Vincent and the Grenadines (VCT)

Vietnam (VNM)*	Yemen (YEM)*	South Africa (ZAF)*	Zambia (ZMB)*
Zimbabwe (ZWE)			

Notes: the wider panel comprises 165 countries and the narrower panel comprises 59 countries, an asterisk marks membership of the narrower panel.

Table A2. Descriptive statistics.

Variable	Mean	Std. dev.	Min.	Max.	N
WS (%)	68.162	289.549	0.027	3850.500	3,677
GDP (US$)	13,064.954	18,435.817	243.077	112,422.146	3,677
POPD	190.747	598.294	1.543	7,965.878	3,677
URB (%)	57.636	22.219	8.044	100.000	3,677
AGVA (%)	11.185	10.786	0.026	79.042	3,677
DR (standardised)	-0.003	0.798	-2.025	2.025	3,677

Appendix B. Technical derivation of the nonparametric varying-coefficient estimator

B.0 Attenuation bias from the unobserved hydrological state

SDG 6.4.2 is the ratio of total freshwater withdrawal to total renewable freshwater resources. Water stress for country i in year t, denoted W_{it}, is taken to respond to the true state of the hydrological system that supplies it, denoted H_{it}, together with controls $\mathbf{x}_{it}$:

$$W_{it} = \beta_i H_{it} + \mathbf{x}_{it}'\boldsymbol{\gamma} + \alpha_i + \lambda_t + \varepsilon_{it}, \tag{B.1}$$

where β_i is a country specific sensitivity of water stress to the hydrological state. H_{it} is not observed; what is observed is D_{it}, a national average drought index constructed from precipitation and evapotranspiration, which decomposes as:

$$D_{it} = H_{it} + v_{it}(A_i), \tag{B.2}$$

where $v_{it}(A_i)$ is the gap between the national climate reading and the hydrological state actually governing supply, and A_i is national land area. The gap widens as A_i falls. A national average is a less complete spatial description of climate in a small country. More consequentially, the mapping from a national climate reading to water supply requires a river basin or reservoir system large enough to integrate rainfall variation across a wide area. Where supply instead depends on groundwater lenses, rainwater harvesting, or desalination, the climate reading and the hydrological state become structurally disconnected rather than merely noisily related, since none of these sources responds to a rainfall anomaly the way a river basin does.

Substituting D_{it} for the unobserved H_{it} and estimating by ordinary least squares does not recover β_i. What is recovered is an attenuated coefficient:

$$\hat{\beta}_i = \beta_i \cdot \frac{\mathrm{Cov}(H_{it}, D_{it})}{\mathrm{Var}(D_{it})} = \beta_i \cdot \big(1 - \phi(A_i)\big), \tag{B.3}$$

where $\phi(A_i) \in [0,1]$ is the share of the variance of D_{it} attributable to the gap rather than to genuine variation in H_{it}, and is decreasing in land area. This is the formal basis for allowing the estimated coefficient on drought to vary smoothly with $\log A_i$ in the structural model of water stress set out above. Where $\phi(A_i)$ is close to one, the coefficient should sit near zero or take an unpredictable sign; as $\phi(A_i)$ falls toward zero, it should move toward β_i itself, correctly signed and larger in magnitude.

B.1 Model and notation

Let $i = 1, \ldots, N$ index countries and $t = 1, \ldots, T_i$ index years. The outcome y_{it} denotes log water stress, the continuous treatment d_{it} denotes the standardised drought index, and $\mathbf{x}_{it}$ denotes a p-vector of controls. The moderator c_i denotes log land area and is time invariant. The structural model is:

$$y_{it} = \beta(c_i)\, d_{it} + g(\mathbf{x}_{it}) + \alpha_i + \lambda_t + \varepsilon_{it}, \qquad E[\varepsilon_{it} \mid d_{it}, \mathbf{x}_{it}, c_i] = 0, \tag{B.4}$$

in which $\beta(\cdot)$ is an unknown smooth function on the support $\mathcal{C}$ of c_i, $g(\cdot)$ is an unknown nuisance function, and α_i and λ_t are country and year effects. The object of interest is $\beta(c)$ at points $c \in \mathcal{C}$, together with the contrast $\beta(c_{\text{high}}) - \beta(c_{\text{low}})$ evaluated at the upper and lower quantiles of the land-area distribution. Model (B.4) is the panel varying-coefficient analogue of the partially linear specification of Robinson (1988), and, with $\beta(c_i)$ restricted to a single constant, reduces to the two-way within nonparametric panel model of Lee and Robinson (2015).

Two features distinguish this specification from the cross-sectional two-step problem treated by Zimmert and Lechner (2026), and one further feature distinguishes it from the kernel-only panel varying-coefficient estimator of Halder and Malikov (2020). The coefficient function is indexed by a time-invariant variable, so the second stage has N rather than $\sum_i T_i$ distinct design points. The additive effects α_i and λ_t must be removed before orthogonalisation, which induces dependence across observations that the cross-sectional theory does not accommodate. Halder and Malikov (2020) estimate an equation of this form with the nuisance function $g(\cdot)$ removed by kernel smoothing rather than by machine learning. The estimator below replaces that kernel step with the cross-fitted first stage described in B.2, and collapses to their smoothed least squares dummy variable estimator if that replacement is undone.

Let $\ddot{v}_{it}$ denote the two-way within transformation of any variable v_{it}, defined as the residual from the projection of v_{it} on the country and year indicator spaces. Applying the transformation to the model gives:

$$\ddot{y}_{it} = \beta(c_i)\,\ddot{d}_{it} + \ddot{g}_{it} + \ddot{\varepsilon}_{it}, \tag{B.5}$$

where $\ddot{g}_{it}$ denotes the within-transformed nuisance term. The transformation leaves $\beta(c_i)$ unaffected because c_i is constant within country.

B.2 The orthogonal moment condition

Define the nuisance functions:

$$\ell_0(\mathbf{x}) = E[\ddot{y}_{it} \mid \ddot{\mathbf{x}}_{it} = \mathbf{x}], \qquad m_0(\mathbf{x}) = E\big[\ddot{d}_{it} \mid \ddot{\mathbf{x}}_{it} = \mathbf{x}\big], \tag{B.6}$$

and collect them in $\eta_0 = (\ell_0, m_0)$. The orthogonalised residuals are:

$$\tilde{y}_{it} = \ddot{y}_{it} - \ell_0(\ddot{\mathbf{x}}_{it}), \qquad \tilde{d}_{it} = \ddot{d}_{it} - m_0(\ddot{\mathbf{x}}_{it}). \tag{B.7}$$

The moment function underlying the estimator is:

$$\psi(W_{it}; \beta, \eta) = \big\{\ddot{y}_{it} - \ell(\ddot{\mathbf{x}}_{it}) - \beta\big[\ddot{d}_{it} - m(\ddot{\mathbf{x}}_{it})\big]\big\}\big[\ddot{d}_{it} - m(\ddot{\mathbf{x}}_{it})\big], \tag{B.8}$$

with $W_{it} = \big(\ddot{y}_{it}, \ddot{d}_{it}, \ddot{\mathbf{x}}_{it}, c_i\big)$. At the true values the conditional moment condition $E[\psi(W_{it}; \beta(c_i), \eta_0) \mid c_i] = 0$ holds by construction.

The Gateaux derivative of the moment in the direction $\eta - \eta_0$, evaluated at η_0, is:

$$\partial_\eta E[\psi(W_{it}; \beta, \eta_0) \mid c_i][\eta - \eta_0] = -E\big[(\ell - \ell_0)\tilde{d}_{it} \mid c_i\big] + \beta\, E\big[(m - m_0)\tilde{d}_{it} \mid c_i\big] - E[\tilde{\varepsilon}_{it}(m - m_0) \mid c_i]. \tag{B.9}$$

Each of the three terms vanishes by iterated expectations, since $E\big[\tilde{d}_{it} \mid \ddot{\mathbf{x}}_{it}, c_i\big] = 0$ and $E[\tilde{\varepsilon}_{it} \mid \ddot{\mathbf{x}}_{it}, c_i] = 0$. The moment is therefore Neyman orthogonal, so first-order errors in $\hat{\eta}$ do not transmit to the estimate of $\beta(c)$.

B.3 Reduction to country-level moments

Because $\beta(\cdot)$ depends on the data only through c_i, the moment condition aggregates within country without loss of information. Define the country-level orthogonal moments:

$$A_i = \textstyle\sum_{t=1}^{T_i} \tilde{d}_{it}^{\,2}, \qquad B_i = \textstyle\sum_{t=1}^{T_i} \tilde{d}_{it}\ \tilde{y}_{it}. \tag{B.10}$$

The pair $(A_i, B_i, c_i)_{i=1}^{N}$ constitutes a sufficient reduction of the second-stage estimating equations, and the effective sample size for the estimation of $\beta(\cdot)$ is N, not $\sum_i T_i$. This reduction

is the formal counterpart of the practical point that a smoother in c_i cannot extract more than N independent design points from the panel.

B.4 The local linear varying-coefficient estimator

Let $K(\cdot)$ denote a second-order kernel, h a bandwidth, and $K_h(u) = h^{-1}K(u/h)$. The estimator solves the locally weighted least-squares problem:

$$\left(\hat{a}(c), \hat{b}(c)\right) = \underset{a,b}{\operatorname{argmin}} \sum_{i=1}^{N} K_h\,(c_i - c) \sum_{t=1}^{T_i} \left[\tilde{y}_{it} - a\,\tilde{d}_{it} - b\,(c_i - c)\,\tilde{d}_{it}\right]^2, \qquad \text{(B.11)}$$

and sets $\hat{\beta}(c) = \hat{a}(c)$. Writing $u_i = c_i - c$ and $w_i = K_h(u_i)$, and defining:

$$S_r(c) = \sum_{i=1}^{N} w_i\, A_i u_i^r, \quad r = 0,1,2, \qquad Q_r(c) = \sum_{i=1}^{N} w_i\, B_i u_i^r, \quad r = 0,1, \qquad \text{(B.12)}$$

the first-order conditions admit the closed form:

$$\hat{\beta}(c) = \frac{S_2(c)\,Q_0(c) - S_1(c)\,Q_1(c)}{S_0(c)\,S_2(c) - S_1(c)^2}. \qquad \text{(B.13)}$$

The local constant estimator obtained by imposing $b = 0$ reduces to $\hat{\beta}_{\mathrm{lc}}(c) = Q_0(c)/S_0(c)$. The local linear form is retained because the contrast of interest is evaluated at the 5th and 95th percentiles of the land-area distribution, which lie in the boundary region where the local constant estimator carries first-order bias of order h rather than h^2.

B.5 Conditions and asymptotic distribution

The following conditions are imposed. The sequence (A_i, B_i, c_i) is independent across i with $E\left[A_i^2\right] < \infty$ and $E\left[B_i^2\right] < \infty$. The density $f(\cdot)$ of c_i is bounded away from zero on a neighbourhood of c and twice continuously differentiable. The coefficient function $\beta(\cdot)$ is twice continuously differentiable on that neighbourhood. The kernel is symmetric with $\int K = 1$, $\mu_2 = \int u^2 K(u)\,du < \infty$ and $\nu_0 = \int K(u)^2 du < \infty$. The bandwidth satisfies $h \to 0$ and $Nh \to \infty$. The nuisance estimator satisfies $\| \hat{\eta} - \eta_0 \| = O_p(r_{NT})$ with the rate condition:

$$\sqrt{Nh}\; r_{NT}^2 \to 0. \qquad \text{(B.14)}$$

Define $q(c) = E\left[T_i^{-1} A_i \mid c_i = c\right]$ and $\Omega(c) = E\left[T_i^{-2}\left(\sum_t \tilde{d}_{it}\,\tilde{\varepsilon}_{it}\right)^2 \mid c_i = c\right]$. Under these conditions:

$$\sqrt{Nh}\left[\hat{\beta}(c) - \beta(c) - 1/2\, h^2 \mu_2\, \beta''(c) + o_p(h^2)\right] \xrightarrow{d} N\left(0,\ \frac{\nu_0\,\Omega(c)}{f(c)\,q(c)^2}\right). \qquad \text{(B.15)}$$

The argument follows the standard local linear expansion. The denominator $S_0 S_2 - S_1^2$ converges after normalisation to $f(c)\,q(c)^2\mu_2$ by the law of large numbers for independent triangular arrays and the continuous mapping theorem. The numerator decomposes into two terms. The first is a smoothing bias term, which the local linear weighting reduces to order h^2 uniformly, including at the boundary. The second is a stochastic term to which the Lindeberg–Feller central limit theorem for independent non-identically distributed arrays applies, the Lindeberg condition being implied by the second-moment restriction on B_i. The contribution of nuisance estimation error is of order r_{NT}^2 by Neyman orthogonality established in B.2, and is asymptotically negligible under the stated rate condition.

The rate condition merits comment. The nuisance functions are estimated on $\sum_i T_i$ observations while the smoother converges at rate $(Nh)^{-1/2}$, so the coupling between the two stages is asymmetric. The requirement $\sqrt{Nh}\, r_{NT}^2 \to 0$ is weaker than the corresponding cross-sectional condition $\sqrt{Nh}\, r_N^2 \to 0$ of Zimmert and Lechner (2026), because $r_{NT} \le r_N$ for any learner whose convergence rate is monotone in sample size. The panel structure therefore relaxes rather than tightens the admissible class of first-stage learners, provided the within transformation is applied before orthogonalisation.

B.6 Variance estimation

Let $\hat{\varepsilon}_{it} = \tilde{y}_{it} - \hat{\beta}(c)\,\tilde{d}_{it}$ denote the second-stage residual and let $\mathbf{z}_{it} = \big(\tilde{d}_{it},\ \tilde{d}_{it}u_i\big)'$. The estimating equations have the sandwich form. The bread is $\widehat{\mathbf{S}} = \sum_i \sum_t w_i\,\mathbf{z}_{it}\mathbf{z}_{it}'$, and two candidate meat matrices are considered:

$$\widehat{\mathbf{M}}_{\mathrm{ind}} = \sum_{i=1}^{N}\sum_{t=1}^{T_i} w_i^2\ \mathbf{z}_{it}\mathbf{z}_{it}'\,\hat{\varepsilon}_{it}^2, \qquad \widehat{\mathbf{M}}_{\mathrm{clu}} = \sum_{i=1}^{N}\Big(\sum_{t=1}^{T_i} w_i\ \mathbf{z}_{it}\,\hat{\varepsilon}_{it}\Big)\Big(\sum_{t=1}^{T_i} w_i\ \mathbf{z}_{it}\,\hat{\varepsilon}_{it}\Big)'. \tag{B.16}$$

The variance estimator $\hat{V}(c) = \big[\hat{\mathbf{S}}^{-1}\widehat{\mathbf{M}}_{\mathrm{clu}}\hat{\mathbf{S}}^{-1}\big]_{11}$ is consistent under the conditions of B.5, whereas the estimator built from $\widehat{\mathbf{M}}_{\mathrm{ind}}$ is inconsistent whenever the orthogonalised scores are serially correlated within country, since it omits the cross-products $\sum_{t\neq s} w_i^2\,\mathbf{z}_{it}\mathbf{z}_{is}'\hat{\varepsilon}_{it}\hat{\varepsilon}_{is}$. The magnitude of the resulting understatement depends on the within-country autocorrelation of the score rather than on T_i alone, and is therefore an empirical quantity rather than a fixed factor.

Inference is reported using the country block bootstrap, in which countries are resampled with replacement and the second stage is recomputed on each resample with the first-stage nuisance held fixed. Holding the nuisance fixed understates uncertainty by a term of order r_{NT}^2, which is negligible under the rate condition of B.5.

B.7 Bandwidth selection

The bandwidth is selected by leave-one-country-out cross-validation on the country-level moments, minimising:

$$\mathrm{CV}(h) = \sum_{i=1}^{N} \frac{\big[B_i - \widehat{\beta}^{(-i)}(c_i;h)\,A_i\big]^2}{A_i}, \tag{B.17}$$

where $\hat{\beta}^{(-i)}$ denotes the estimator computed with country i omitted from both the kernel weights and the moment sums. Cross-validation is conducted at the country level because the effective sample for the second stage is N, as established in B.3. A rule-of-thumb bandwidth constructed from $\sum_i T_i$ rather than N undersmooths by a factor of $(\sum_i T_i\,/N)^{1/5}$ and produces coefficient functions that are not monotone in the moderator even when the underlying function is.

The criterion is monotone decreasing in h over the admissible range for the data used here, so the selected bandwidth corresponds to the globally linear limit of the local linear estimator. This is itself an interpretable finding: the data support a coefficient function that is linear in log land area, and the linear interaction specification reported in the main text is the correct summary of the heterogeneity rather than a restriction imposed on it.

References

Alesina, A. (2003). The size of countries: Does it matter? Journal of the European Economic Association, 1(2-3), 301-316. https://doi.org/10.1162/154247603322390946

Alesina, A., & Spolaore, E. (1997). On the number and size of nations. The Quarterly Journal of Economics, 112(4), 1027-1056. https://doi.org/10.1162/003355300555411

Alesina, A., & Spolaore, E. (2005). War, peace, and the size of countries. Journal of Public Economics, 89(7), 1333-1354. https://doi.org/10.1016/j.jpubeco.2003.07.013

Alley, W. M., Healy, R. W., LaBaugh, J. W., & Reilly, T. E. (2002). Flow and storage in groundwater systems. Science, 296(5575), 1985-1990. https://doi.org/10.1126/science.1067123

Athey, S., Tibshirani, J., & Wager, S. (2019). Generalized random forests. The Annals of Statistics, 47(2), 1148-1178. https://doi.org/10.1214/18-AOS1709

Auffhammer, M. (2018). Quantifying economic damages from climate change. Journal of Economic Perspectives, 32(4), 33-52. https://doi.org/10.1257/jep.32.4.33

Bearth, N., & Lechner, M. (2025). Causal machine learning for moderation effects. Journal of Business and Economic Statistics, 43(4), 1158-1169. https://doi.org/10.1080/07350015.2025.2479646

Blakeslee, D., Fishman, R., & Srinivasan, V. (2020). Way down in the hole: Adaptation to long-term water loss in rural India. American Economic Review, 110(1), 200-224. https://doi.org/10.1257/aer.20180976

Breusch, T. S., & Pagan, A. R. (1980). The Lagrange multiplier test and its applications to model specification in econometrics. The Review of Economic Studies, 47(1), 239-253. https://doi.org/10.2307/2297111

Briguglio, L. (1995). Small island developing states and their economic vulnerabilities. World Development, 23(9), 1615-1632. https://doi.org/10.1016/0305-750x(95)00065-k

Burke, M., Hsiang, S. M., & Miguel, E. (2015). Global non-linear effect of temperature on economic production. Nature, 527(7577), 235-239. https://doi.org/10.1038/nature15725

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. The Review of Economics and Statistics, 90(3), 414-427. https://doi.org/10.1162/rest.90.3.414

Cashman, A. (2013). Water security and services in the Caribbean. Inter-American Development Bank. https://doi.org/10.18235/0009089

Chernozhukov, V., Newey, W. K., & Singh, R. (2022). Automatic debiased machine learning of causal and structural effects. Econometrica, 90(3), 967-1027. https://doi.org/10.3982/ECTA18515

Clarke, P. S., & Polselli, A. (2026). Double machine learning for static panel models with fixed effects. Econometrics Journal, 29(1), 69-86. https://doi.org/10.1093/ectj/utaf011

Cook, B. I., Mankin, J. S., & Anchukaitis, K. J. (2018). Climate change and drought: From past to future. Current Climate Change Reports, 4(2), 164-179. https://doi.org/10.1007/s40641-018-0093-2

Dai, A. (2013). Increasing drought under global warming in observations and models. Nature Climate Change, 3(1), 52-58. https://doi.org/10.1038/nclimate1633

Dalin, C., Wada, Y., Kastner, T., & Puma, M. J. (2017). Groundwater depletion embedded in international food trade. Nature, 543(7647), 700-704. https://doi.org/10.1038/nature21403

De Hoyos, R. E., & Sarafidis, V. (2006). Testing for cross-sectional dependence in panel-data models. The Stata Journal, 6(4), 482-496. https://doi.org/10.1177/1536867x0600600403

Dell, M., Jones, B. F., & Olken, B. A. (2014). What do we learn from the weather? The new climate-economy literature. Journal of Economic Literature, 52(3), 740-798. https://doi.org/10.1257/jel.52.3.740

Deryng, D., Conway, D., Ramankutty, N., Price, J., & Warren, R. (2014). Global crop yield response to extreme heat stress under multiple climate change futures. Environmental Research Letters, 9(3), 034011. https://doi.org/10.1088/1748-9326/9/3/034011

Elimelech, M., & Phillip, W. A. (2011). The future of seawater desalination: Energy, technology, and the environment. Science, 333(6043), 712-717. https://doi.org/10.1126/science.1200488

Friedberg, R., Tibshirani, J., Athey, S., & Wager, S. (2021). Local linear forests. Journal of Computational and Graphical Statistics, 30(2), 503-517. https://doi.org/10.1080/10618600.2020.1831930

Garrick, D. E., Hall, J. W., Dobson, A., Damania, R., Grafton, R. Q., Hope, R., Hepburn, C., Bark, R., Boltz, F., De Stefano, L., O'Donnell, E., Matthews, N., & Money, A. (2017). Valuing water for sustainable development. Science, 358(6366), 1003-1005. https://doi.org/10.1126/science.aao4942

Garrick, D., De Stefano, L., Yu, W., Jorgensen, I., O'Donnell, E., Turley, L., Aguilar-Barajas, I., Dai, X., de Souza Leão, R., Punjabi, B., Schreiner, B., Svensson, J., & Wight, C. (2019). Rural water for thirsty cities: A systematic review of water reallocation from rural to urban regions. Environmental Research Letters, 14(4), 043003. https://doi.org/10.1088/1748-9326/ab0db7

Gheuens, J., Nagabhatla, N., & Perera, E. D. P. (2019). Disaster-risk, water security challenges and strategies in small island developing states (SIDS). Water, 11(4), 637. https://doi.org/10.3390/w11040637

Grafton, R. Q., Williams, J., Perry, C. J., Molle, F., Ringler, C., Steduto, P., Udall, B., Wheeler, S. A., Wang, Y., Garrick, D., & Allen, R. G. (2018). The paradox of irrigation efficiency. Science, 361(6404), 748-750. https://doi.org/10.1126/science.aat9314

Gungoa, V., & Kebede, S. (2024). Climate variability impact on groundwater quality in Small Island Developing States: Mauritius Island as a case study. Frontiers in Water, 6, 1432908. https://doi.org/10.3389/frwa.2024.1432908

Halder, S. C., & Malikov, E. (2020). Smoothed LSDV estimation of functional-coefficient panel data models with two-way fixed effects. Economics Letters, 192, 109239. https://doi.org/10.1016/j.econlet.2020.109239

Härdle, W. (1990). Applied nonparametric regression. Cambridge University Press. https://doi.org/10.1017/ccol0521382483

Hoekstra, A. Y., & Hung, P. Q. (2005). Globalisation of water resources: International virtual water flows in relation to crop trade. Global Environmental Change, 15(1), 45-56. https://doi.org/10.1016/j.gloenvcha.2004.06.004

Hoekstra, A. Y., Mekonnen, M. M., Chapagain, A. K., Mathews, R. E., & Richter, B. D. (2012). Global monthly water scarcity: Blue water footprints versus blue water availability. PLoS ONE, 7(2), e32688. https://doi.org/10.1371/journal.pone.0032688

Hornbeck, R., & Keskin, P. (2014). The historically evolving impact of the Ogallala aquifer: Agricultural adaptation to groundwater and drought. American Economic Journal: Applied Economics, 6(1), 190-219. https://doi.org/10.1257/app.6.1.190

Judson, R. A., & Owen, A. L. (1999). Estimating dynamic panel data models: A guide for macroeconomists. Economics Letters, 65(1), 9-15. https://doi.org/10.1016/s0165-1765(99)00130-5

Kattenberg, M. A. C., Scheer, B. J., & Thiel, J. H. (2023). Causal forests with fixed effects for treatment effect heterogeneity in difference-in-differences (CPB Discussion Study). CPB Netherlands Bureau for Economic Policy Analysis.

Klosin, S., & Vilgalys, M. (2022). Estimating continuous treatment effects in panel data using machine learning with an agricultural application (arXiv Working Study No. 2207.08789). https://arxiv.org/abs/2207.08789

Koenker, R., & Bassett, G. (1978). Regression quantiles. Econometrica, 46(1), 33-50. https://doi.org/10.2307/1913643

Koenker, R. (2004). Quantile regression for longitudinal data. Journal of Multivariate Analysis, 91(1), 74-89. https://doi.org/10.1016/j.jmva.2004.05.006

Kummu, M., Guillaume, J. H. A., de Moel, H., Eisner, S., Flörke, M., Porkka, M., Siebert, S., Veldkamp, T. I. E., & Ward, P. J. (2016). The world's road to water scarcity: Shortage and stress in the 20th century and pathways towards sustainability. Scientific Reports, 6, 38495. https://doi.org/10.1038/srep38495

Lee, J., & Robinson, P. M. (2015). Panel nonparametric regression with fixed effects. Journal of Econometrics, 188(2), 346-362. https://doi.org/10.1016/j.jeconom.2015.03.004

Li, Q., & Racine, J. S. (2007). Nonparametric econometrics: Theory and practice. Princeton University Press.

Li, Q., & Racine, J. S. (2008). Nonparametric estimation of conditional CDF and quantile functions with mixed categorical and continuous data. Journal of Business and Economic Statistics, 26(4), 423-434. https://doi.org/10.1198/073500107000000250

Liu, J., Yang, H., Gosling, S. N., Kummu, M., Flörke, M., Pfister, S., Hanasaki, N., Wada, Y., Zhang, X., Zheng, C., Alcamo, J., & Oki, T. (2017). Water scarcity assessments in the past, present, and future. Earth's Future, 5(6), 545-559. https://doi.org/10.1002/2016ef000518

Machado, J. A. F., & Santos Silva, J. M. C. (2019). Quantiles via moments. Journal of Econometrics, 213(1), 145-173. https://doi.org/10.1016/j.jeconom.2019.04.009

Mekonnen, M. M., & Hoekstra, A. Y. (2010). A global and high-resolution assessment of the green, blue and grey water footprint of wheat. Hydrology and Earth System Sciences, 14(7), 1259-1276. https://doi.org/10.5194/hess-14-1259-2010

Naumann, G., Alfieri, L., Wyser, K., Mentaschi, L., Betts, R. A., Carrao, H., Spinoni, J., Vogt, J., & Feyen, L. (2018). Global changes in drought conditions under different levels of warming. Geophysical Research Letters, 45(7), 3285-3296. https://doi.org/10.1002/2017gl076521

Pagan, A., & Ullah, A. (1999). Nonparametric econometrics. Cambridge University Press. https://doi.org/10.1017/cbo9780511612503

Parker, S. Y., Parchment, K. F., & Gordon-Strachan, G. M. (2023). The burden of water insecurity: A review of the challenges to water resource management and connected health risks associated with water stress in small island developing states. Journal of Water and Climate Change, 14(12), 4404-4423. https://doi.org/10.2166/wcc.2023.239

Perrone, D., & Jasechko, S. (2019). Deeper well drilling an unsustainable stopgap to groundwater depletion. Nature Sustainability, 2(8), 773-782. https://doi.org/10.1038/s41893-019-0325-z

Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. Econometrica, 74(4), 967-1012. https://doi.org/10.1111/j.1468-0262.2006.00692.x

Pesaran, M. H. (2021). General diagnostic tests for cross-sectional dependence in panels. Empirical Economics, 60(1), 13-50. https://doi.org/10.1007/s00181-020-01875-7

Pokhrel, Y., Felfelani, F., Satoh, Y., Boulange, J., Burek, P., Gädeke, A., Gerten, D., Gosling, S. N., Grillakis, M., Gudmundsson, L., Hanasaki, N., Kim, H., Koutroulis, A., Liu, J., Papadimitriou, L., Schewe, J., Müller Schmied, H., Stacke, T., Telteu, C., Thiery, W., Veldkamp, T., Zhao, F., & Wada, Y. (2021). Global terrestrial water storage and drought severity under climate change. Nature Climate Change, 11(3), 226-233. https://doi.org/10.1038/s41558-020-00972-w

Racine, J. (1997). Consistent significance testing for nonparametric regression. Journal of Business and Economic Statistics, 15(3), 369-378. https://doi.org/10.1080/07350015.1997.10524714

Robinson, P. M. (1988). Root-N-consistent semiparametric regression. Econometrica, 56(4), 931-954. https://doi.org/10.2307/1912705

Rosa, L., Chiarelli, D. D., Rulli, M. C., Dell'Angelo, J., & D'Odorico, P. (2020). Global agricultural economic water scarcity. Science Advances, 6(18), eaaz6031. https://doi.org/10.1126/sciadv.aaz6031

Ryan, N., & Sudarshan, A. (2022). Rationing the commons. Journal of Political Economy, 130(1), 210-257. https://doi.org/10.1086/717045

Sachs, J. D., Schmidt-Traub, G., Mazzucato, M., Messner, D., Nakicenovic, N., & Rockström, J. (2019). Six transformations to achieve the Sustainable Development Goals. Nature Sustainability, 2(9), 805-814. https://doi.org/10.1038/s41893-019-0352-9

Sachs, J., Schmidt-Traub, G., Kroll, C., Lafortune, G., & Fuller, G. (2021). Sustainable Development Report 2020. Cambridge University Press. https://doi.org/10.1017/9781108992411

Sekhri, S. (2014). Wells, water, and welfare: The impact of access to groundwater on rural poverty and conflict. American Economic Journal: Applied Economics, 6(3), 76-102. https://doi.org/10.1257/app.6.3.76

Semenova, V., Goldman, M., Chernozhukov, V., & Taddy, M. (2023). Inference on heterogeneous treatment effects in high-dimensional dynamic panels under weak dependence. Quantitative Economics, 14(2), 471-510. https://doi.org/10.3982/QE1670

Sippel, S., Meinshausen, N., Fischer, E. M., Székely, E., & Knutti, R. (2020). Climate change now detectable from any single day of weather at global scale. Nature Climate Change, 10(1), 35-41. https://doi.org/10.1038/s41558-019-0666-7

Taylor, R. G., Scanlon, B., Döll, P., Rodell, M., Van Beek, R., Wada, Y., Longuevergne, L., Leblanc, M., Famiglietti, J. S., Edmunds, M., Konikow, L., Green, T. R., Chen, J., Taniguchi, M., Bierkens, M. F. P., MacDonald, A., Fan, Y., Maxwell, R. M., Yechieli, Y., Gurdak, J. J., Allen, D. M., Shamsudduha, M., Hiscock, K., Yeh, P. J.-F., Holman, I., & Treidel, H. (2013). Ground water and climate change. Nature Climate Change, 3(4), 322-329. https://doi.org/10.1038/nclimate1744

Veldkamp, T. I. E., Wada, Y., Aerts, J. C. J. H., Döll, P., Gosling, S. N., Liu, J., Masaki, Y., Oki, T., Ostberg, S., Pokhrel, Y., Satoh, Y., Kim, H., & Ward, P. J. (2017). Water scarcity hotspots travel downstream due to human interventions in the 20th and 21st century. Nature Communications, 8, 15697. https://doi.org/10.1038/ncomms15697

Vicente-Serrano, S. M., López-Moreno, J. I., Gimeno, L., Nieto, R., Morán-Tejeda, E., Lorenzo-Lacruz, J., Beguería, S., & Azorin-Molina, C. (2011). A multiscalar global evaluation of the impact of ENSO on droughts. Journal of Geophysical Research, 116(D20), D20109. https://doi.org/10.1029/2011jd016039

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica, 48(4), 817-838. https://doi.org/10.2307/1912934

Wisser, D., Frolking, S., Douglas, E. M., Fekete, B. M., Vörösmarty, C. J., & Schumann, A. H. (2008). Global irrigation water demand: Variability and uncertainties arising from agricultural and climate data sets. Geophysical Research Letters, 35(24), L24408. https://doi.org/10.1029/2008GL035296

Yatchew, A. (2003). Semiparametric regression for the applied econometrician. Cambridge University Press. https://doi.org/10.1017/cbo9780511615887

Zimmert, M., & Lechner, M. (2026). Nonparametric estimation of conditional average treatment effects under high-dimensional confounding. Econometrics Journal, 29(2), 290-319. https://doi.org/10.1093/ectj/utaf026

Declarations

Funding

The author declares that no funds, grants, or other support were received during the preparation of this manuscript.

Competing interests

The author has no competing interests to declare that are relevant to the content of this article.

Ethical approval

Not applicable. This study analyses publicly available country level secondary data and involves no human participants or animals, so no ethical approval was required.

Consent to participate

Not applicable. This study involves no human participants.

Consent to publish

Not applicable. This study contains no data relating to any identifiable individual person.

Clinical trial number

Not applicable.

Data availability

All data analysed in this study are publicly available. Water stress, income, population density, urbanisation, agricultural value added, and land area are drawn from the World Bank World Development Indicators, available at https://databank.worldbank.org/source/world-development-indicators, and the drought index is constructed by the author from the ERA5-Land reanalysis distributed by the European Centre for Medium-Range Weather Forecasts, available at https://www.ecmwf.int/en/era5-land. Both datasets are openly accessible and require no permission for reuse.

Author contributions

Muhammad Afnan Arif: Conceptualization, Methodology, Software, Formal analysis, Investigation, Resources, Data curation, Visualization, Writing – original draft.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- replicationpackagesdglandareav19.zip