WORLD RESOURCES INSTITUTE

Working Paper

AQUEDUCT GLOBAL MAPS 2.1: CONSTRUCTING DECISION-RELEVANT GLOBAL WATER RISK INDICATORS

FRANCIS GASSERT, MATT LUCK, MATT LANDIS, PAUL REIG, AND TIEN SHIAO

ABSTRACT

The availability of freshwater resources to meet human demands has emerged as a top-tier global issue for both environment and development. However, many decision-makers lack the technical expertise to fully understand hydrological information. In response to growing concerns from private sector actors around water availability, water quality, climate change, and increasing demand, we employed the composite index approach as a robust communication tool to translate hydrological data into intuitive indicators of water-related risks.

We grouped 12 indicators into a framework identifying spatial variation in water risks. For 6 of the 12 indicators, we used an ensemble of time series estimators, spatial regression, and a sparse hydrological model to generate novel datasets of water supply and use. We adapted the remaining six indicators from existing publications. We chose aggregation methods to maximize transparency and communicability, and to allow for dynamic weighting to reflect different users' sensitivities to water-related risks. We are currently unable to validate overall index scores because no datasets of water-related losses exist in the public domain. Data availability, specifically for major infrastructure (e.g., interbasin transfers) and *in-situ* water quality and river gauge measurements, is the primary constraining factor in model accuracy.

CONTENTS

Suggested Citation: Gassert, F., M. Luck, M. Landis, P. Reig, and T. Shiao. 2014. "Aqueduct Global Maps 2.1: Constructing Decision-Relevant Global Water Risk Indicators." Working Paper. Washington, DC: World Resources Institute. Available online at http://www.wri.org/publication/aqueduct-global-maps-21-indicators

The resulting *Aqueduct Water Risk Atlas* (Aqueduct) is a publicly available, global database and interactive tool that maps indicators of water-related risks. Aqueduct enables comparison across large geographies to identify regions or assets deserving of closer attention. This paper documents the methodology used to generate the hydrological metrics and indicators in the *Aqueduct Water Risk Atlas* (see Reig et al. 2013[1] for explanation of the Aqueduct framework and indicator selection). The data and maps are publicly available so that others may build off this effort.

1 INTRODUCTION

The availability and quality of freshwater to meet human needs has emerged as a top-tier global issue for environment and development,[2] and the number of people affected by water shortages has increased over time.[3] Risks associated with water availability are further compounded by uncertainties in the distribution of future climatic and rainfall patterns. High and unsustainable water use forces competition or compromises that may lead to conflict among users.[4] Water managers must balance the needs of growing populations, food irrigation requirements, and energy production, as well as threats to ecosystems and the services they provide. Furthermore, declining data collection efforts and gauging station networks in the world's rivers and lakes, and lack of consistent global metrics limit our ability to adequately address this problem.[5] In an effort to fill this gap, the World Resources Institute's Aqueduct Water Risk Atlas (Aqueduct) compiles advancements in hydrological modeling, remotely sensed data, and published datasets into a freely accessible online platform.

Aqueduct is a publicly available, global database and interactive tool that maps indicators of water-related risks for decisionmakers worldwide. Aqueduct uses a combination of geospatial and statistical models to translate global hydrological data into straightforward indicators and aggregated scores that can inform a broad range of corporate, governmental, and civil society users. We grouped 12 global indicators into a framework designed in response to the growing concerns from private and public sector actors around water scarcity, water quality, climate change, and increasing demand for freshwater. These indicators are intended for comparative analysis across large geographies to identify regions or assets deserving of closer attention, and are not appropriate for local or site-specific analyses. This tool and the underlying data and methodology are open and available to serve public interests.

The 12 indicators in the Aqueduct framework were selected in three steps. First, we conducted a literature review of relevant water issues, existing water indicators, and data sources. Then we evaluated potential data sources through a comparative analysis of their spatial and temporal coverage, granularity, relevance to water users, consistency, and credibility of sources. Final selection took place in consultation with industry, public-sector, and academic water experts. We sought to select indicators that covered the full breadth of water-related risks, while minimizing overlap and potential confusion resulting from an overabundance of indicators. A complete description of the framework and indicator selection can be found in Reig et al. (2013).[6]

Six of Aqueduct's 12 indicators were developed by the World Resources Institute (WRI) and ISciences LLC, using a combination of publicly available datasets and modeling techniques. Sections 2 and 3 describe the modeling techniques and six resulting indicators, respectively. The remaining six indicators were adapted from existing published sources and are described in Section 4. Section 5 describes the indicator scoring and aggregation methodology. Finally, in Section 6 we discuss the results and implications of this work.

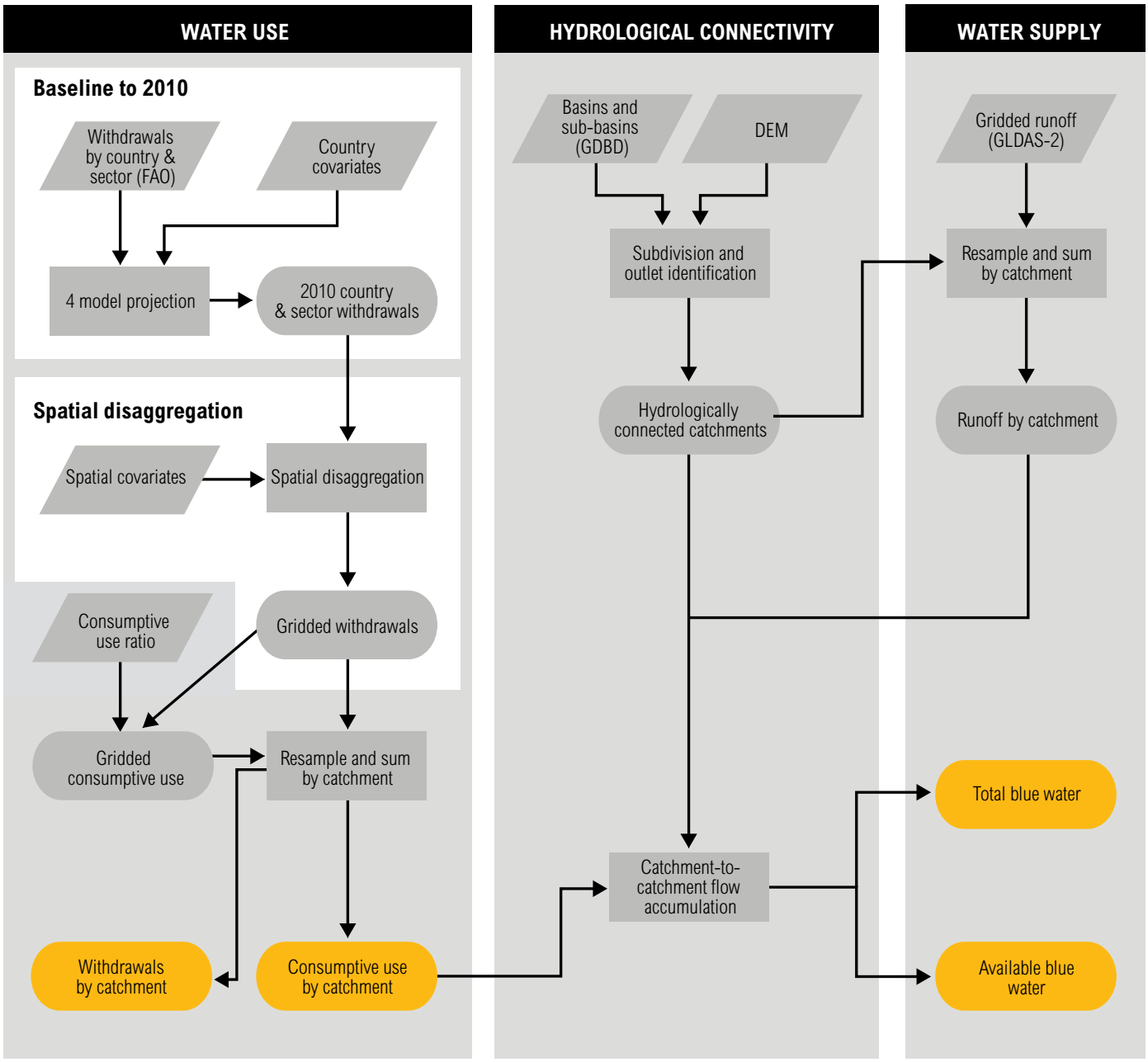
2 WATER USE AND SUPPLY MODELING

To create six of Aqueduct's 12 indicators, we generated two metrics of spatially explicit water use and two metrics of spatially explicit supply (Figure 1). Use and supply are estimated at a hydrological catchment scale, which is required to spatially partition runoff and withdrawals and accumulate values downstream.

2.1 Catchments

For this exercise, we define *catchments* as areas of land that drain to a single outlet point. Catchments were extracted from the Global Drainage Basin Database (GDBD),[7] which comprises 11,476 complete river basins[8] more finely resolved into 73,074 sub-basin polygons attributed with basin and Pfafstetter[9] codes. To extract consistently sized catchments relative to the scale of our other variables, we subdivided the 11,476 complete GDBD basins using a recursive algorithm that splits basins above a threshold 100,000 km_2 into the next lower level of sub-basin polygons by Pfafstetter code. Once completed, we found the outlet of each resulting catchment by using the GDBD digital elevation model (DEM) to identify the lowest point in the catchment and the adjacent downstream

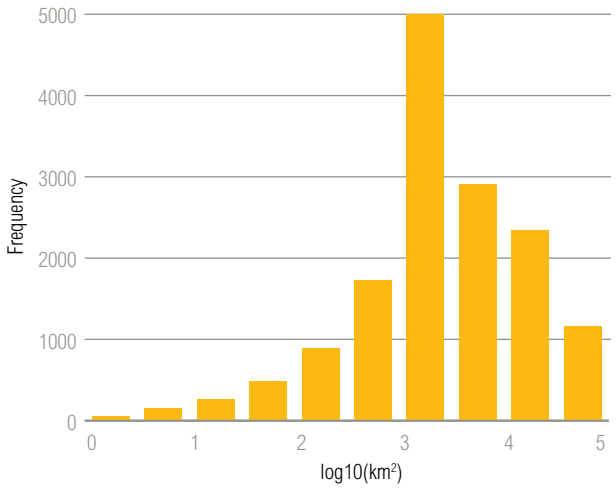
Figure 1 | **Water Supply and Use Model Schematic**



Note: Parallelograms are inputs and rounded shapes are outputs. The two final catchment-scale water-use metrics and two water-supply metrics are highlighted in yellow.

Figure 2 | **Aqueduct Catchments**

a. Distribution of Catchment Size



b. Global Map of Catchments

Source: WRI Aqueduct. Catchments extracted from Matsutomi et al. 2009.

catchment or sink. If we found multiple outlets for one catchment, we selected a single outlet at random. This situation occurs in GDBD because coastal interbasins, or intervening zones, were often combined forming a single coastal catchment between larger basins. There were also a few instances where GDBD aggregated delta areas of large river mouths that resulted in disjunct multi-part polygons; these were manually split and treated as separate catchments.

Splitting of the GDBD basins produced 15,006 catchments with a median and mean size of 2,282 and 8,804 km², respectively. Basins along the coastlines and numerous inland sinks tend to be small while upland catchments are larger, leading to a distribution of basin size that is highly skewed to the right, with many smaller basins than larger ones (Figure 2).

2.2 Water Use

We estimated two metrics of spatially explicit water use: *water withdrawal* (the total amount of water abstracted from freshwater sources for agricultural, domestic, and industrial uses), and *consumptive use* (the portion of withdrawn water that evaporates or is incorporated into a product and is thus no longer available for downstream use). A third metric, *non-consumptive use*, was derived as withdrawal minus consumptive use.

2.2.1 Water Withdrawal

We used withdrawal data from the Food and Agriculture Organization of the United Nations (FAO) Aquastat database,[10] which contains water use data collected from national governments as well as FAO's internal estimates, or from Gleick et al.[11] where FAO data was not available. Aquastat reports annual water use by country and sector (agriculture, domestic, and industrial, such that a country's total annual water use is the sum of the three sectors) though reporting is inconsistent and varies by country.[12] The most recent year for which withdrawal data were available was 2010; however, because of inconsistencies in reporting, not all countries had data for this year. As detailed in the following subsections, we first projected the water withdrawals to a common year (2010) and, in order to use these values at the catchment scale, spatially disaggregated values from countries and re-aggregated them to catchments.

Table 1 | **Withdrawal Projection Variables by Sector**

	VARIABLE	ABBREVIATION	SOURCE
RESPONSE VARIABLE	log(Water withdrawals)	Y	FAO Aquastat[a] Pacific Institute[b]
EXPLANATORY VARIABLES FOR ALL SECTORS	log(GDP)	GDP	World Bank[c]
	log(Population)	POP	World Bank
	Average annual precipitation	PRCP	FAO Aquastat
	Total renewable water supply	RWS	FAO Aquastat
	Sectoral water withdrawal ratio	WWR	Authors' calculation (Y – log(RWS))
AGRICULTURAL ONLY	log(Area Equipped for irrigation)	IGAREA	FAOSTAT[d] FAO Aquastat Freydank and Siebert[e]
	log(Agricultural land area)	AGR	World Bank
INDUSTRIAL ONLY	log(CO_2 emissions)	CO_2	World Bank
	log(Electricity, net generation)	ELEC	Energy Information Administration[f]
	log(Coal production)	COAL	Energy Information Administration
DOMESTIC ONLY	Urban population (%)	URB	World Bank

Notes:

a. Food and Agriculture Organization of the United Nations, "AQUASTAT - FAO's Information System on Water and Agriculture," accessed June 05, 2013, http://www.fao.org/nr/water/aquastat/main/index.stm.

b. Peter Gleick et al., *The World's Water Volume 7* (Washington, DC: Island Press, 2011), http://worldwater.org/data.html.

c. GDP reported in constant 2000 US$ at purchasing parity power. World Bank, "World Development Indicators," 2012, http://data.worldbank.org/data-catalog/world-development-indicators.

d. Food and Agriculture Organization of the United Nations, "FAOSTAT," 2012, http://faostat3.fao.org/faostat-gateway/go/to/home/E.

e. Katharina Freydank and Stefan Siebert, *Towards Mapping the Extent of Irrigation in the Last Century: Time Series of Irrigated Area Per Country*, 2008, http://publikationen.ub.uni-frankfurt.de/frontdoor/index/index/docId/5916.

f. U.S. Energy Information Administration (EIA), "International Energy Statistics," 2012, http://www.eia.gov/cfapps/ipdbproject/IEDIndex3.cfm.

2.2.1.1 Baseline Withdrawal 2010

For countries with withdrawal data reported in Aquastat for 2008 or later (24 of 178 countries with sufficient data available), we used the reported value as the 2010 estimate. For countries without these data, we projected withdrawals for a baseline year of 2010 using equations that relate water use to country-level variables for which there were data in 2010 (see Table 1 for a list of the variables). These relationships were estimated using regression models. Because there is no clear consensus on the best regression modeling approach for cross-section time-series data such as these, projections were based on a combination of two regression models. Specifically, we used a mixed-effects

(or multilevel) model,[13] and a fixed-effects model using the within-transform.[14] Both models account for changes over time within a country, but they differ in the details of how the models are fit.

Table 1 lists response and explanatory variables for each sector. The response variable, water withdrawals, is based on data from FAO Aquastat or from Gleick et al. (1–3 countries depending on sector). Explanatory variables were modified by filling temporal gaps in the data (less than 1 percent of values), using single imputation *via* loess regression[15] and simple linear regression. For country-variable combinations that are missing entirely (e.g., Myanmar GDP), we used multiple imputation based on the Amelia package in *R*,[16] filling between 1 percent and 2 percent of country-variable combinations depending on the sector.

The two regression models were as follows. The mixed-effects estimator (henceforth ME) predicts water withdrawals (Y) for country, i, and year, t, as a function of k explanatory variables (X), with coefficients (β) as well as separate intercepts (b_o) and slopes (on the *YEAR* variable; b_{YEAR}) for each country.

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \cdots + \beta_k X_{kit} + b_{oi} + b_{YEAR,i} YEAR + \varepsilon_{it}$$

This approach explicitly partitions variation in water withdrawals among between-country variation (the variation in the intercepts and slopes) and within-country variation (the residual variation). Considering the United States as an example, this model chooses variables and coefficients to account for both the country's high agricultural withdrawals relative to other countries, as well as its observed increase in withdrawals over time. Because this model estimates the change in withdrawals as well as the average level of withdrawals for a particular country, it can be used to estimate withdrawals even for countries that had no observed withdrawal data (as long as they have all the covariates). These models were fit using maximum likelihood with the R package *nlme*.[17]

We used a second estimator, a fixed-effects approach model, using the within-transform (henceforth FE). The within-transform explicitly removes between-country variation by subtracting the country mean across all years from each variable:

$$(Y_{it} - \bar{Y}_i) = \beta_1 (X_{1it} - \bar{X}_{1i}) + \beta_2 (X_{2it} - \bar{X}_{2i}) + \cdots + \beta_k (X_{kit} - \bar{X}_{ki}) + (\varepsilon_{it} - \bar{\varepsilon}_i)$$

While the ME approach can explain a high proportion of variation in water use (see below), one disadvantage is that the variable selection process can be dominated by variables that primarily explain between-country variation. In this application, modeling within-country variation is more relevant. We use the FE approach to force the variable selection process to focus only on explaining within-country variation. One shortcoming is that the FE approach cannot be used to estimate 2010 withdrawals for countries that have no reported withdrawal data, since the mean withdrawals are unknown and cannot be added back to the predicted yearly deviation. This applies to 5–9 countries, depending on sector, out of 178. For those countries, we used the ME estimators exclusively for prediction. Both approaches have strengths and weaknesses, and neither proved superior. Therefore, we used estimates derived from both models.

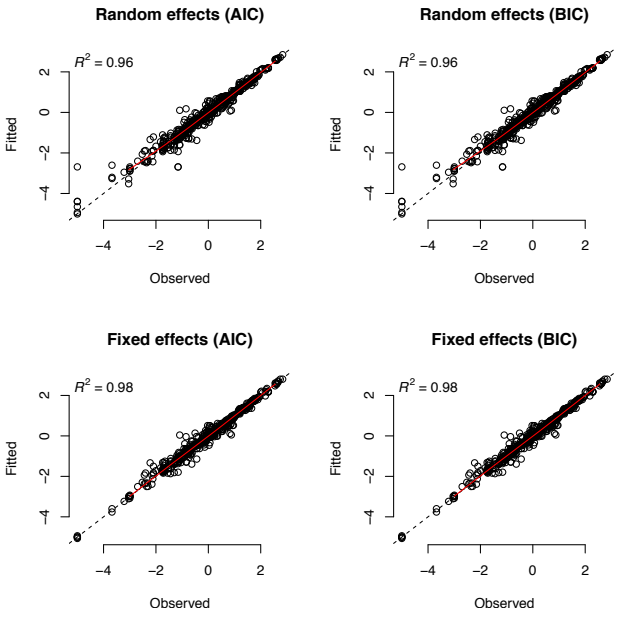
With each modeling approach, we fit a full model containing all predictors, polynomial terms on key variables such as GDP, and selected interactions among up to three variables. We then reduced each model using backward elimination of variables by two different goodness-of-fit criteria: Akaike's Information Criterion (*AIC*)[18] and the Bayesian Information Criterion (*BIC*).[19] These criteria measure goodness-of-fit of a model penalized by the number of parameters used; *BIC* penalizes additional variables more heavily, thus models selected by *BIC* tend to have fewer variables included. Although *AIC* is more widely used, some authors prefer *BIC*. Both *AIC* and *BIC* were used with each model (ME and FE) resulting in a total of four withdrawal estimates for each country.

Overall, models predicted total variation across years and countries well, with total R^2 between 0.94 and 0.99 (Table 2; Figure 3). We used the average of the predicted 2010 value from each of the resulting four models (ME-AIC, ME-BIC, FE-AIC, FE-BIC) as the final 2010 estimate for each sector.

Table 2 | **Proportion of Variation Explained for Each Sector by Modeling Approach**

SECTOR	MODEL	TOTAL VARIATION	BETWEEN VARIATION	WITHIN VARIATION
Agricultural	ME – AIC	0.96	0.97	0.13
	ME – BIC	0.96	0.97	0.12
	FE – AIC	0.98	—	0.09
	FE – BIC	0.98	—	0.07
Industrial	ME – both	0.96	0.83	0.55
	FE – AIC	0.94	—	0.16
	FE – BIC	0.94	—	0.15
Domestic	ME – AIC	0.99	0.93	0.72
	ME – BIC	0.99	0.92	0.78
	FE – both	0.98	—	0.49

Figure 3 | **Observed vs. Modeled Values for Four Models of Agricultural Water Withdrawals**



Note: Scale in log[km^3]. The dashed black line indicates the 1:1 line, and the red line indicates a loess smoother.

2.2.1.2 Spatial Disaggregation

Once we attained temporally consistent country and sector specific withdrawal estimates, we downscaled these estimates from countries to catchments on a sector-by-sector basis. For each sector, we selected *k* spatially explicit gridded datasets representative of that sector to maximize predictive power over country withdrawals (Table 3). We summed these gridded datasets from pixels (p) to countries (i), and fit zero-intercept regression estimators of 2010 withdrawals ($\hat{Y}$) (Table 4).

$$\hat{X}_i = \sum_{p \in i} X_p$$

$$\hat{Y}_i = \beta_1 \hat{X}_{1i} + \beta_2 \hat{X}_{2i} + \cdots + \beta_k \hat{X}_{ki} + \varepsilon$$

Table 3 | **Spatial Disaggregation Explanatory Variables by Sector**

SECTOR	VARIABLE	ABBREVIATION	RESOLUTION	SOURCE
Agricultural	Global Map of Irrigation Areas	GMIA	5 arc-min	Siebert et al.[a]
Industrial	Nighttime Lights 2010	NTL	30 arc-sec	NOAA[b]
Domestic	Nighttime Lights 2010	NTL	30 arc-sec	NOAA
	Gridded Population of the World 2010	GPW	2.5 arc-min	CIESIN[c]
	Interaction	NTL×GPW		

Notes:

a. Stefan Siebert et al., "Global Map of Irrigation Areas" (Johann Wolfgang Goethe University, Frankfurt am Main, Germany / Food and Agriculture Organization of the United Nations, Rome, Italy, 2007), http://www.fao.org/nr/water/aquastat/irrigationmap/index.stm.

b. National Oceanographic and Atmospheric Administration (NOAA) and National Geophysical Data Center (NGDC), "Version 4 DMSP-OLS Nighttime Lights Time Series," 2010, http://www.ngdc.noaa.gov/dmsp/downloadV4composites.html.

c. Center for International Earth Science Information Network (CIESIN), Columbia University, United Nations Food and Agriculture Organization (FAO), and Centro Internacional de Agricultura Tropical (CAIT), "Gridded Population of the World Version 3 (GPWv3): Population Count Grid, Future Estimates," 2005, http://sedac.ciesin.columbia.edu/gpw.

Table 4 | **Proportion of Country Withdrawal Variation Explained by Sector**

SECTOR	$R^2(p)$
Agricultural	0.93 (<0.01)
Industrial	0.84 (<0.01)
Domestic	0.95 (<0.01)

We used the fitted country regression coefficients (β) to project withdrawals over each pixel (P_p) and normalized the projected withdrawals by dividing by the sum of projected withdrawals for all pixels within the containing country (i) to estimate withdrawals within each pixel (U_p) for each sector.

$$P_p = \beta_1 X_{1p} + \beta_2 X_{2p} + \cdots + \beta_k X_{kp}$$

$$U_p = \frac{P_p}{\sum_{q \in (i \mid i \ni p)} P_q}$$

This method guarantees that the total withdrawal for each country is conserved, that is, the sum of withdrawals for all pixels within a country is equal to the country withdrawal. Withdrawals for each catchment (j) are then the sum of the estimated withdrawals for each pixel in the catchment.

$$U_j = \sum_{p \in j} U_p$$

Finally, we summed the three sectors (agricultural, domestic, and industrial) to create total withdrawal.

2.2.2 Consumptive Use

We estimated consumptive use by multiplying water withdrawals by estimates of the portion of withdrawn water that is consumed per sector (agricultural, domestic, and industrial). Shiklomanov and Rodda estimate withdrawals (U_r) and consumptive use (C_r) for each sector in their 26 "natural-economic" regions of the world.[20] From this, we calculated consumptive use ratios of C_r / U_r, and gridded the 26 region polygons to match the pixel size of the

sectoral water withdrawals. We then multiplied the gridded ratios by gridded withdrawals and summed by catchment to get consumptive use by sector per catchment (C_j).

$$C_p = U_p \left(\frac{C_r}{U_r}\right); C_j = \Sigma_{p \in j} C_p$$

The total consumptive use in each catchment is the sum of consumptive use for each sector. We derived non-consumptive use by subtracting consumptive use from withdrawals. Withdrawals and especially consumptive use are largely dominated by irrigated agriculture (Figure 4).

2.3 Water Supply

We computed two metrics of annual water supply: *total blue water* (*Bt*) and *available blue water* (*Ba*). Total blue water approximates naturalized river discharge whereas available blue water is an estimate of surface water availability that removes water that is consumed upstream.

2.3.1 Runoff Model

We computed water supply from runoff (R), which is the amount of water that falls as precipitation (P) and is available to flow across the landscape after evapotranspiration (Et) and changes in soil moisture storage (ΔS) are accounted for (i.e., $R = P - Et - \Delta S$). We extracted monthly runoff from the Global Land Data Assimilation System Version 2.0 1°×1° land surface model (GLDAS-2.0, 1948–2010 monthly),[21] which uses the Noah v.3.3 land surface model and Princeton climate-forcing data to initialize meteorological variables such as radiation, temperature, and precipitation to observed conditions. Runoff from Noah treats *Et* from all cropland as dryland (nonirrigated). Therefore, *Et* from rain-fed agricultural is accounted for as an input to runoff, and any water withdrawn for irrigating crops is in excess of *Et*.

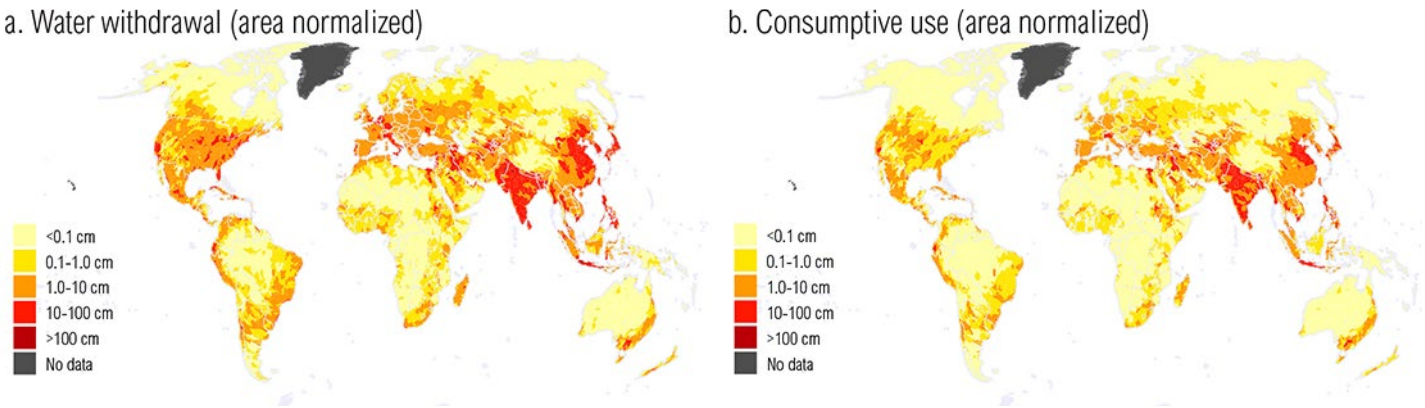
We selected this dataset after evaluating four models for accuracy, spatial resolution, and time coverage (see Section 8.2 for details). GLDAS-2 has been validated against gauged discharge and found to yield acceptable results for annual discharge (least bias), timing of peak flow, and intra- and inter-annual variability, and accurately simulated discharge in high-latitude, mid-latitude, and tropical rivers. For example, the percent error between global mean annual discharge between observed Global Runoff Data Centre (GRDC) and simulated by Noah (v.2.7.1) over a 63 year period was 3.5 percent.[22]

We calculated monthly runoff by catchment (R_j) by summing the subsurface (baseflow or lateral flows of shallow groundwater) and surface runoff components of the Noah land-surface model. Annual runoff is the sum of the runoff of 12 calendar months. We omitted the first two years (1948–49) to minimize error that might be introduced during the initial spinup period.

2.3.2 Flow Accumulation

We estimated water supply using a catchment-to-catchment flow accumulation approach, which aggregates water by catchment and transports it to the next downstream

Figure 4 | **Area-Normalized Total Withdrawals and Consumptive Use, 2010 Estimates**



Source: WRI Aqueduct.

catchment. For the calculation of our indicators, this catchment-to-catchment flow accumulation method has two advantages over conventional pixel-to-pixel flow accumulation. First, it is better suited to our disaggregation of withdrawals data because it relaxes the requirement that withdrawals be located with spatial precision to the pixel level. Second, aggregating runoff by catchment allows water to be freely transported within a catchment and eliminates the problem that use or demand in one pixel cannot be supplied from a river in an adjacent pixel. This method assumes that all water in a catchment is effectively available to all users, and that no water from outside a catchment is available to the catchment. Since the catchment-to-catchment flow accumulation method does not include a dynamic temporal element (the water flows completely through the system from headwater to mouth in each time step without retention), it is only appropriate for time steps longer than it takes flow to transit a catchment. Given the relatively small spatial scale of our catchments and the monthly and yearly time steps used here, this is a valid assumption across most of the world except for areas below dams with large reservoir capacity.

We computed total blue water (*Bt*) for each catchment, *j*, by recursively summing runoff (*R*) from adjacent upstream catchments (the catchments that have *j* as their downstream outlet):

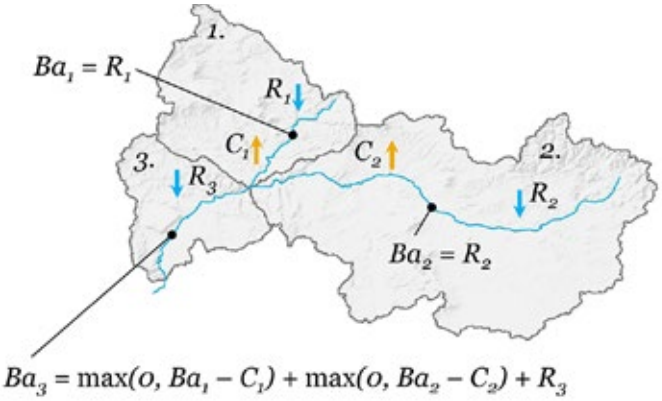
$$Bt_j = \begin{cases} R_j + \sum_{(u|\text{outlet}(u)=j)} Bt_u, \{u|\text{outlet}(u)=j\} \neq \emptyset \\ R_j, else \end{cases}$$

First-order catchments are defined as those without upstream catchments, so *Bt* in such catchments is equal to runoff. This equation was also used to flow accumulate other input values, for example, storage capacity and use, for several indicators (Section 3).

Available blue water (*Ba*) denotes the total amount of water available to a catchment accounting for upstream consumptive uses. We computed *Ba* as runoff plus all water flowing into the catchment from adjacent upstream catchments where consumptive use (*C*) is removed in upstream catchments prior to being counted (Figure 5):

$$Ba_j = \begin{cases} R_j + \sum_{(u|\text{outlet}(u)=j)} max\,(0, Ba_u - C_u), \{u|\text{outlet}(u)=j\} \neq \emptyset \\ R_j, else \end{cases}$$

Figure 5 | **Schematic Diagram of Catchment-to-Catchment Flow Accumulation**



Note: Available blue water (*Ba*) in catchment 3 is equal to runoff plus the sum of *Ba* minus consumptive use (*C*) for the adjacent upstream catchments. Since catchments 1 and 2 have no upstream catchments, *Ba* is equal to runoff. Note that consumption is not counted against *Ba* in the current catchment but is counted in catchments further downstream.

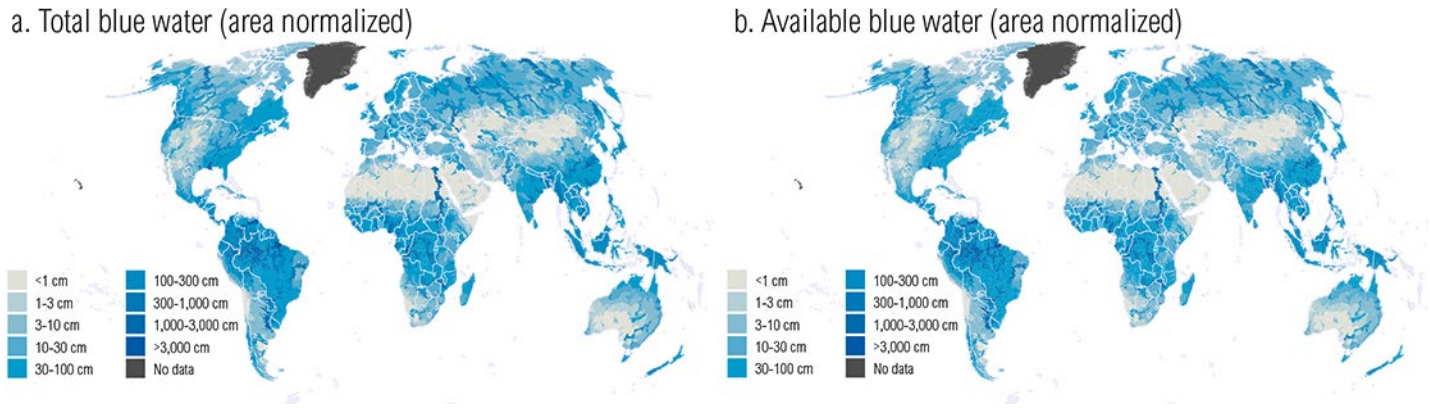
First-order catchments have no upstream catchments, so *Ba* is simply runoff. In the event that consumptive use exceeds *Ba*, we assumed that any excess use is satisfied by alternative water sources (e.g. groundwater storage) and set the volume of water to be transferred to zero.

Although both *Ba* and *Bt* primarily measure surface water supply, the Noah land surface model also estimates shallow subsurface flow that may eventually reemerge on the surface (e.g., as springs) or contribute to the baseflow of rivers.[23] Changes in existing water stores (e.g., deep groundwater and meltwater from glaciers) are not accounted for in our water supply model. Interbasin transfers of water and desalination are also unaccounted for because of insufficient global data, which may lead to underestimation of water supply particularly in coastal regions. The spatial pattern of *Ba* is very similar to that of *Bt* (Figure 6), though there are noticeable differences in regions with high consumptive use.

3 MODELED USE AND SUPPLY INDICATORS

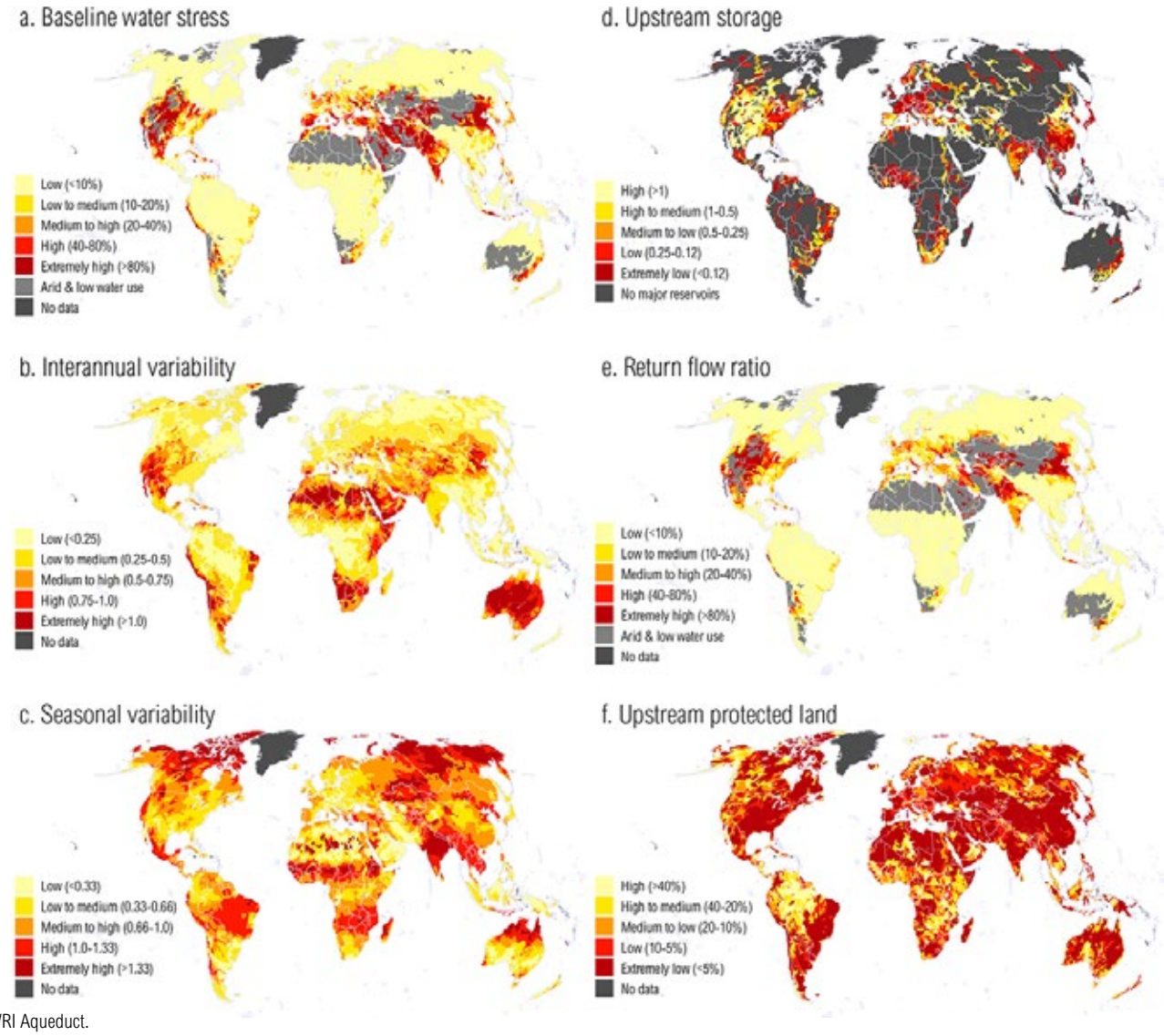
From the modeled use and supply datasets we derived six indicators (Figure 7) that are described below.

Figure 6 | **Area-Normalized Mean Annual Total and Available Blue Water, 1950–2008**



Source: WRI Aqueduct.

Figure 7 | **Modeled Use and Supply Indicators**



Source: WRI Aqueduct.

3.1 Baseline Water Stress

Baseline water stress (*BWS*) measures the ratio of total annual water withdrawal (*Ut*) to average annual available blue water (*Ba*), a commonly used indicator also known as relative water demand.[24] It is important to note that most estimates of relative water demand do not account for upstream consumptive use as we do here. We used a long time series of supply (1950–2010) to reduce the effect of multi-year climate cycles and to allow us to ignore complexities of short-term water storage (e.g., dams, floodplains) for which global operational data is nonexistent.[25] Baseline water stress thus measures chronic stress rather than drought stress.

$$r_{BWS} = \frac{Ut_{2010}}{\text{mean}_{[1950,2010]}(Ba)}$$

We masked catchments with less than 0.012 m/m²/year of withdrawal and 0.03 m/m²/year of available blue water as "arid and low water use" since catchments with low values were more prone to error in our estimates of baseline water stress. Additionally, although current use in such catchments is low, any new withdrawals could easily push them into higher stress categories. We scored these areas as maximum risk for the purposes of aggregation (Section 5).

3.2 Inter-Annual Variability

Inter-annual variability (IAV) measures the variation in natural water supply between years and is the catchment-specific coefficient of variation of *Bt*, calculated as the standard deviation of total blue water (*Bt*) divided by the mean of total blue water.

$$r_{IAV} = \frac{\text{sd}_{[1950,2010]}(Bt)}{\text{mean}_{[1950,2010]}(Bt)}$$

Inter-annual variability ignores human influences such as diversions and infrastructure, and focuses on natural variation in surface water supply. This overestimates perceived variability in highly modified river basins.

3.3 Seasonal variability

Seasonal variability (*SV*) estimates within-year variation of water supply. We computed mean total blue water for each of the calendar months (Bt_m), and divided the standard deviation of the 12 monthly values by the overall mean monthly total blue water.

$$\overline{Bt}_m = \text{mean}_{[1950,2010]}(Bt_{i,m}),\ \text{m} \in \{jan...dec\}$$

$$r_{SV} = \frac{\text{sd}_{[jan...dec]}(\overline{Bt}_m)}{\text{mean}_{[jan...dec]}(\overline{Bt}_m)}$$

Seasonal variability ignores human influences such as diversions and infrastructure, but instead attempts to measure natural variation in surface water supply. While the runoff model captures short-term soil moisture storage, other forms of natural seasonal storage such as floodplains and wetlands are not captured and may lead to error in seasonal variability estimates.

3.4 Upstream Storage

Upstream storage (*STOR*) measures how many years of total blue water storage capacity exist upstream of or within the given catchment. We extracted storage capacity from the Global Reservoir and Dam (GRanD) database, which attempts to capture all large dams and reservoirs (those with capacity greater than 0.1 km³ across the world).[26] Since reporting of dams and reservoirs varies by region, this indicator has greater influence in areas with more complete datasets of reservoirs and dams (e.g., the

United States). We accumulated upstream and within-basin storage capacity (*US*) in the same way we computed Bt, then divided by mean total blue water. This indicator is intended to counteract overestimates of variability in highly modified basins. We considered areas that are not covered by GRanD as having insufficient data. These catchments are coded as "no major dams" and excluded from scoring and aggregation.

$$r_{STOR} = \frac{US}{\text{mean}_{[1950,2010]}(Bt)}$$

Higher upstream storage indicates that a basin can withstand longer periods of deficits in available water with reservoir storage and also implies that risk of flooding may be lower. We did not attempt to account for the negative impacts of reservoir storage such as increased evaporation and changes to downstream flow regime.

3.5 Return Flow Ratio

Return flow ratio (RFR) measures the ratio of non-consumptive use upstream and within the given catchment relative to the mean available blue water. We accumulated upstream and within-basin non-consumptive use (*URF*) in the same way we computed Bt. We then divided by mean available blue water.

$$r_{RFR} = \frac{URF}{\text{mean}_{[1950,2010]}(Ba)}$$

This estimates the percent of available blue water that has been previously used and discharged, and it indicates reliance on water treatment infrastructure and natural features such as buffers and wetlands to maintain water quality, based on the assumption that water use degrades downstream water quality. We did not attempt to capture the effects of human or natural water treatment, nor did we differentiate between types of water uses to estimate pollutant loads.

3.6 Upstream Protected Land

Upstream protected land (PROT) measures the proportion of total blue water that originated from protected areas. We extracted protected areas from the World Database on Protected Areas,[27] excluding IUCN category V protected areas,[28] as well as a large number of unclassified lands, breeding centers, municipal parks, cultural and historic sites, and exclusively marine areas, as these areas do not explicitly protect ecological services. We estimated mean total blue water using only runoff originating in protected lands (*PBt*) in the same way we computed *Bt*. We then divided by mean total blue water to derive indicator values for percentage of water sourced from upstream protected land.

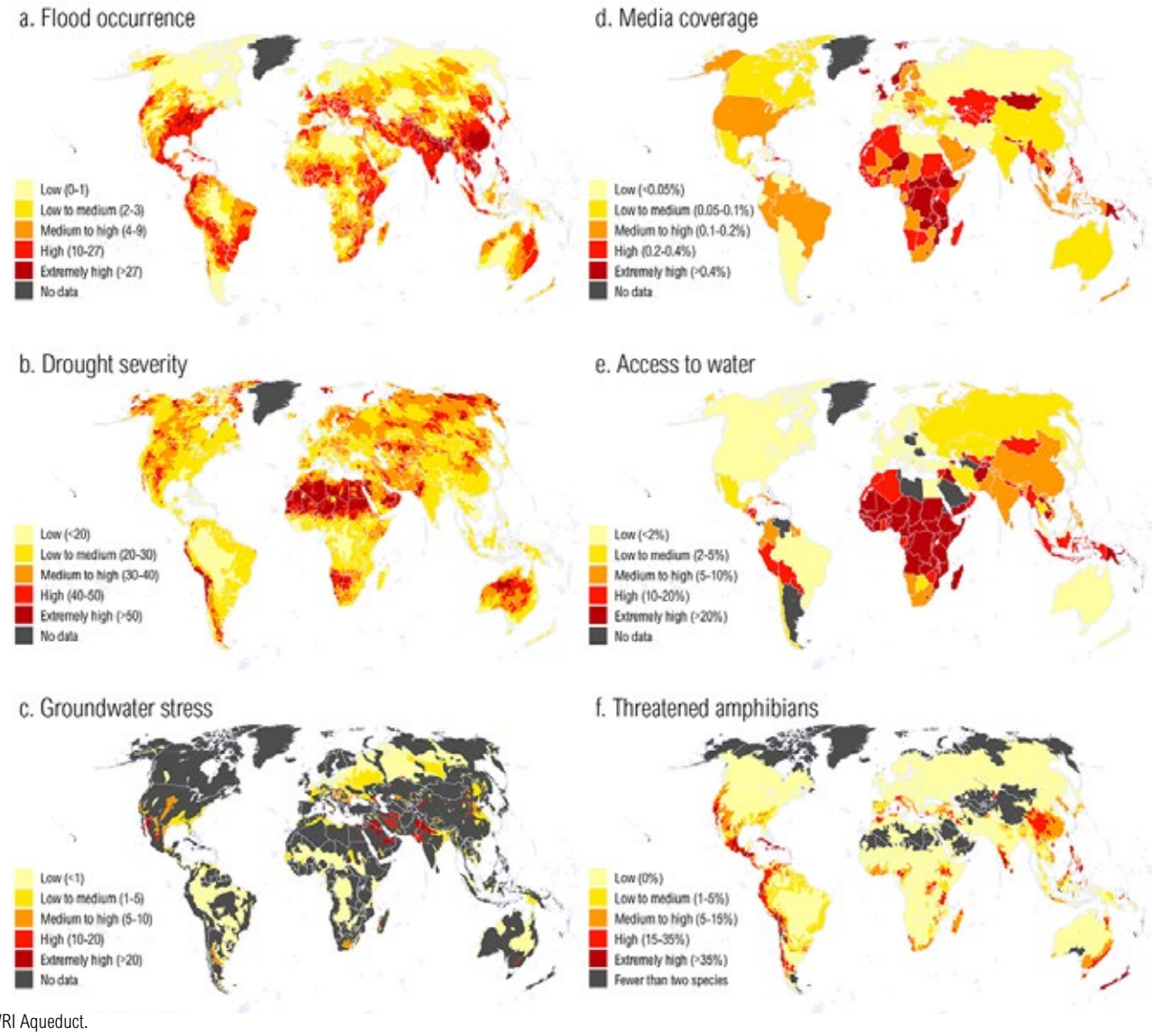
$$r_{PROT} = \frac{\text{mean}_{[1950,2010]}(PBt)}{\text{mean}_{[1950,2010]}(Bt)}$$

Upstream protected land does not attempt to quantify the effectiveness of the protected area or account for external pressures on the watershed, which may have an outsize impact on the health of freshwater ecosystems.[29]

4 EXTERNALLY SOURCED INDICATORS

The remaining 6 of the 12 indicators used in Aqueduct were adapted from external sources. Of these 6, (Figure 8), 4 were generated with minimal processing, and the remaining 2 are completely unmodified from their published form.

Figure 8 | **Externally Sourced Indicators**



Source: WRI Aqueduct.

4.1 Flood Occurrence

Flood occurrence (FO) measures the number of floods recorded in each catchment between 1985 and 2011. Reported flood extent polygons were taken from the Global Flood Observatory.[30] Polygons are estimated from remote sensing, governmental, and media reports of affected regions. Extent polygons (E) are then spatially joined to catchments (j) to count the total number of floods that may have affected each catchment over the recorded period.

$$r_{FO,j} = \text{count}\left(\{E | E \cap j \neq \emptyset\}\right)$$

The flood occurrence indicator differs from the preceding hydrological indicators in several ways. First, floods are extreme events that are not captured in our long-term measurements. Second, this indicator counts actual observations rather than modeled occurrence, reflecting the influence of flood control infrastructure as well as events such as flash floods and coastal flooding, which are not easily captured by current hydrological models. Finally, this dataset's method of flood extent estimation ignores local typography and may substantially overestimate the extent of flood impact.

4.2 Drought Severity

Drought severity (DRO) measures the mean severity of drought events from 1901 to 2008 as recorded in a modeled 1° × 1° gridded data set by Sheffield and Wood.[31] To produce this dataset, they generated a monthly soil moisture hydrograph for each grid cell, and defined drought runs as continuous periods in which soil moisture falls under the 20th percentile of the monthly hydrograph ($q(\Theta)$<20%). Severity (S) of a drought run beginning at time, t_i, is the length (D) times the intensity (I) of the drought, with the length measured in months and intensity equal to the average number of points by which soil moisture falls beneath the 20th percentile.

$$S = \sum_{t=t_i}^{t+D-1} 20\% - q(\Theta)_t \qquad i.e.\ S = I \times D$$

We resampled the gridded mean severity dataset and averaged it across our hydrological catchments.

$$r_{DRO,j} = \sum_{p \subset j} \text{mean}(S)_p$$

By definition, all regions of the world experience some form of drought 20 percent of the time. The drought severity indicator emphasizes those regions where soil moisture deficits are longer and drier, thereby making them harder to adapt to and mitigate. Regions that experience decadal or multidecadal variations in precipitation are more likely to fall into this category.

4.3 Groundwater Stress

Groundwater stress (GW) measures the ratio of groundwater withdrawal relative to its sustainable recharge rate over a given aquifer. Gleeson et al. define groundwater footprint (GF) "as $A[C/(R - E)]$ where C, R, and E are respectively the area-averaged annual abstraction of groundwater, recharge rate, and the groundwater contribution to environmental stream flow," estimated at a 1° gridded resolution, and A "is the areal extent of any region of interest where C, R, and E can be defined."[32] Groundwater stress is groundwater footprint divided by aquifer area.

$$r_{GW} = \frac{GF}{A}$$

Gleeson et al. only report values for major known aquifers, excluding areas with local, shallow, or complex groundwater systems, as well as areas with a recharge rate less than 2 mm/year. We excluded areas without reported values from scoring and aggregation. Groundwater stress estimates the sustainability of groundwater withdrawals but cannot predict the total volume of water stored in aquifers.

4.4 Media Coverage

Media coverage (MC) measures the number of news articles about water issues in a country relative to the total number of articles about the country. We used Google News to search a string of keywords including "water shortage" or "water pollution," and a country—for example, "water shortage + Egypt"—limiting results to the 10 years from January 1, 2002 to December 31, 2011.[33] For each country, we summed the number of articles on water shortage and water pollution (W) and divided by the total number of articles on any topic found when searching for only the country name (T).

$$r_{MC} = \frac{W}{T}$$

Media coverage estimates the relative importance of water issues in the English-language media. Areas that have been dominated by conflict (e.g., Iraq) or other high-profile issues (e.g., Venezuela) appear lower in value despite apparent water issues.

4.5 Access to Water

Access to water (WC) measures the proportion of population without access to improved drinking water sources by country and estimates the coverage of drinking water infrastructure.

$$r_{WC} = \frac{P_{TOTAL} - P_{ACCESS}}{P_{TOTAL}}$$

This indicator is taken from the WHO/UNICEF Joint Monitoring Programme (JMP) for Water Supply and Sanitation 2010 dataset, which defines an improved drinking water source as "one that, by nature of its construction or through active intervention, is protected from outside contamination."[34] The JMP aggregates this dataset from nationwide health surveys.

4.6 Threatened Amphibians

Threatened amphibians (AMPH) measures the percentage of amphibian species classified by the International Union for Conservation of Nature (IUCN) as threatened in each catchment. We joined the IUCN Red List of Threatened Species for amphibians with an IUCN database on estimated amphibian species ranges, making several name corrections in joining the data. We then spatially joined the estimated ranges (R) with the hydrological catchments to count the number of threatened and nonthreatened amphibian species. We excluded catchments with fewer than two extant species from aggregation and scoring.

$$r_{AMPH,j} = \frac{\text{count (threatened (R))}}{\text{count (R)}}, for\ (R|R \cap j \neq \emptyset\})$$

The threatened amphibians indicator excludes species that are extinct or locally extirpated, and thus estimates current levels of vulnerability rather than total degradation. Furthermore, this indicator does not attempt to differentiate among the causes of species decline.

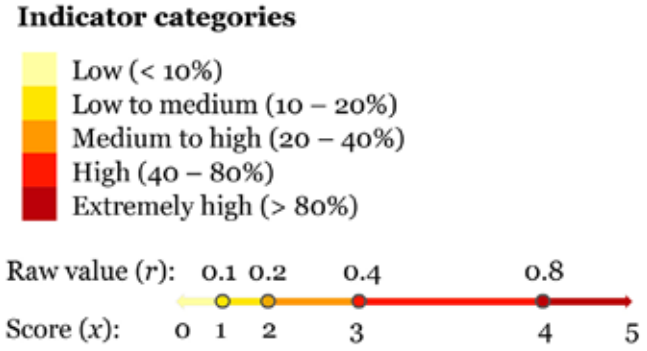
5 INDICATOR SCORING AND AGGREGATION

After computing raw values of indicators, we normalized indicators for display and aggregation. We used a linear weighted aggregation approach, and displayed indicators on an online interactive platform to enable dynamic aggregation.

5.1 Thresholds and Normalization

The first step in aggregation is to place all indicators on a comparable scale. We normalized indicators over a set of thresholds, which were chosen to divide indicators into five categories. For each indicator, we determined thresholds using existing literature, the range and distribution of indicator values, and expert judgment. For example, for baseline water stress, our thresholds reflect thresholds used by other withdrawal-to-availability indicators.[35]

Figure 9 | **Example Indicator Raw Values Normalized Over Five Categories to Scores between 0 and 5**



We then mapped raw values over the thresholds using continuous functions, normalizing the indicators to scores between 0 and 5, such that scores less than or equal to 1 correspond with the lowest category, and scores greater than 4 correspond with the highest category (Figure 9). The normalization functions vary by indicator but generally follow either a simple linear or a logarithmic form (Table 5).

The threshold method of indicator normalization has several advantages and one notable disadvantage. The foremost advantage is that it creates clear categories, and enables scores to be matched to guidelines. Relative to purely statistical methods of normalization, the threshold method is independent of the distribution of data, and thus is unaffected by extreme values, allowing for comparison across different samples or new datasets.

However, the scoring of indicators, whether based on established guidelines or statistical distributions, is subjective. By defining thresholds, we assign meaning to specific indicator values. To maintain transparency in the process, we display the relationship between raw values and categories and enable users to access the raw indicator values.

Table 5 | **Indicator Normalization Functions from Raw Indicator Values, *r*, to Scores, *x*.**

INDICATOR	FUNCTION	LOWEST CATEGORY ($x \leq 1$)	HIGHEST CATEGORY ($x > 4$)
BWS	$x = \max\left(0, \min\left(5, \frac{\ln(r) - \ln(0.1)}{\ln(2)} + 1\right)\right)$	$r \leq 10\%$	$r > 80\%$
IAV	$x = \max\left(0, \min\left(5, \frac{r}{0.25}\right)\right)$	$r \leq 0.25$	$r > 1$
SV	$x = \max\left(0, \min\left(5, \frac{r}{1/3}\right)\right)$	$r \leq 1/3$	$r > 1 1/3$
STOR	$x = \max\left(0, \min\left(5, \frac{\ln(r) - \ln(0.1)}{-\ln(2)} + 1\right)\right)$	$r \geq 1$	$r < 0.125$
RFR	$x = \max\left(0, \min\left(5, \frac{\ln(r) - \ln(0.1)}{\ln(2)} + 1\right)\right)$	$r \leq 10\%$	$r > 80\%$
PROT	$x = \max\left(0, \min\left(5, \frac{\ln(r) - \ln(0.4)}{-\ln(2)} + 1\right)\right)$	$r \geq 40\%$	$r < 5\%$
FO	$x = \max\left(0, \min\left(5, \frac{\ln(r) - \ln(1)}{\ln(3)} + 1\right)\right)$	$r \leq 1$	$r > 27$
DRO	$x = \max\left(0, \min\left(5, \frac{r - 10}{10}\right)\right)$	$r \leq 20$	$r > 50$
GW	$x = \begin{cases} \min\left(5, \frac{\ln(r) - \ln(5)}{\ln(2)} + 1\right), r_{GW} \geq 5 \\ \max\left(0, \frac{\ln(r+1.5) - \ln(5)}{\ln(2)} + 1\right), r_{GW} < 3.5 \end{cases}$	$r \leq 1$	$r > 20$
MC	$x = \max\left(0, \min\left(5, \frac{\ln(r) - \ln(0.0005)}{\ln(2)} + 1\right)\right)$	$r \leq 0.05\%$	$r > 0.4\%$
WC	$x = \max\left(0, \min\left(5, \frac{\ln(r) - \ln(0.025)}{\ln(2)} + 1\right)\right)$	$r \leq 2.5\%$	$r > 20\%$
AMPH	$x = \max\left(0, \min\left(5, \frac{\ln(r + 0.05)}{\ln(2)}\right)\right)$	$r = 0\%$	$r > 35\%$

Note: BWS = Baseline water stress, IAV = Inter-annual variability, SV = Seasonal variability, STOR = Upstream storage, RFR = Return flow ratio, PROT = Upstream protected land, FO = Flood occurrence, DRO = Drought severity, GW = Groundwater stress, MC = Media coverage, WC = Access to water, AMPH = Threatened amphibians.

5.2 Weighting

Exposure to water-related risks varies with the characteristics of water users. To obtain aggregated water-risk scores, we allow users to modify the weight of each indicator to match their exposure to the different aspects of water risk. We expose five possible weights, or descriptors of importance, on a base two exponential scale, which is preferred over a linear scale because of the human tendency to categorize intensity by orders of magnitude of difference.[36] Users can also exclude indicators completely from aggregation.

To determine a default set of indicator weights, we used input from six staff water experts following the principles of the Delphi technique. This technique uses a series of intensive questionnaires interspersed with controlled opinion feedback to obtain the most reliable consensus of opinion from a group of experts.[37] The Delphi technique is intended for use in judgment situations, in which pure model-based statistical methods are not practical or possible because of the lack of appropriate historical data, and thus where some form of human judgment input is necessary.[38] The lack of consistent information on exposure to water risks and the subjective nature of indicator weights made this technique an ideal fit.

Additionally, we developed pre-set weighting schemes for nine industry sectors based on information provided in corporate water disclosure reports[39] and input from industry experts to reflect the particular risks and challenges faced by each water-intensive sector (Figure 10). For each sector, we modified the default indicator weights based on the relative importance of each indicator to the sector based on information disclosed by companies on their exposure to, and losses from, water-related risks. To validate the industry-sector pre-set weighting schemes, we presented preliminary weighting schemes to industry representatives from the nine sectors and solicited feedback on the relative importance of each indicator for their sector.

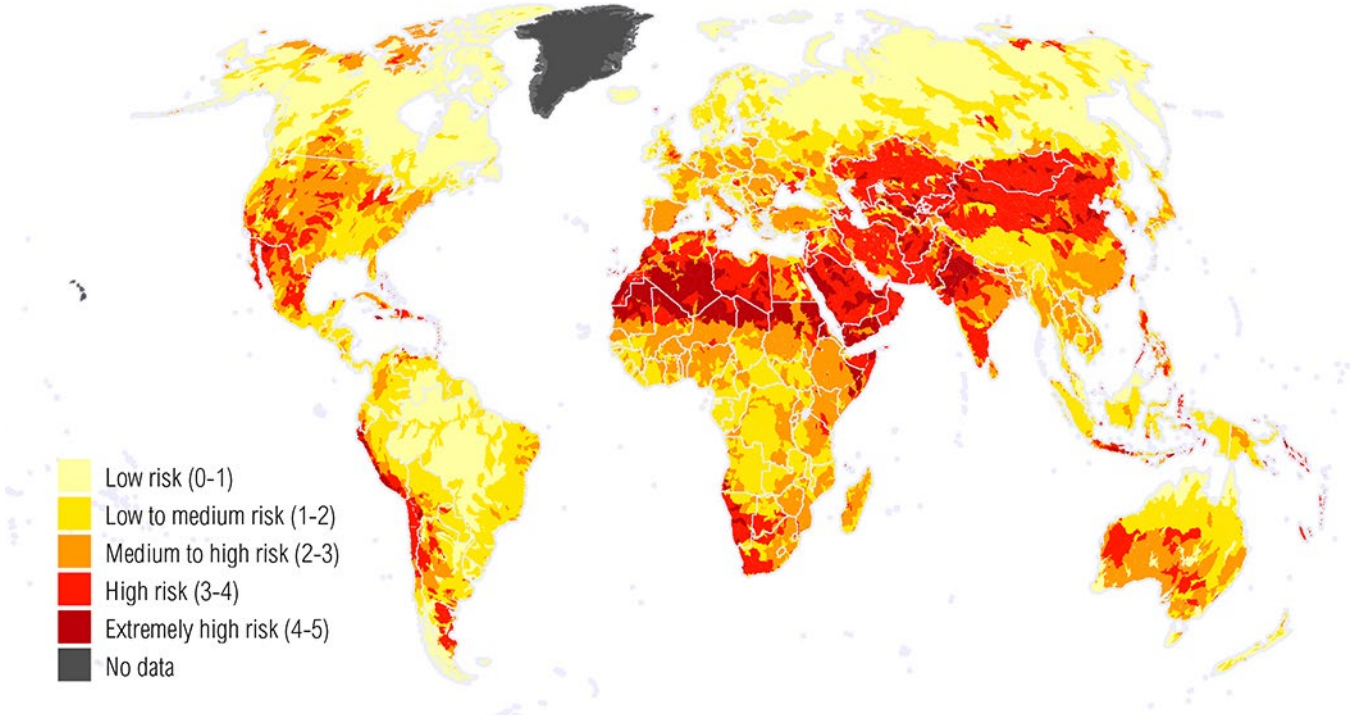
5.3 Indicator Aggregation

Finally, to obtain estimates of overall water risk, the Aqueduct tool combines individual indicator scores into aggregated scores using linear aggregation. Specifically, for any set of indicators, I, it computes a weighted average (a) for each area (j) as the sum of the indicator scores (x) times their weights (w), divided by the sum of all the weights. Indicators in areas for which there are no data are excluded from the weighted average for those areas.

$$a_j = \frac{\sum x_{ij} w_i}{\sum w_i}, (i \in I | x_{ij} \neq NULL)$$

Figure 10 | **Weighting Schemes**

Figure 11 | **Overall Water Risk Default Weighting**



Source: WRI Aqueduct.

Since the weighted average pulls all indicator values toward the mean, we rescaled the aggregated scores to extend through the full range of values (0–5) to generate a final displayed score (s).

$$s_j = 5 \left(\frac{a_j - min\ (a)}{\max\ (a) - min\ (a)} \right)$$

This approach more clearly communicates the full range of relative risk values given the user's chosen weights.

6 DISCUSSION

The Aqueduct Water Risk Atlas provides the means to compare spatial variations in potential water issues across the globe. The composite overall water risk index (Figure 11) highlights regions of high risk and provides an entry point for users into the database, while individual indicators more clearly measure specific water resource characteristics. Moreover, by identifying regions that face the highest risks, Aqueduct may be used to prioritize where more detailed local-scale analysis is merited.

These global indicators are best suited for comparative analyses across large geographies to identify regions or assets deserving of closer attention, and are not appropriate for catchment or site-specific analyses. Indeed, global-scale indices such as this face significant limitations in their ability to accurately and objectively capture most aspects of the underlying phenomena at specific places.

At the composite index level, the selection of aggregation methods is an inherently subjective process that creates value by simplifying complex phenomena. Although we make every effort to create a robust and objective framework, the academic and professional discourse remains without a single best way to combine and compare diverse metrics into a composite index. Therefore, results, to some extent, reflect the judgment of the authors and expert

advisors. However rigorous the results produced, this exercise inevitably runs up against the limits of describing the complexity of water risks with a single number.

Furthermore, validation of the composite index of water risk is a challenge because of the difficulty of collecting risk-event data. Consistently measured, geographically explicit data on economic or environmental losses caused by water resource constraints would greatly aid in this exercise. Instead, we rely heavily on the robustness of input data and methods as well as expert review panels to ensure a meaningful product.

Other possible water-related risks also escape our index. The complex and qualitative nature of regulatory and reputational drivers of risk complicates researchers' ability to create useful metrics. Water policies and regulations can limit or redistribute the use of water in ways for which we are unable to account.

Barriers, such as inconsistent availability of data, as well as the unwillingness or lack of capacity for governments to collect and share water data, hamper the construction of consistent global water information.[40] Water infrastructure is likewise undermeasured, as there are no published global datasets of major water transfer projects, infrastructural losses, or reservoir evaporation. Similarly, methods for evaluating freshwater ecosystem services such as flood attenuation and pollution control at the global scale remain underdeveloped. While improvements in modeling methodologies could overcome some of these barriers, collection and sharing of data by national governments is essential to the creation of accurate global water information.

Nothing can fully replace locally calibrated information on water infrastructure, policy, and management practices. Nonetheless, Aqueduct's global metrics and associated maps provide an easily accessible entry point to evaluating water risks, enabling audiences to better understand the relative importance of major water resource constraints among specific geographies, regardless of expertise. The World Resources Institute plans to continue to improve the Aqueduct indicator and aggregation methodology. We welcome comments and suggestions from interested parties. For more information on the Aqueduct Water Risk Atlas please visit www.wri.org/aqueduct.

ENDNOTES

1. Paul Reig, Tien Shiao, and Francis Gassert, "Aqueduct Water Risk Framework," Working Paper (Washington, DC: World Resources Institute, January 2013), http://www.wri.org/publication/aqueduct-water-risk-framework.
2. Geoff Dabelko and Meaghan Parker, *Seven Ways 7 Billion People Affect the Environment and Security* (Washington, DC: Woodrow Wilson Center for Scholars, 2013), http://www.newsecuritybeat.org/2013/01/ways-billion-people-affect-environment-security-policy-brief/; World Economic Forum, *Global Risks 2013, Eighth Edition* (Cologny/Geneva, Switzerland, 2013), http://www3.weforum.org/docs/WEF_GlobalRisks_Report_2013.pdf.
3. Matti Kummu et al., "Is Physical Water Scarcity a New Phenomenon? Global Assessment of Water Shortage over the Last Two Millennia," *Environmental Research Letters* 5, no. 3 (July 16, 2010): 034006, doi:10.1088/1748-9326/5/3/034006.
4. C.J. Vorosmarty et al., "Global Threats to Human Water Security and River Biodiversity," *Nature* 467, no. 7315 (2010): 555–61, http://www.nature.com/nature/journal/v467/n7315/full/nature09440.html#/methods.
5. Caroline A. Sullivan, "Quantifying Water Vulnerability: A Multi-Dimensional Approach," *Stochastic Environmental Research and Risk Assessment* 25, no. 4 (2011): 627–40, doi:10.1007/s00477-010-0426-8.
6. Paul Reig, et al. "Aqueduct Water Risk Framework," 2013.
7. Yuji Masutomi et al., "Development of Highly Accurate Global Polygonal Drainage Basin Data," *Hydrological Processes* 23, no. 4 (February 15, 2009): 572–584, doi:10.1002/hyp.7186.
8. *Catchment* and *basin* are often used synonymously. However in this paper, we use basin to refer to the entire area of land that drains to an ocean or closed inland water body (e.g. the Nile River Basin), and catchment as a more general term that can refer to portions of a larger river basin that share a common outlet (e.g. portions of the main stem or tributaries of a large river).
9. Otto Pfafstetter, "Classification of hydrographic basins: coding methodology," unpublished manuscript, Departamento Nacional de Obras de Saneamento, 1989: 1-2; K.L Verdin and J.P Verdin, "A Topological System for Delineation and Codification of the Earth's River Basins," *Journal of Hydrology* 218, no. 1–2 (May 1999): 1–12, doi:10.1016/S0022-1694(99)00011-6.
10. Food and Agriculture Organization of the United Nations, "AQUASTAT - FAO's Information System on Water and Agriculture."
11. Peter Gleick et al., *The World's Water Volume 7* (Washington, DC: Island Press, 2011), http://worldwater.org/data.html.
12. The primary differentiating factor between domestic (i.e. urban) and industrial or agricultural withdrawals is that domestic withdrawals are supplied by a public distribution system while the others are not. Agricultural withdrawals are the amount above and beyond rain-fed agriculture (the effect of which is accounted for in land surface modeling). See: A Kohli, K Frenken, and C Spottorno, *Disambiguation of Water Statistics*, May 2012.
13. Judith D. Singer and John B. Willett, *Applied Longitudinal Data Analysis* (New York: Oxford University Press, 2003); Jose C. Pinheiro and Douglas M. Bates, *Mixed-Effects Models in S and S-Plus* (New York: Springer Verlag, 2000).
14. Badi H. Baltagi, *Econometric Analysis of Panel Data*, 4th ed. (Chichester: John Wiley & Sons, 2008).
15. W. S. Cleveland, E. Grosse, and W. M. Shyu, "Local Regression Models," in *Statistical Models in S*, edited by J.M. Chambers and T. J. Hastie (Wadsworth and Brooks / Cole, 1992).
16. James Honaker and Gary King, "What to Do About Missing Values in Time-Series Cross-Section Data," *American Journal of Political Science* 54, no. 2 (April 2010): 561–81, doi:10.1111/j.1540-5907.2010.00447.x.
17. Jose C. Pinheiro et al., "NIme: Linear and Nonlinear Mixed Effects Models. R Package Version 3.1-108," 2013, http://cran.r-project.org/web/packages/nlme/nlme.pdf.
18. H. Akaike, "A New Look at the Statistical Model Identification," *IEEE Transactions on Automatic Control* 19, no. 6 (1974): 716–23, doi:10.1109/TAC.1974.1100705.
19. G. E. Schwarz, "Estimating the Dimension of a Model," *Annals of Statistics* 6, no. 2 (1978): 461–64, doi:10.1214/aos/1176344136.
20. I. A. Shiklomanov and John C. Rodda, eds., *World Water Resources at the Beginning of the Twenty-First Century* (Cambridge University Press, 2004).
21. M. Rodell et al., "The Global Land Data Assimilation System," *Bulletin of the American Meteorological Society* 85, no. 3 (2004): 381–94, http://dx.doi.org/10.1175/BAMS-85-3-381.
22. Benjamin F. Zaitchik, Matthew Rodell, and Francisco Olivera, "Evaluation of the Global Land Data Assimilation System Using Global River Discharge Data and a Source-to-Sink Routing Scheme," *Water Resources Research* 46, no. 6 (2010): 1–17, doi:10.1029/2009WR007811.
23. NCAR Research Applications Laboratory, "Land Surface Modeling: The Community Noah Land Surface Model (LSM)," https://www.rap.ucar.edu/research/land/technology/lsm.php, accessed January 11, 2014.
24. Amber Brown and Marty D Matlock, "A Review of Water Scarcity Indices and Methodologies," white paper (The Sustainability Consortium, 2011), http://www.sustainabilityconsortium.org/wp-content/themes/sustainability/assets/pdf/whitepapers/2011_Brown_Matlock_Water-Availability-Assessment-Indices-and-Methodologies-Lit-Review.pdf.
25. Some authors have attempted to account for dams and other storage features using basic operational rules. See, for example: Marc F. P. Bierkens, L. P. H. van Beek, and Yoshihide Wada, "Global Monthly Water Stress: 1. Water Balance and Water Availability," *Water Resources Research*, 2011, doi:10.1029/2010WR009791; L. P. H. van Beek et al., "Global Monthly Water Stress: 2. Water Demand and Severity of Water Stress," *Water Resources Research*, 2011, doi:10.1029/2010WR009792.
26. B. Lehner et al., "High-Resolution Mapping of the World's Reservoirs and Dams for Sustainable River-Flow Management," *Frontiers in Ecology and the Environment* 9, no. 9 (May 2011): 494–502, doi:10.1890/100125.
27. International Union for Conservation of Nature (IUCN) and United Nations Environment Programme World Conservation Monitoring Centre (UNEP-WCMC), "The World Database on Protected Areas," June 2012, http://protectedplanet.net/.

28. Category V protected areas are defined as "areas where the interaction of people and nature over time has produced an area of distinct character with significant ecological, biological, cultural and scenic value and where safeguarding the integrity of this interaction is vital to protecting and sustaining the area and its associated nature conservation and other values." Category V is one of the more flexible categories, allowing for historical and contemporary development such as ecotourism. Nigel Dudley, ed., *Guidelines for Applying Protected Area Management Categories* (Gland, Switzerland: IUCN, 2008), doi:10.2305/IUCN.CH.2008.PAPS.2.en.
29. Robin Abell, J. David Allan, and Bernhard Lehner, "Unlocking the Potential of Protected Areas for Freshwaters," *Biological Conservation*, 2007, doi:10.1016/j.biocon.2006.08.017.
30. G.R. Brakenridge, "Global Active Archive of Large Flood Events", Dartmouth Flood Observatory, University of Colorado, http://floodobservatory.colorado.edu/Archives/index.html.
31. Justin Sheffield and Eric F Wood, "Projected Changes in Drought Occurrence Under Future Global Warming from Multi-Model, Multi-Scenario, IPCC AR4 Simulations," *Climate Dynamics* 31 (2008): 79–105, http://link.springer.com/article/10.1007/s00382-007-0340-z.
32. Tom Gleeson et al., "Water Balance of Global Aquifers Revealed by Groundwater Footprint," *Nature* 488, no. 7410 (2012): 197–200, doi:10.1038/nature11295.
33. World Health Organization (WHO) and the United Nations Children's Fund (UNICEF), "WHO / UNICEF Joint Monitoring Programme (JMP) for Water Supply and Sanitation," 2012, www.wssinfo.org.
34. International Union for Conservation of Nature (IUCN), "The IUCN Red List of Threatened Species," 2010, http://www.iucnredlist.org/technical-documents/spatial-data#amphibians.
35. Charles J. Vörösmarty et al., "Global Water Resources: Vulnerability from Climate Change and Population Growth," *Science* 289, no. 5477 (2000): 284–88, http://www.sciencemag.org/cgi/doi/10.1126/science.289.5477.284; United Nations Commission on Sustainable Development (UNCSD), *Comprehensive Assessment of the Freshwater Resources of the World* (New York, 1997), http://www.un.org/ga/search/view_doc.asp?symbol=E/CN.17/1997/9&Lang=E.
36. T. Evangelos, *Multi-criteria Decision Making Methods: A Comparative Study*, (Dordrecht: Kluwer Academic Publishers), 2000).
37. Gene Rowe and George Wright, "The Delphi Technique as a Forecasting Tool: Issues and Analysis," *International Journal of Forecasting* 15, no. 4 (October 1999): 353–75, http://www.sciencedirect.com/science/article/pii/S0169207099000187.
38. Norman Dalkey and Olaf Helmer, "An Experimental Application of the DELPHI Method to the Use of Experts," *Management Science* 9 , no. 3 (April 01, 1963): 458–67, doi:10.1287/mnsc.9.3.458.
39. Carbon Disclosure Project, *Collective Responses to Rising Water Challenges*, 2012, https://www.cdproject.net/CDPResults/CDP-Water-Disclosure-Global-Report-2012.pdf; Carbon Disclosure Project, *Raising Corporate Awareness of Global Water Issues*, 2011, https://www.cdproject.net/CDPResults/CDP-Water-Disclosure-Global-Report-2011.pdf; Berkley Adrio, *Clearing the Waters: A Review of Corporate Water Risk Disclosure in SEC Filings*, June 2012, http://www.ceres.org/resources/reports/clearing-the-waters-a-review-of-corporate-water-risk-disclosure-in-sec-filings/view. IFC-ESAT (2005) "Sector Fact Sheet: Textiles and Apparel," available at: http://firstforsustainability.org/documents/factsheet_textiles.pdf. UBS Investment Research (2012) Q Series: Water Risk to Businesses.
40. Sullivan, Caroline A, "Quantifying Water Vulnerability: A Multi-Dimensional Approach," *Stochastic Environmental Research and Risk Assessment* 25, no. 4 (2011): 627–640, doi:10.1007/s00477-010-0426-8.
41. ISciences L.L.C., "Freshwater Sustainability Analyses: Interpretive Guidelines," November 2011, http://docs.wri.org/aqueduct/freshwater_sustainability_analyses.pdf.
42. B. M. Fekete, C. J. Vörösmarty, and W. Grabs, "High-Resolution Fields of Global Runoff Combining Observed River Discharge and Simulated Water Balances," *Global Biogeochemical Cycles* 16, no. 3 (2002): 10–15, doi:10.1029/1999GB001254.
43. Ibid.; D.R. Legates and C.J. Willmott. Mean Seasonal and Spatial Variability in Gauge-Corrected, Global Precipitation," *Journal of Climatology* 10 (1990):111–27; D.R. Legates and C.J. Willmott, "Mean Seasonal and Spatial Variability in Global Air Temperature," *Theoretical Applied Climatology* 41(1990):11–21.
44. Zaitchik, Rodell, and Olivera, "Evaluation of the Global Land Data Assimilation System Using Global River Discharge Data and a Source-to-Sink Routing Scheme."
45. Reichle, Rolf H., Randal D. Koster, Gabriëlle J. M. De Lannoy, Barton a. Forman, Qing Liu, Sarith P. P. Mahanama, and Ally Touré. "Assessment and Enhancement of MERRA Land Surface Hydrology Estimates." Journal of Climate 24, no. 24 (December 2011): 6322–6338. doi:10.1175/JCLI-D-10-05033.1.
46. NASA Goddard Earth Sciences Data and Information Services Center, "GLDAS Version 2 (GLDAS-2) Data Have Been Released," accessed May 22, 2013, http://disc.sci.gsfc.nasa.gov/gesNews/gldas_2_data_release.
47. Zaitchik, Rodell, and Olivera, "Evaluation of the Global Land Data Assimilation System Using Global River Discharge Data and a Source-to-Sink Routing Scheme."

APPENDIX A SUPPLEMENTARY INFORMATION

A1 2010 Withdrawal Projection Model Supplementary Figures

Table A1a | **Coefficients (± standard error) for Four Models of Domestic Water Withdrawals**

VARIABLE	ME-AIC	ME-BIC	FE-AIC AND BIC
Intercept	-8.1 ± 0.97	-7.9 ± 0.59	
YEAR	0.005 ± 0.0017	0.005 ± 0.0017	
POP	1 ± 0.16	1.22 ± 0.098	0.02 ± 0.42
GDP	0.1 ± 0.16	-0.1 ± 0.089	-0.5 ± 0.26
WWR	0.4 ± 0.27	0.5 ± 0.25	
PRCP	0.00012 ± 0.00003	0.00013 ± 0.00003	
URB	1 ± 1.1	0.8 ± 0.11	5 ± 1.7
POP:WWR	0.16 ± 0.042	0.2 ± 0.038	
GDP:WWR	-0.13 ± 0.038	-0.15 ± 0.033	
POP:URB	0.4 ± 0.17		
GDP:URB	-0.2 ± 0.17		-0.4 ± 0.17
POP:GDP			0.11 ± 0.038
POP:PRCP			0.0005 ± 0.00013

Source: WRI Aqueduct.
Note: Mixed effects models use variable selection based on AIC and BIC, and fixed effects models are based on AIC and BIC. Variable abbreviations as in Table 1.

Table A1b | **Coefficients (± standard error) for Four Models of Industrial Water Withdrawals**

VARIABLE	ME-AIC AND BIC	FE-AIC	FE-BIC
Intercept	-17 ± 5		
YEAR	0.014 ± 0.0049		
ELEC	-3 ± 1.3		
GDP	1.3 ± 0.49	-48.3 ± 24	-0.5 ± 0.45
GDP_2		4 ± 2.2	
GDP_3		-0.13 ± 0.067	
WWR	0.27 ± 0.051		
CO_2	3 ± 1.5		
POP	0.6 ± 0.095		
COAL	0.06 ± 0.018		
GDP:ELEC	0.3 ± 0.13		
GDP:CO_2	-0.3 ± 0.15		
GDP:WWR		-0.8 ± 0.2	-0.8 ± 0.18

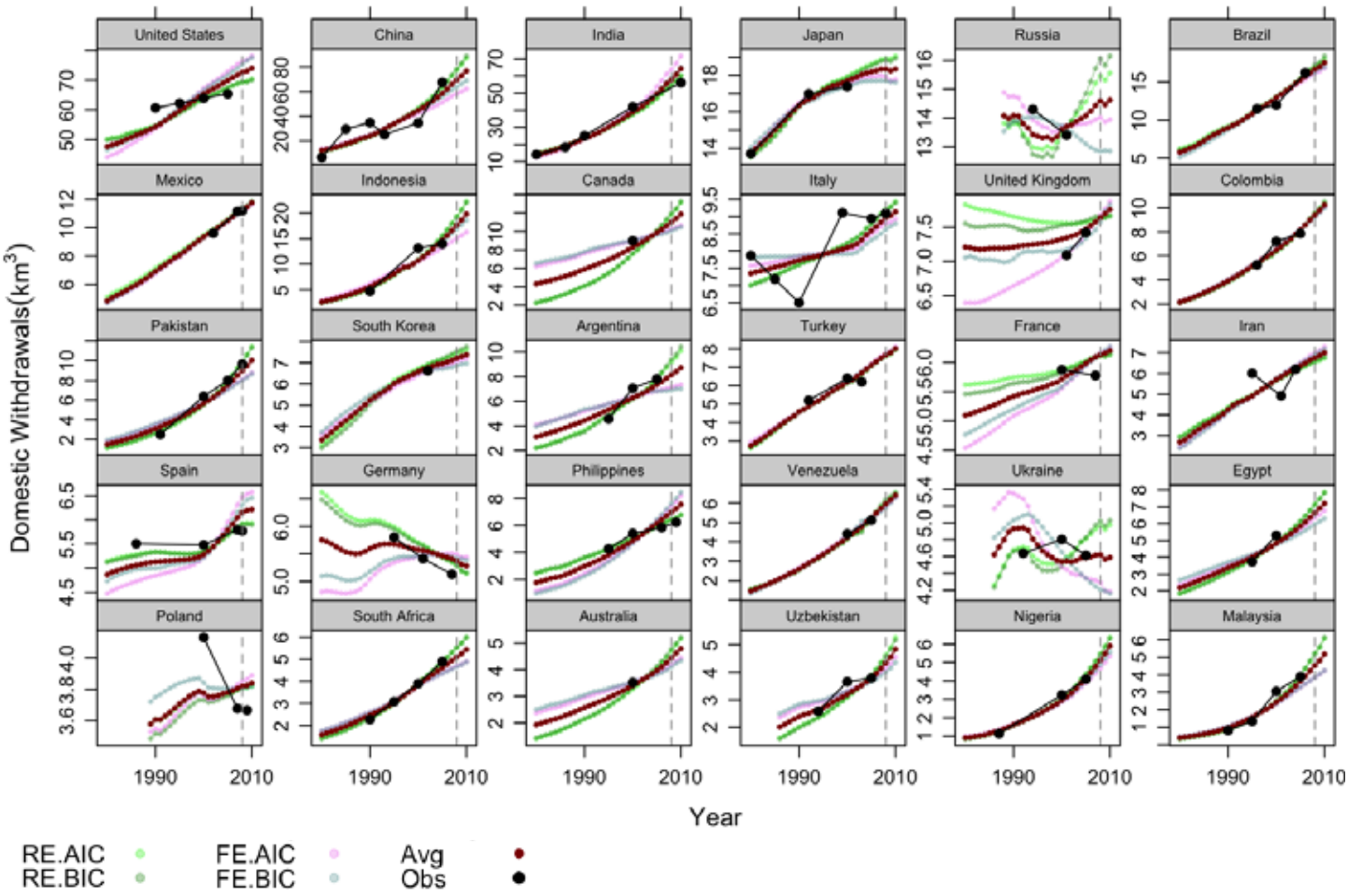
Source: WRI Aqueduct project.
Note: Mixed effects models use variable selection based on AIC and BIC, and fixed effects models are based on AIC and BIC. Variable abbreviations as in Table 1.

Table A1c | **Coefficients (± standard error) for Four Models of Agricultural Water Withdrawals**

VARIABLE	ME-AIC	ME-BIC	FE-AIC	FE-BIC
Intercept	-2 ± 1.5	-2.8 ± 0.66		
IGAREA	0.5 ± 0.36	0.53 ± 0.056	0.2 ± 0.16	0.1 ± 0.15
GDP	-0.2 ± 0.15	-0.16 ± 0.074	2 ± 1.2	-0.03 ± 0.094
GDP2			-0.16 ± 0.096	
AGR	0.4 ± 0.047	0.38 ± 0.046	-0.4 ± 0.28	
POP	0.22 ± 0.084	0.22 ± 0.083	-2 ± 1.6	0.7 ± 0.21
PRCP	0.00045 ± 0.00004	0.00044 ± 0.00004		
WWR	1 ± 0.49	1.3 ± 0.28		
POP:GDP			0.2 ± 0.15	
IGAREA:GDP	0.004 ± 0.035			
IGAREA:WWR	0.2 ± 0.14	0.05 ± 0.019	0.32 ± 0.098	0.25 ± 0.093
GDP:WWR	-0.05 ± 0.051	-0.08 ± 0.031	-0.15 ± 0.093	
GDP:WWR:IGAREA	-0.02 ± 0.014			
IGAREA:PRCP			0.0005 ± 0.00017	0.0005 ± 0.00017
GDP:PRCP			-0.0002 ± 0.00015	

Source: WRI Aqueduct project.
Note: Mixed effects models use variable selection based on AIC and BIC, and fixed effects models are based on AIC and BIC. Variable abbreviations as in Table 1.

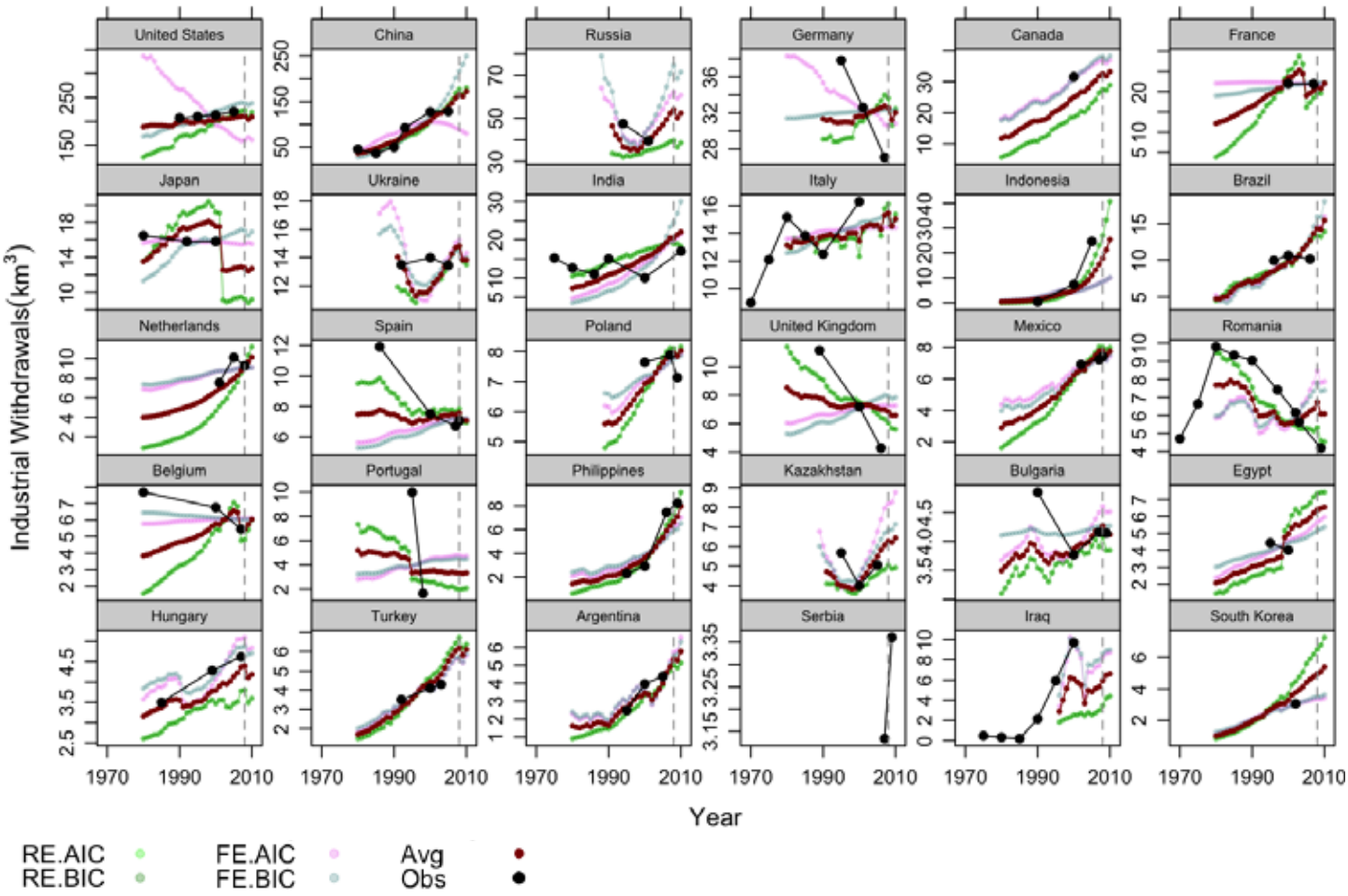
Figure A1a | **Domestic Withdrawals for the 30 Countries with the Highest Sectoral Water Use**



Source: WRI Aqueduct.

Note: Countries are listed in decreasing order of water use. Note differences in the y-axis limits. Model fits are shown for each of the four estimated models (see section 2.2.2 for descriptions of the models) as well as the average of all four models (Avg) and the observed points (Obs).

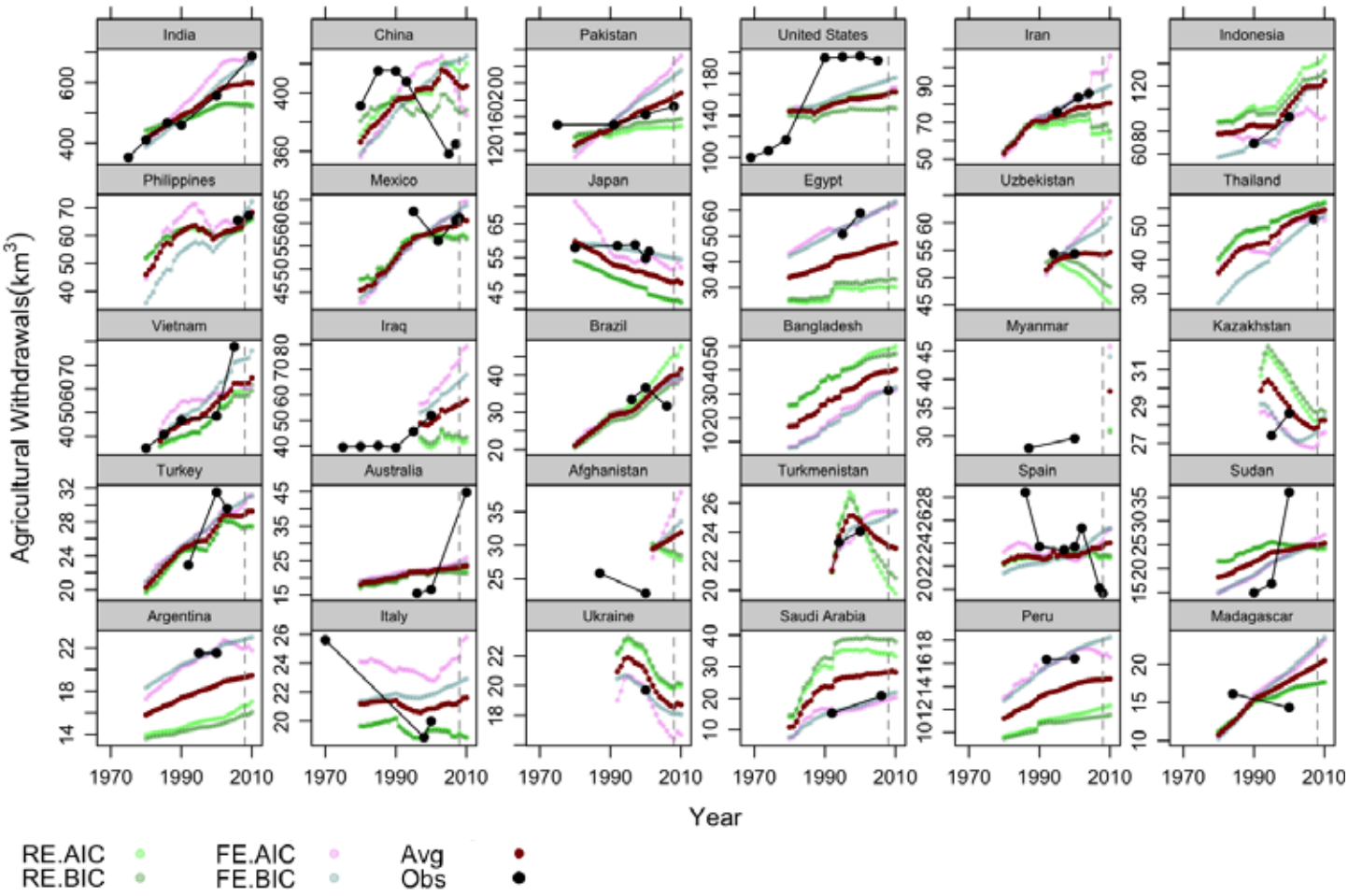
Figure A1b | **Industrial Withdrawals for the 30 Countries with the Highest Sectoral Water Use**



Source: WRI Aqueduct.

Note: Countries are listed in decreasing order of water use. Note differences in the y-axis limits. Model fits are shown for each of the four estimated models (see section 2.2.2 for descriptions of the models) as well as the average of all four models (Avg) and the observed points (Obs).

Figure A1c | **Agricultural Withdrawals for the 30 Countries with the Highest Sectoral Water Use**



Source: WRI Aqueduct.

Note: Countries are listed in decreasing order of water use. Note differences in the y-axis limits. Model fits are shown for each of the four estimated models (see section 2.2.2 for descriptions of the models) as well as the average of all four models (Avg) and the observed points (Obs).

A2 Runoff Model Comparison

Water supply estimates in the previous iteration of Aqueduct water stress maps (developed by ISciences, L.L.C. for the Coca-Cola Company)[41] were based on the University of New Hampshire Global Runoff Data Center (UNH-GRDC) dataset.[42] The UNH-GRDC dataset was developed specifically with the goal of maximizing the accuracy of runoff measurements, and includes extensive gauge data in its construction. Despite its accuracy, several aspects make it unsuitable for use here.

First, the UNH-GRDC dataset uses climatologically averaged precipitation, temperature, and gauged river discharge data across approximately 40 years, so it cannot be used to estimate temporal variability. Additionally, the climatology covers an unspecified time period, since the underlying data are based on inconsistent observed data over several decades ending no later than 1990.[43] Finally, because of an issue with an underlying equation in the water balance model, the UNH-GRDC dataset makes arid and semi-arid areas too dry. Nonetheless, because of its demonstrated accuracy over most of the rest of the world, we used this dataset as a reference against which to evaluate bias in selecting a runoff dataset.

We evaluated three other datasets for runoff bias, spatial resolution, and temporal extent (Table A2a), and assumed a climatological period of 1950–90 for the UNH-GRDC dataset for purposes of comparison. Note that, only GLDAS-2 can be directly compared with UNH-GRDC because only it has an overlapping temporal extent. Therefore, GLDAS-2 was directly compared with UNH-GRDC, and the other two models were compared with GLDAS-2.

Table A2a | **Evaluated Global Runoff Models for Annual Water Supply**

DATASET	RESOLUTION (DEGREES)	TEMPORAL EXTENT
UNH-GRDC[a]	0.5	1950-1990 climate norm
GLDAS-2 Noah[b]	1.0	1948-2008
NCEP CFSR Noah[c]	0.3	1979-2010
MERRA-Land[d]	0.5 × 0.66	1980-present

Source: WRI Aqueduct.

Notes:

a. D.R. Legates and C.J. Willmott. Mean Seasonal and Spatial Variability in Gauge-Corrected, Global Precipitation," *Journal of Climatology* 10 (1990):111–27;. D.R. Legates and C.J. Willmott, "Mean Seasonal and Spatial Variability in Global Air Temperature," Theoretical *Applied Climatology* 41(1990):11–21.

b. M. Rodell et al., "The Global Land Data Assimilation System," *Bulletin of the American Meteorological Society* 85, no. 3 (2004): 381–94, http://dx.doi.org/10.1175/BAMS-85-3-381.

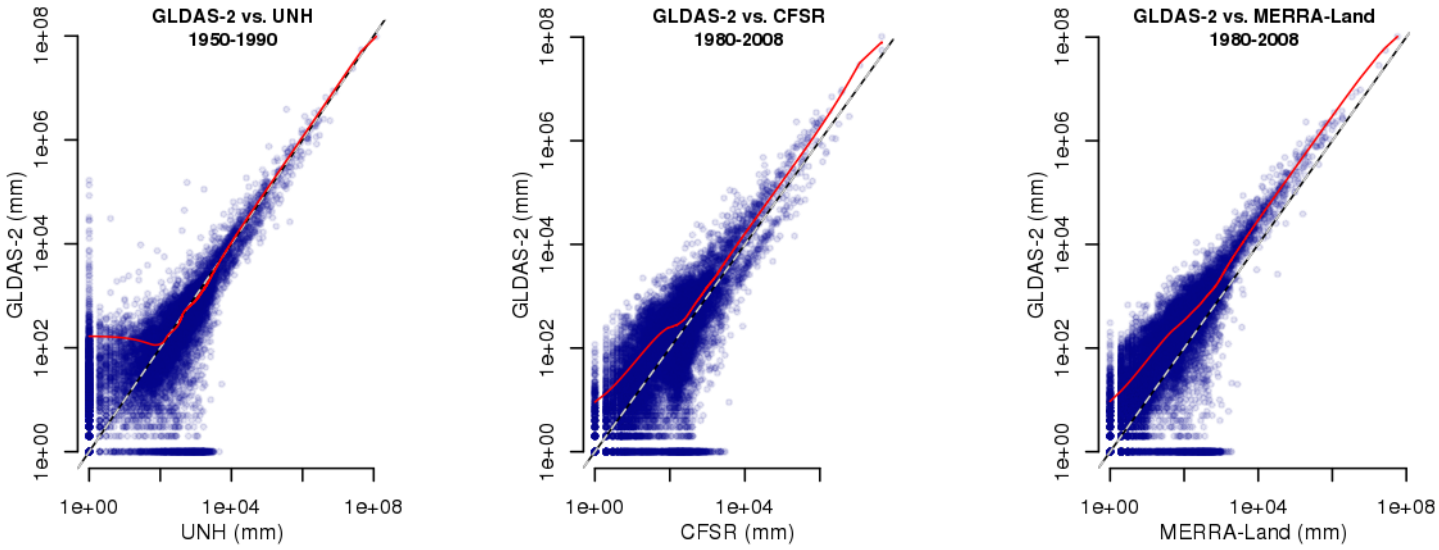
c. S. Saha et al., "The NCEP Climate Forecast System Reanalysis," *Bulletin of the American Meteorological Society* August (2010): 1015–1057, doi:10.1175/2010Bams3001.1.

d. Rolf H. Reichle et al., "Assessment and Enhancement of MERRA Land Surface Hydrology Estimates," *Journal of Climate* 24, no. 24 (December 2011): 6322–6338, doi:10.1175/JCLI-D-10-05033.1.

Of the three datasets, to our knowledge only GLDAS-2 runoff has been validated against gauged discharge in terms of absolute deviations.[44] MERRA-Land has been compared with gauged data, but only in relative terms.[45] That is to say, MERRA-Land anomalies have been assessed for accuracy, but absolute estimates have not been assessed. In fact, in our comparisons of long-term mean runoff from GLDAS-2, CFSR, and MERRA-Land vs. UNH-GRDC, we found that MERRA-Land was the driest of the four datasets (Figure A2), but that GLDAS-2 closely matched the UNH-GRDC data with little overall bias over most of the globe. The exception is in arid to semi-arid regions (note departure of the red line from the 1:1 line in the lower left panel of Figure A2), as discussed above.

In addition to its accuracy, several factors cause us to favor GLDAS-2 over CFSR and MERRA-Land. First, it has the longest period of record, extending from 1948 to 2008. Second, there are substantial plans to extend the model period to the present, and update the resolution to 0.25 degree.[46] Although GLDAS-2 currently uses only the Noah land surface model to produce runoff estimates, it will eventually be provided as an ensemble of four different land surface models. This is important because the underlying land surface model can be a significant source of variation in runoff estimates,[47] and therefore an ensemble is preferred to minimize the bias resulting from any one model. Future enhancements in GLDAS should enable us to maintain continuity in Aqueduct's water supply estimates while improving overall accuracy.

Figure A2a | **Comparison of Flow Accumulated Runoff (*Bt*) among Four Datasets**



Source: WRI Aqueduct project.

Note: See Table A2a for characteristics of datasets. Points represent individual catchments (n = 14998); the dashed line indicates a 1:1 relationship; the red line indicates a loess smoother, and the time period indicates the interval over which the runoff values were averaged. Departures of the red line from the 1:1 line indicate bias of one dataset relative to the other.

ABOUT WRI

The World Resources Institute (WRI) is a global environmental and development think tank that focuses on the intersection of the environment and socioeconomic development. We go beyond research to put ideas into action, working globally with governments, business, and civil society to build transformative solutions that protect the earth and improve people's lives.

ABOUT THE AUTHORS

Francis Gassert is a research analyst with the Food, Forests, and Water Program at WRI, where he manages data collection and GIS analysis for the Aqueduct project.
Contact: fgassert@wri.org

Matt Landis is a research scientist at ISciences, L.L.C., where he develops and applies hydrological algorithms and models.

Matt Luck is a research scientist at ISciences, L.L.C., where he develops and applies hydrological algorithms and models.

Paul Reig is an associate with the Food, Forests, and Water Program at WRI, where he leads the design and development of the Aqueduct project.
Contact: preig@wri.org

Tien Shiao is a senior associate with the Food, Forests, and Water Program at WRI, where she leads the application and road testing of the Aqueduct project.
Contact: tshiao@wri.org

ACKNOWLEDGMENTS

This publication was made possible thanks to the ongoing support of the World Resources Institute Markets and Enterprise Program and the Aqueduct Alliance. The authors would like to thank the following people for providing invaluable insight and assistance: Nicole Grohoski, Thomas Parris, Pragyajan Rai, Tianyi Luo, Robert Kimball, Andrew Maddocks, Tien Shiao, Betsy Otto, Charles Iceland, Daryl Ditz, and Kirsty Jenkinson as well as Nick Price and Hyacinth Billings for graphic support and final editing. For their guidance and feedback during the development of the Aqueduct Water Risk Atlas Global Maps, the authors would also like to thank:

Robin Abell, World Wildlife Fund
James Bradbury, World Resources Institute
David Cooper, World Resources Institute
Martina Flörke, University of Kassel
Tom Gleeson, McGill University
Cy Jones, World Resources Institute
Jaap Kwadijk, Deltares
Oliver Maennicke, World Wildlife Fund
Carmen Revenga, World Wildlife Fund
Katie Reytar, World Resources Institute
Mindy Selman, World Resources Institute
Justin Sheffield, Princeton University
Richard Vogel, Tufts University
Yoshihide Wada, Utrecht University
Hessel Winsemius, Deltares
Lijin Zhong, World Resources Institute