Using natural experiments and counterfactuals for causal assessment: River salinity and the Ganges water agreement

Gopal Penny[1,2], M. Shahjahan Mondal[3], Subir Biswas[3], Diogo Bolster[1,2], Jennifer L. Tank[2,4], Marc F. Müller[1,2]

[1]Department of Civil & Environmental Engineering & Earth Sciences, University of Notre Dame, USA
[2]Environmental Change Initiative, University of Notre Dame, USA
[3]Institute of Water and Flood Management, Bangladesh University of Engineering and Technology, Bangladesh
[4]Department of Biological Sciences, University of Notre Dame, USA

Key Points:

- Natural experiments and counterfactuals are promising tools for causal assessment
- We generate counterfactuals to assess the effect of the Ganges water treaty on streamflow and salinity
- The Ganges water treaty plays a modest yet important role reducing salinity in the delta in Bangladesh

Corresponding author: Gopal Penny, `gopalpenny@gmail.com`

Abstract

The effect of environmental policy on water resources is often challenging to evaluate due to dynamic interactions between people and water, particularly in data-scarce watersheds. Increasing interactions between society and hydrology present a need to understand causal relations for improved assessment and prediction in complex human-water systems. Conventional approaches to causal assessment in hydrology are sometimes insufficient due to data-scarcity or system complexity. We argue that natural experiments present a promising and complementary avenue for assessing causal relations in such systems. In this spirit, we exploit a natural experiment to assess causal effects of the Ganges water treaty, between India and Bangladesh, on streamflow and channel salinity in the Ganges delta in Bangladesh. We apply causal inference to assess the effect of the treaty on streamflow and use the results to generate synthetic ensembles of streamflow and salinity under a realistic scenario (with the treaty) and a counterfactual scenario (without the treaty). We then use synthetic streamflow ensembles to model salinity ensembles. The Ganges water treaty increased dry season streamflow in Bangladesh by approximately 18%, and decreased channel salinity by approximately 10%. The treaty has the greatest effect on salinity in Bangladesh in the driest years, but the overall effect is small compared with natural variability. We show that our approach accounts for natural hydrologic variability to accurately assess the causal effect of the treaty, compared with a naive approach which greatly overestimates the effect. This research demonstrates the value of natural experiments for causal assessment in coupled human-water systems.

1 Introduction

Human society has extensively modified rivers throughout the world [*Ceola et al.*, 2019; *Vörösmarty et al.*, 2010], and incorporating anthropogenic interactions with the water cycle remains a major challenge in hydrology [*Sivapalan et al.*, 2012; *Gleeson et al.*, 2019]. Many recent efforts have focused on understanding dynamic interactions and feedbacks between humans and water resources with the aim of developing new insights into the resilience, trajectories, and co-evolutionary behavior of coupled human-water systems [e.g., *Montanari et al.*, 2013; *Pande and Sivapalan*, 2016; *Penny and Goddard*, 2018]. Integral to that effort is the need to identify and evaluate the causal mechanisms that produce hydrologic change in heavily modified catchments [*Ehret et al.*, 2014]. This task is particularly urgent in regions undergoing rapid change and in need of an adequate policy response, often despite very limited observational data [*Sivapalan et al.*, 2014; *Thompson et al.*, 2013]. Accurately framing the causes of hydrologic change in these regions is critical to evaluate water resources interventions and predict water availability amidst rapid social and environmental change [*Wine*, 2018; *Srinivasan et al.*, 2016]. Causal assessment is therefore a critical task for hydrology but is fraught with challenges for a variety of reasons including the complexity of hydrologic processes [*Savenije*, 2009], limited observational data [*Hrachowitz et al.*, 2013], and the issue of equifinality whereby an observed outcome could result from multiple, unobserved causal pathways [*Beven*, 2006]. These challenges are particularly vexing for causal assessment when considering scales pertinent to water management which must contend with considerable variability of climatic drivers, complexity of hydrologic processes, and landscape heterogeneity.

To illustrate these challenges, consider a simple example in which we want to estimate the effect of a river diversion, beginning at time t_0, on streamflow downstream of the diversion (Figure 1). This generic example is non-trivial if the diversion behavior is unknown, a situation that is analogous to the case study of the Ganges river, introduced in Section 2.1. A simple yet naive approach to this problem would be to compare observed flow regimes downstream of the diversion before and after the t_0. Such an approach is "naive" because it does not account for confounding factors or processes, such as changing rainfall or upstream water use, that could affect the flow regime and bias the results.

An ideal experiment would circumvent the problem of confounding variables by removing biases between treatment and control groups. Indeed, this is commonly done in other fields (e.g., medical studies) through the use of randomized control trials, but a similar approach in hydrology is unlikely to succeed. In the example of the river diversion, *if* the days in which the diversion operates could be randomly assigned, the difference in flow observations between the treatment (diversion is ‘on’) and control (diversion is ‘off’) could be interpreted as the causal effect of the diversion because all unobserved confounding factors would be controlled for (or ‘averaged out’) through randomization [see *Angrist and Pischke*, 2008]. Unfortunately, such an approach is generally impractical in real water systems.

Because of the challenges noted above, hydrologists often consider causal effects in terms of *counterfactuals*: hypothetical scenarios of what would have occurred under causal conditions that differ from reality [*Lewis*, 1973, 2004]. The counterfactual theory of causation asserts that the causal effect of a driver is the difference between the actual outcome and the hypothetical outcome of the counterfactual scenario in which the driver takes on a different value or is excluded entirely. Counterfactuals are commonly used in hydrology to assess causal relationships, typically through observing or modeling alternate scenarios.

For instance, in the river diversion scenario, the counterfactual for downstream flow represents what downstream flow would be if the diversion did not exist. This counterfactual could be determined from streamflow records directly upstream and downstream of the diversion *if* such records are available because, absent the diversion, the two would be equal (Figure 1a). In more complex scenarios, counterfactuals could be similarly constructed from observations using results from paired catchment studies [*Brown et al.*, 2005], space-for-time substitution [*Wagener et al.*, 2010], or large-sample hydrology [*Gupta et al.*, 2014]. However, observations of policy-relevant variables in hydrology are often challenging to obtain [see *Hrachowitz et al.*, 2013], which motivates the use of models to bridge observation gaps [*Müller and Thompson*, 2019] and assess causal relationships [*Beven*, 2012].

Models are commonly used in hydrology to generate counterfactuals and assess causation in a process of *predictive inference* [*Ferraro et al.*, 2019]. In predictive inference, models are created and calibrated to reproduce observations, after which drivers are manipulated to assess the causal effect on outcome variables. Continuing with the river diversion example, a counterfactual for downstream flow could be created by using a hydrologic model to simulate streamflow in the absence of the diversion (Figure 1b), calibrating on downstream observations before the diversion takes effect. The difference between the model (the counterfactual) and observations (after t_0) is the causal effect of the diversion. Numerous approaches are available for generating plausible counterfactuals in hydrology [e.g., see *Stedinger and Taylor*, 1982].

Despite its common application, however, causal assessment through predictive inference can be challenging to implement given the complexity of hydrologic processes and the issue of equifinality [e.g., see *Savenije*, 2009]. For instance, observational data might be inadequate to sufficiently calibrate and validate the critical processes in a hydrologic model. The equifinality thesis asserts that limited data not only affects the uncertainty of parameters, but can also lead to uncertainty in the dominant hydrologic processes [*Beven*, 2006]. Furthermore, in intensively modified catchments, these concerns are amplified by the fact that human decision-making and water resources are mutually dependent and potentially nonstationary [*Sivapalan and Blöschl*, 2015]. A model calibrated before the diversion begins may not be valid in the presence of the diversion, complicating predictive inference.

As an alternative, we advocate the use of natural experiments to assess causal relationships and generate plausible counterfactuals. Natural experiments differ from traditional approaches to causal assessment in hydrology in that they seek to mimic ideal experiments

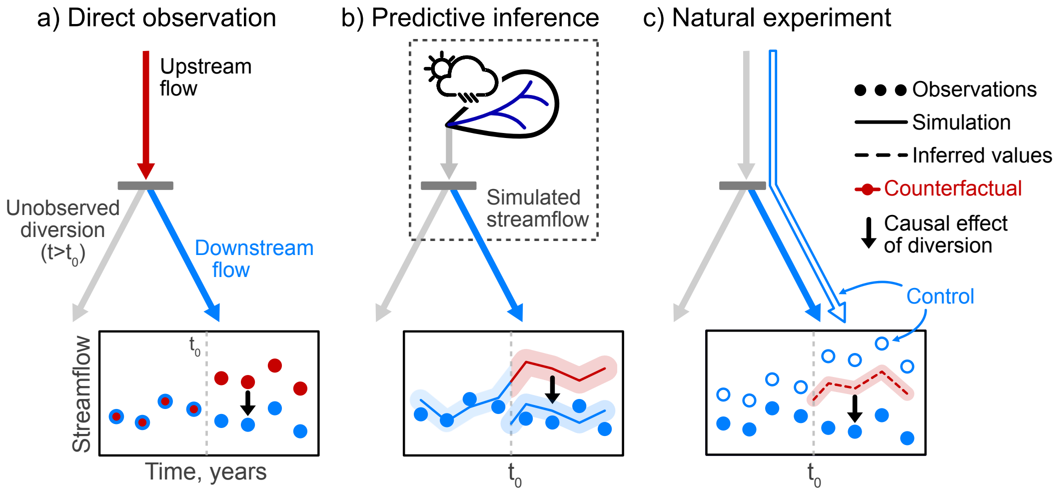

Figure 1. Causal assessment of a river diversion through the use of counterfactuals determined from (a) direct observation, (b) predictive inference, and (c) natural experiments. The causal effect of the diversion (black arrows) is the difference between downstream flow (light blue) and the counterfactual (red). In this simple case, the counterfactual of downstream flow could be taken directly from the observations of upstream flow *if* there is no selection bias before and after t_0 (direct observation). If selection bias exists or such data are not available, the counterfactual must be simulated (predictive inference) or inferred through the use of a control (natural experiment). We argue that natural experiments merit increased consideration in hydrology because they can offer valuable advantages over direct observation and predictive inference, which are more commonly applied to hydrologic problems.

by leveraging known characteristics (e.g., events, rules, or processes) of the investigated system, in order to identify treatment and control groups in a manner that is *as good as* random [see *Angrist and Pischke*, 2008]. In the river diversion example, a natural experiment might arise from the arbitrary nature of diversion rules (Figure 1c). If the diversion is turned 'on' and 'off' following rules that are unrelated to streamflow and its drivers, the outcome is arguably equivalent to that of the randomized experiment described above, where the diversion is turned 'on' and 'off' at random. Natural experiments require that the observational data meet specific criteria (i.e., an external and plausibly exogenous forcing), which sometimes presents obstacles to their implementation. Nevertheless, natural experiments have long been used for policy assessment in social and economic sciences [see *Angrist and Pischke*, 2008] and evaluation of disturbances in ecology [*Stewart-Oaten et al.*, 1986]. This type of approach is also emerging in studies of coupled human-water systems [*Müller and Levy*, 2019] including transboundary water resources [*Müller et al.*, 2016] and water quality [*Sigman*, 2005; *Keiser and Shapiro*, 2019], agricultural water use [*Deryugina and Konar*, 2017], effects of global trade on water and nutrient use [*Dang and Konar*, 2018; *Dang et al.*, 2018], and hydrologic consequences of land use change [*Levy et al.*, 2018].

Here we use a natural experiment to attribute hydrologic change in the context of the Ganges Water Agreement between India and Bangladesh [*GOB and GoI*, 1996]. The agreement was signed in 1996 and allocates Ganges river flow to India and Bangladesh during the dry season when streamflow is essential to both countries [*Hossain*, 1998]. Dry season streamflow is insufficient to meet the needs of both India and Bangladesh, and future freshwater supply in the region will likely be further threatened due to increasing competition for irrigation supply and increasing abstraction in the upstream watershed [*Mukherjee et al.*, 2018]. The agreement expires in 2026, creating an immediate policy need to evaluate its causal effect on both streamflow and salinity. Such an evaluation is hampered by the complexity and heterogeneity of socio-hydrologic processes throughout the basin and by the absence of data directly upstream of the diversion. These obstacles make it challenging to construct a counterfactual, either explicitly from observations (Figure 1a) or based on a fully calibrated hydrologic model (Figure 1b). Instead, we utilize a natural experiment to infer the causal effect of the treaty (similar to Figure 1c) by leveraging specific features of the agreement, which preferentially allocates flow to India and Bangladesh in six alternating 10-day periods, and properties of streamflow, including high autocorrelation in dry season. To illustrate the importance of careful treatment of causal relations, we also conduct a naive assessment by estimating the effect of the treaty as the difference in streamflow before and after the treaty was signed.

We use the results from the natural experiment to construct an ensemble of counterfactual scenarios of streamflow (Sections 2.3) and river salinity (Section 2.4), and evaluate these outcomes in relation to natural variability (Sections 3.1 and 3.2). We also show that the naive approach would considerably overestimate the effect of the treaty compared with our quasi-experimental approach. We conclude by evaluating results and discussing policy implications (Section 4), and contextualizing the potential benefits of natural experiments and counterfactuals for causal assessment of coupled human-water systems (Section 5).

2 Methods

2.1 The Ganges water sharing agreement

The Ganges river originates in the Himalayas of India and Nepal before flowing through the Gangetic plain and arriving at Farakka barrage, where water is diverted into the Hooghly distributary which flows south towards Kolkata. The Ganges continues east into Bangladesh, joining the Brahmaputra and Meghna rivers, and finally discharges into the Bay of Bengal (Fig. 2). The massive extent of the watershed encompasses considerable social and hydrologic complexity spanning multiple cultures and hydrologic regimes, with 600 million people over an area of 1 million km^2 and 520 km^3 of streamflow per year on average. The hydrology exhibits strong seasonality, with precipitation and streamflow driven by the summer Monsoon (June–September) and followed by an extended dry season (January–May).

Flow availability in the Ganges is essential in both India and Bangladesh and the transboundary river has been the subject of regional disputes for at least 65 years [*Hossain*, 1998]. Dry season flow has immediate effects on salinity, health, and economic well-being in both countries [*Hossain*, 1998]. In India, flow in the Hooghly has important consequences for Kolkata including siltation of the port [*Ray*, 1978] and salinity in drinking water [*Ganguly and Roy*, 2018]. In Bangladesh, salinization poses great threats to health and livelihoods of delta inhabitants [*Salehin et al.*, 2018; *Rahman et al.*, 2019], with much of the population at risk of hypertension from excessive salt intake [*Talukder et al.*, 2016], agrarian livelihoods in danger [*Mondal et al.*, 2019], the largest mangrove forest in the world, the Sundarbans, at risk of disappearance, and salinization strongly associated with forced internal migration [*Chen and Mueller*, 2018].

Beginning in 1975, a portion of Ganges flow has been diverted at Farakka barrage into the Hooghly river to de-silt and sustain the port of Kolkata in India [*Mirza*, 2004]. Operation of the barrage was governed by temporary agreements through 1988, followed by nearly 10 years absent agreement until the 1996 treaty was signed into effect for a period of 30 years [*Hossain*, 1998]. The 1996 treaty (hereafter referred to as *the treaty*) covers the January 1–May 31 period each year, with specific allocation requirements during six 10-day allocation periods March 11 – May 10 (Table 1). As the treaty is due for renewal in 2026, a quantitative estimate of the specific (causal) effect of the treaty on historical flow and salinity of the Ganges is relevant and timely and will inform future predictions.

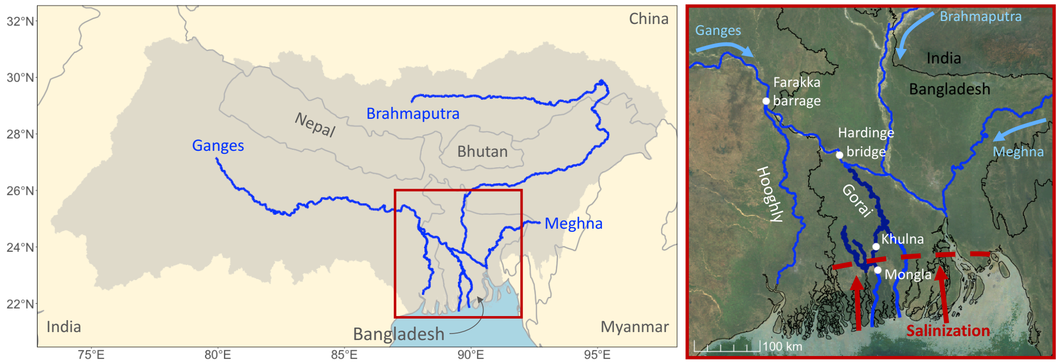

Figure 2. Ganges watershed map. (left) The delta receives inflow from the Ganges, Brahmaputra, and Meghna rivers. (right) Since the 1970s, a portion of Ganges water has been diverted at Farakka barrage into the Hooghly river. Salinization due to low delta inflow is especially concerning in southwestern Bangladesh. Data used in this study include streamflow records at Hardinge bridge and salinity records at Khulna and Mongla stations. Dark blue channels indicate the extent of the HEC-RAS hydrodynamic mixing model (Section 2.4).

Table 1. Ganges agreement flow allocation.

Inflow to Farakka	Share of India	Share of Bangladesh
January 1 – May 31 (except allocation periods)		
<70,000 cusec	50%	50%
70,000-75,000 cusec	Remainder	35,000 cusec
>75,0000 cusec	40,000 cusec	Remainder
Alternating 10-day allocation periods (March 11 – May 10)		
Bangladesh allocation period	Remainder	35,000 cusec
India allocation period	35,000 cusec	Remainder

A *cusec* is a cubic feet per second ($0.028 m^3 s^{-1}$)

2.2 Data Availability

At the time of this study, the availability of salinity and streamflow records were sparse, especially along the main stem of the Ganges. As records of streamflow at Farakka barrage were not available, we relied primarily on streamflow records at Hardinge bridge in Bangladesh, 160 kilometers downstream of the barrage (Table 2). These daily streamflow observations were not continuous, and in five years the dry season (January 1–May 31) contained at least 20 consecutive days without any observational data. The remaining years contained a maximum of three consecutive missing days within the dry season. The years with 20 or more consecutive missing days were included in our causal assessment of streamflow, but removed from the synthetic streamflow dataset (see Section 2.3).

We had access to salinity records at Khulna station and six additional stations comprising the boundaries to the HEC-RAS model, including Mongla station (Table 2). We had concerns over data integrity because the data contained unexpected extreme salinity (low or high) in some years, exhibited a variable relationship between salinity and streamflow, and, in 2014, salinity at Khulna station (upstream) exceeded salinity at Mongla station (downstream). However, these datasets were the only available resource for *in situ* salinity mea-

Table 2. Data availability.

Location[1]	Measurement	Availability
Hardinge bridge	Streamflow	1985-2016[2]
Khulna station	Salinity	1988-2017[3]
Mongla station	Salinity	2001-2011[4]
HEC-RAS boundaries (5)	Salinity	2014, 2017

[1] All data obtained from the Bangladesh Water Development Board.
[2] Years 1997, 2001–2003, and 2013 contained 20 or more consecutive missing days in dry season.
[3] Missing data from 1990, 1994, 2000, 2004, and 2009-2012.
[4] Missing data from 2004.

surements and, in general, they exhibited self consistency such that salinity increased and peaked in dry season and variability was within allowable tolerance. We note that the relationship between salinity and streamflow exhibited a notable change in 2013–2015 (see Section S2, Figure S2). Because this change was unlikely related to the 1996 treaty, we focused our salinity analysis on years prior to 2013 wherever possible.

2.3 Causal inference and streamflow counterfactual

In order to generate a counterfactual and assess the effect of the Ganges treaty on streamflow, we employ a natural experimental approach similar to *difference-in-differences* designs used by economists [*Angrist and Pischke*, 2008] and Before-After-Control-Impact (BACI) studies used by ecologists [*Stewart-Oaten et al.*, 1986; *Underwood*, 1994]. This approach is used for causal inference in a particular type of natural experiment where a treatment (here, the treaty) is imposed on only one of two comparable groups at a particular point in time [see *Angrist and Pischke*, 2008; *Müller and Levy*, 2019]. The causal effect of the treatment ($\Delta X_{treatment}$) is evaluated as:

$$\Delta X_{treatment} = (X_{T,post} - X_{T,pre}) - (X_{C,post} - X_{C,pre}),$$

where X is the considered outcome, averaged across observation units, for the treatment (T) and control (C) groups before (pre) and after ($post$) treatment. In effect, the change in the control group ($X_{C,post} - X_{C,pre}$) controls for the change that would have occurred in the treatment group in the absence of treatment. The critical assumption is that observation units in the treatment group would have followed the same *trajectory* as the control group if the treatment had not been applied. This requirement is formalized in the stable unit treatment value assumption (SUTVA), which stipulates two components: (i) no interference among units where the treatment of some units affects the treatment of others, and (ii) no hidden treatments or unobserved variables affecting the potential outcomes of the experimental units. The ideal control would therefore capture all non-random variability and nonstationarity within the system and allow analysis to proceed as if the units of observation were randomly assigned to the treatment or control category, such that the only difference between groups is the effect of the treatment itself.

To evaluate the Ganges treaty (i.e., the treatment), we assigned daily streamflow records (the evaluated outcome) into treatment and control groups. Flows during allocation periods (March 11–May 10) were assigned to the treatment group to evaluate the average effect of the treaty, and the control group was comprised of streamflow records 20 days before (Feb 19–Mar 10) and after (May 11–May 31) the allocation periods. The SUTVA assumption requires that the March 11 and May 10 dates in the treaty are arbitrary and exogenous, meaning that they are not, themselves, determined by any hydrologic considerations at the relevant time scales. In other words, it should make no difference from a hydrology perspective that the the bounds fall on any particular day around March 11 and May 10. Evidence from historical negotiations suggests that political considerations (not hydrologic ones) were invoked when determining these dates [*Hossain*, 1998], and streamflow recession analysis indicated similar trajectories of dry-season streamflow before and after the implementation of the treaty (see SI, Section S1).

Formally, we operationalize the difference-in-differences approach and estimate the effect of the agreement on streamflow with the linear regression

$$\ln Q = \alpha_0 + \alpha_\tau D_\tau + \alpha_\gamma D_\gamma + \alpha_{treat} D_\gamma D_\tau + \varepsilon \tag{1}$$

where $\ln Q$ is the logarithm of daily streamflow observations, D_τ is a 'dummy' variable equal to 1 for observations after 1996 (i.e., after the treaty was signed) and 0 otherwise, D_γ is 1 for observations in the treatment group and 0 for the control group, and ε is the error term. The regression coefficients represent the intercept (α_0), the Time Effect (α_τ), the Group Effect (α_γ), and the treatment effect (α_{treat}). We specify two separate regressions to assess the *causal* effect of the treatment: (1) a specification as described in Equation 1, and (2) a similar specification that differentiates the effect of the agreement during the India and Bangladesh allocation periods, so that the group term consists of two dummy variables ($D_\gamma = \{D_{\gamma,IN}, D_{\gamma,BD}\}$) with distinct regression coefficients. The error term ε in Equation 1 represents regression errors, which we assume are uncorrelated with the treatment but correlated within each year. We therefore clustered errors by year, such that the variance covariance matrix Σ_{ij} takes a value of σ^2 if i and j were observed during the same year and 0 otherwise. Regression coefficients α_i were estimated using ordinary least squares (OLS) and clustered standard errors σ^2 using the 'sandwich' approach [specifications 1 and 2, see *Williams*, 2000]. The causal effect of the treatment on streamflow was assessed by considering the sign, amplitude, and statistical significance (student-t test) of the regression coefficient α_{treat}.

In order to demonstrate the value of the difference-in-differences causal inference approach, we estimated the effect of the treaty in two additional regression specifications using a *naive* approach in which we excluded data from the control period and used only data from the allocation periods. The regression output in these cases represents a simple comparison of streamflow before and after the treaty was signed. These regressions included (3) a specification with only the time effect, $\ln Q = \alpha_0 + \alpha_\tau D_\tau$, and (4) a specification that captured the before-after comparison in the differentiated group periods, $\ln Q = \alpha_0 + \alpha_\gamma D_\gamma + \alpha_{treat} D_\gamma D_\tau$, with $D_\gamma = \{D_{\gamma,IN}, D_{\gamma,BD}\}$. These naive specifications 3 and 4 mimic the causal specifications 1 and 2, respectively, but cannot be used for causal assessment because they do not account for unobserved drivers through the use of a control.

Lastly, we used a similar framework to generate 1000 realizations of synthetic streamflow time series with the agreement (the "treated" scenario) and without the agreement (the "counterfactual" scenario). For this purpose, we defined a fifth regression specification (5) which included the regression variables from specification (2) as well as year fixed effects to control for interannual climate variations, and quadratic calendar days to control for seasonal variability. The latter was modeled as a parabola that captures the stream recession in dry season and recovery at the beginning of wet season. Because the focus is now on prediction, these added controls increase prediction accuracy while preserving the causal inference results from the (more parsimonious) specification (2). Coefficients from specification (5) were estimated using generalized least squares with cluster-robust standard errors at the annual level [*Cameron et al.*, 2011]. Coefficient estimates were then used to generate 1000 random realizations of synthetic streamflow timeseries at Hardinge bridge, with the agreement ($\{\alpha_{treat,IN}, \alpha_{treat,BD}\} \neq 0$) and without the agreement ($\{\alpha_{treat,IN}, \alpha_{treat,BD}\} = 0$) (see Appendix A for details). These two ensembles excluded years with 20 or more consecutive missing days (Table 2). We used both ensembles of synthetic streamflow time series as inputs to the salinity model to assess the effect of the agreement on river salinity (Section 2.4).

2.4 Salinity model and counterfactual

The difference-in-differences approach taken for streamflow (Section 2.3) was not amenable to salinity because we were unable to identify an adequate control period that would adhere to SUTVA. River salinity in the delta accumulates throughout the dry season, so that salinity on any given day is the cumulative effect of complex mixing processes within the delta over the course of the dry season. Selecting an adequate control would be necessary to adhere to the "no hidden treatments" assumption of SUTVA, but selection of such a control would likely violate the "no interference" assumption because the rate of salinity accumulation depends on both streamflow levels and salinity gradients, which complicates any definition of treatment and control groups. For these reasons, we utilized a more conventional predictive inference approach to understand the effect of the treaty on salinity, building a HEC-RAS transport model [*USACE*, 2016] to simulate salinity within a portion of the delta. This approach relies on the intuition that the Ganges treaty *only* affected river salinity through its effect on streamflow. The cost of such an approach is that we must assume that the model sufficiently captures the dynamic processes and boundary conditions that affect salinity so that the model is accurate even when applied to the hypothetical counterfactual scenarios.

The natural experimental approach for streamflow and the predictive inference approach for salinity were complimentary. The synthetic streamflow data generated from regression specification (5) were used to generate synthetic salinity within the delta, and both synthetic datasets (streamflow and salinity) were necessary boundary conditions for the HEC-RAS model. We used a one-dimensional mixing model to generate the synthetic salinity data from streamflow data. The model included discharge and first order exchange to simulate salinity concentration (C) at Khulna station (where salinity data is available) as a function of streamflow at Hardinge Bridge (Q_H):

$$\frac{dC}{dt} = -aQ_H C + be^{-dQ_H}(C_D - C), \tag{2}$$

where $C_D = 35ppt$ is downstream (ocean) salinity and parameters a, b, and d are calibration constants. The two terms on the right-hand side of the equation represent discharge and first order exchange, respectively. The exchange term be^{-dQ_H} decreases (or increases) with higher (or lower) streamflow as the channel becomes dominated by inflow (or tidal oscillations).

The mixing model was calibrated on Khulna salinity across all years through 2008 in which data were available and nearly continuous, both for streamflow at Hardinge bridge and salinity at Khulna station (Table 2). Our calibration therefore captures the period in which dredging of Gorai channel was initiated [beginning in 1998, *de Groot and van Groen*, 2001], but not any changes in the relationship between Hardinge streamflow and Khulna salinity occurring after 2008 (see Section 2.2). The physical model was validated using a leave-one-out approach across the same years used for calibration (see Section S2 for details). The validated physical model was forced by the ensembles of synthetic streamflow (treated and counterfactual) to simulate corresponding ensembles of salinity at Khulna, and these ensembles were extrapolated to salinity at the boundaries of the HEC-RAS domain (Appendix B).

The HEC-RAS model was calibrated manually in two steps, using only observed data from 2014 and 2017 (see Section S3 and Figure S3 for further details). The hydrodynamic model was first calibrated by fitting Manning's roughness coefficient using observed water level in the Gorai channel (see Fig. 2), with a reasonable Nash Sutcliffe Efficiency (NSE=0.81). The dispersal coefficient of the salinity mixing model was then calibrated using observed salinity data at Khulna station (NSE=0.82). These calibrated coefficients were used to simulate the spatial and temporal distribution of salinity, as forced by the ensemble-median realization of synthetic streamflow and salinity for the treaty and counterfactual scenarios.

3 Results

3.1 Causal effect of the agreement on streamflow

Table 3 reports the regression results describing the effect of the Ganges treaty on streamflow, with columns 1–2 corresponding to the causal difference-in-differences regression specifications, columns 3–4 corresponding to the naive specifications, and column 5 corresponding to the specification used to generate synthetic data, as described in Section 2.3. The effect of the treaty is captured by the Treatment Effect, except in column 3 which evaluates the effect of the treaty as the post-treaty Time Effect. Because the dependent variable was specified as $\ln Q$, the estimated coefficient for each effect can be interpreted as a percent change (divided by 100) in streamflow during the considered period due to the treaty.

The Ganges treaty resulted in an 8.2±7.4% (Mean ± Standard error) increase (statistically insignificant) in streamflow in the combined allocation periods (Table 3, Column 1). When differentiating between India and Bangladesh allocation periods, the treaty produced an 18.5±7.6% increase (statistically significant) in the Bangladesh periods and a -1.6±8.5% decrease (statistically insignificant) in streamflow in the India periods (Column 2). The effect of the treaty is similar in the regression with fixed effects (Column 5). This suggests that our results are robust to interannual (e.g., climate driven) streamflow variability. By comparison, the naive approach exaggerates the effect of the treaty, estimating a 36% increase in streamflow in the combined allocation periods (Column 3) and 46% and 26% increases in the Bangladesh and India allocation periods, respectively (Column 4).

In specifications 1 and 2, the Time Effect corresponds to the increase in streamflow in the control period before and after the treaty was signed. In other words, streamflow increased by 27.8±10.9% (statistically significant) in the control period, which explains the discrepancy between the causal and naive estimates. The naive specifications estimate the *net* change in streamflow in the allocation periods before and after the treaty, whereas the causal specifications estimate the change in streamflow *relative* to the control (e.g., 46.3% - 27.8% = 18.5% increase in the Bangladesh allocation period). The increase in streamflow during the control period reflects the non-stationary hydrology of the river and has important implications for the interpretation of results, which we discuss further in Section 4.1.

In both causal specifications (Table 3, columns 1–2), the Group Effect (α_γ) was negative. This is consistent with the decline in streamflow during the allocation periods which coincided with the driest part of the year. The group effect is diminished when fixed effects are included (Column 5), because the inclusion of Day and Day^2 captures most of the dry season recession and recovery of the hydrograph.

The regression specification with fixed effects (Table 3, Column 5) was used to generate synthetic streamflow with the agreement ("treated" scenario) and without the agreement ("counterfactual" scenario). Synthetic streamflow reproduced observed intra-annual streamflow variability with an R^2 coefficient of 0.75, suggesting that the model captures the most relevant sources of variability. The synthetic 95% prediction interval contained nearly all observations (see Figure 3, blue). Similarly, the synthetic counterfactual data (without treaty) reproduced observed streamflow variability outside of the allocation periods (Figure 3, red). The regression model used to generate synthetic streamflow explicitly captured interannual climate variations by controlling for year fixed effects (Figure 4). As a consequence, synthetic flow data reproduced annual flow volumes during the Bangladesh and India allocation periods (Figure 4, blue line vs. black dots). In contrast, counterfactual (without treaty) flows clearly deviated from observations during the Bangladesh allocation periods (Figure 4, red line vs. black dots), indicating the effect of the treaty. Note that these deviations were greater when observed streamflow was lower, highlighting that the minimum allocation requirement in the agreement had the strongest effect in the driest years. Results also suggest that the net effect of the agreement (given by the difference between long-term mean values of the syn-

Table 3. Linear regressions: Effect of the Ganges treaty on downstream flow.

		Dependent variable:				
		Log streamflow at Hardinge Bridge (Bangladesh)				
		Causal		Naive		Synthetic
		(1)	(2)	(3)	(4)	(5)
Treatment Effect	BD Alloc. Period		0.185**		0.463***	0.208***
(Ganges treaty)			(0.076)		(0.131)	(0.074)
	IN Alloc. Period		−0.016		0.262*	−0.001
			(0.083)		(0.146)	(0.082)
	Combined	0.082				
		(0.074)				
Group Effect	BD Alloc. Period		−0.343***			−0.086*
			(0.065)			(0.051)
	IN Alloc. Period		−0.297***		0.045**	−0.082
			(0.069)		(0.019)	(0.054)
	Combined	−0.320***				
		(0.066)				
Time Effect	(post-treaty)	0.278**	0.278**	0.360***		0.816***
		(0.109)	(0.109)	(0.136)		(0.044)
Intercept		6.855***	6.855***	6.536***	6.513***	8.420***
		(0.091)	(0.091)	(0.115)	(0.118)	(0.232)
Covariates		None	None	None	None	Fixed effects‡
Observations		2,623	2,623	1,616	1,616	2,623
R^2		0.215	0.226	0.163	0.181	0.750

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$
Standard errors were clustered by year in all regressions
‡Regression covariates included fixed effects for year, day, day^2

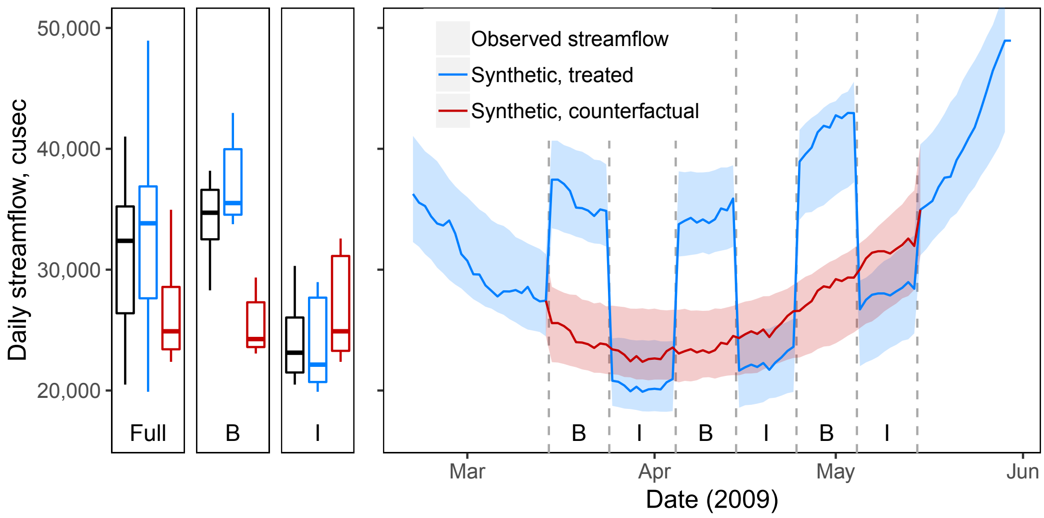

Figure 3. Dry season streamflow timeseries at Hardinge bridge in 2009. The shaded areas indicate 95% prediction intervals around daily streamflow. The high and low 10-day periods follow the 10-day allocation periods alternating between India and Bangladesh. Boxplots show within-year variability of daily data during the full March 11–May 10 period and for Bangladesh and India allocation periods, respectively.

thetic streamflow with and without the agreement) was notably smaller than the variability in annual streamflow (Figure 4).

3.2 Causal effect of the agreement on salinity

Synthetic flows with the agreement gave rise to salinity levels that were smaller than those caused by counterfactual synthetic flows (without the agreement). The agreement reduced salinity levels by 9.8% for the entire allocation period (March 11–May 10), and by 13.1% at the end of the allocation periods (May 10). The absolute reductions in salinity were considerably larger when salinity was higher. These reductions resulted primarily from the slower rate of salinity intrusion during Bangladesh allocation periods. Although the allocation periods ended May 10, salinity typically peaked May 22–25 (Figure 5a). Similar to streamflow, the average effect of the agreement (10%) was considerably smaller than the overall interannual variability of peak salinity (Figure 5b).

The HEC-RAS model highlighted intra-annual and interannual variability in salinity intrusions from the Bay of Bengal (Figure 6). Freshwater discharge, (given by salinity ≤ 1 ppt) extended fully into the delta at the beginning of the treaty period, but receded considerably into the Gorai channel during the allocation periods. Brackish water (salinity=2 ppt) exhibited similar behavior, with sharp salinity gradients towards the end of dry season. An indicative drinking water salinity limit of 5 ppt intruded upstream of Khulna at the end of the allocation period and even further on the date of maximum intrusion (note that there is no official salinity guideline for drinking water [*Rahman et al.*, 2019]). The effect of the agreement on intrusion distance was small compared with the interannual variability in the allocation period and subsequent date of maximum intrusion (Figure 6b).

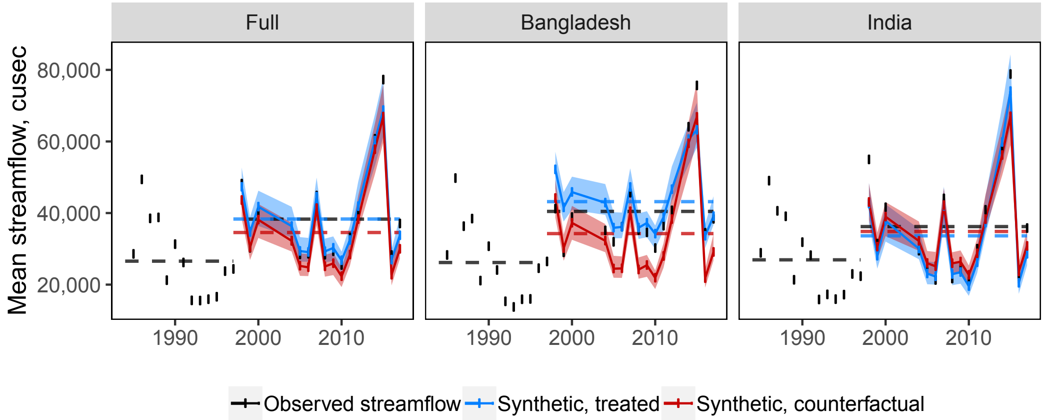

Figure 4. Dry season streamflow at Hardinge bridge including the annual average of daily flow from (a) the full period Mar 10–May 10, (b) Bangladesh allocation periods, and (c) India allocation periods. The thick dashed lines represent the long-term mean and shaded ribbons indicate the annual mean of the 95% daily prediction interval.

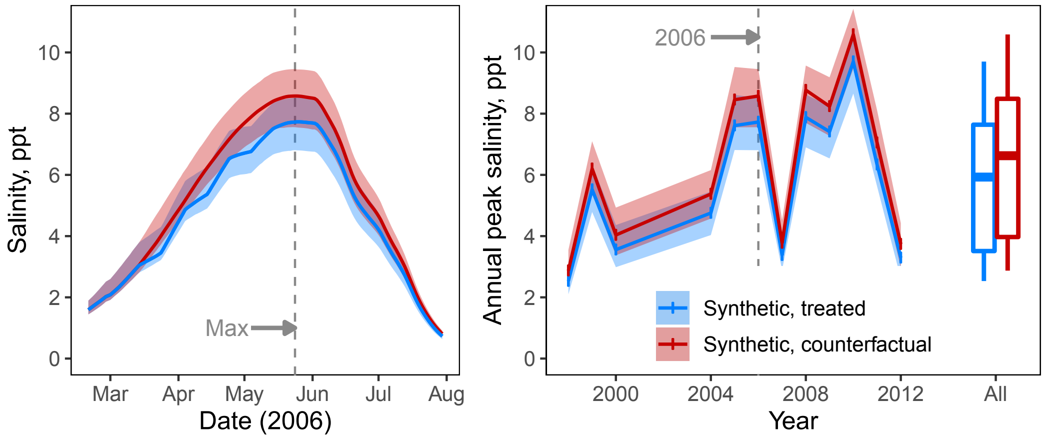

Figure 5. Delta salinization at Khulna station, with (treated) and without (counterfactual) the agreement. The change in salinity due to the agreement was less than natural variability, but greater when salinity was high due to low streamflow.

4 Discussion

4.1 Endogeneity and causal inference

Careful consideration of causal inference in an empirical context is necessary to isolate the effect of the agreement from potential confounding sources of hydrologic change. The effect of the treaty cannot be estimated by simply comparing streamflow before and after the treaty because any changes could be caused by confounding factors not associated with the treaty, such as precipitation or land use change. Indeed, such a naive approach overestimates the effect of the treaty by 28%, which is the change in streamflow within the control period.

The critical challenge to our empirical approach was the identification of an appropriate control period such that the control and treatment groups would have followed similar trajectories in the absence of the treaty in order to meet the criteria of SUTVA. The "no in-

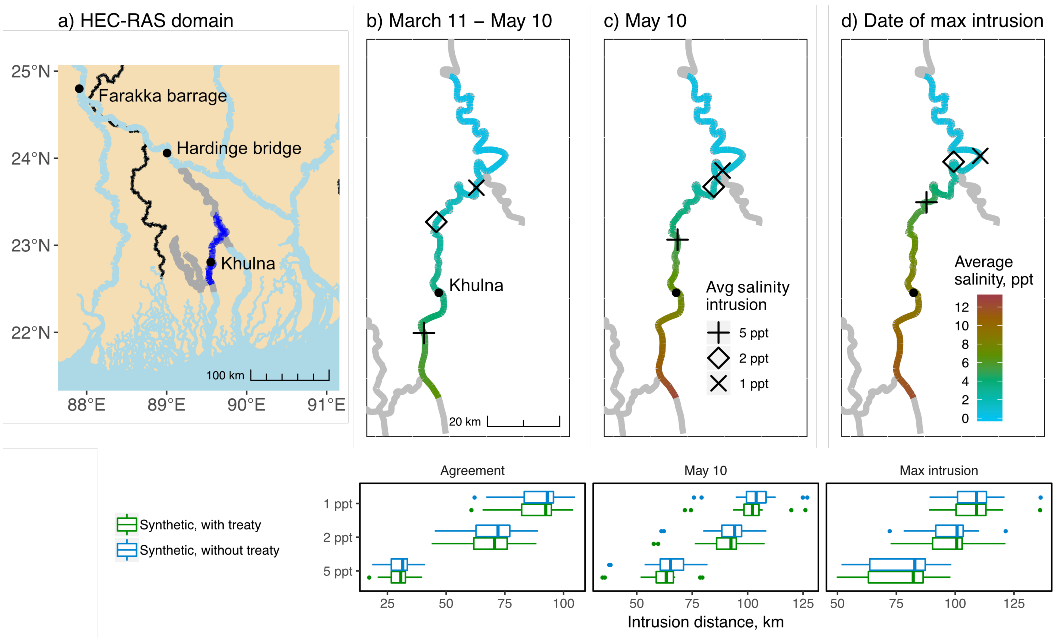

Figure 6. Maps of delta salinization (top) and boxplots of salinity intrusion distance (bottom). (a) HEC-RAS model domain (gray) and location of salinity profile maps (dark blue) situated in the Ganges-Brahmaputra-Meghna delta. (b) Average salinity during the allocation period, Mar 11–May 10, including intrusion of freshwater (1 ppt), brackish water (2 ppt), and suggested upper bound for drinking water (5 ppt). (c) Average salinity at the end of the agreement period, May 10. (d) Average salinity on the date of maximum intrusion.

terference" assumption is realized through clustering of daily streamflow observations within each year. The "no hidden treatments" assumption is affirmed by two arguments: (i) we assume India acts in its own interest and does not unnecessarily suppress its streamflow allocation at the start of dry season, which is supported by (ii) comparable dry season recession constants before and after the treaty, indicating similar diversion rates by India in the dry season before the allocation periods (i.e., January 1–March 10, see Table S2). The finding that dry season streamflow increased *independent* of the treaty is further justified by the fact that streamflow increased in both November and December by 37% (Table S1), *outside* the period governed by the treaty (January 1–May 31). The statistically significant increase in streamflow in the control period since the signing of the treaty (Time effect in Table 3, columns 1–2) likely results from an observed increase in precipitation in October and earlier snowmelt in May, both of which are unrelated to the treaty. In other words, the control period functions exactly as it should by isolating the agreement from other unobserved, nonstationary hydrologic processes.

4.2 Data scarcity and non-stationarity

Empirical assessment of salinity in this scenario is impractical due to limited availability of high-integrity data, pushing us towards a modeling approach to assess the causal effect of the treaty on salinity. The parsimonious salinity model attributed a 10% decrease in salinity to the agreement and the HEC-RAS model predicted a small decrease in salinity intrusion distances. Using streamflow to model the effect of the agreement on salinity was complicated by multiple challenges. The available data were sparse and irregularities within the dataset suggested possible concerns with data integrity which we were forced to accept because of lack of alternatives. Additionally, predictions are challenged by the fact that the Ganges delta is in constant flux due to erosion, sedimentation, and other anthropogenic processes including dredging, land use change, groundwater pumping, and aquaculture practices [e.g., see *Rahman et al.*, 2019; *Mondal et al.*, 2019].

In particular, validation of the mixing model (Section S2) suggests a nonstationary relationship between streamflow at Hardinge bridge and salinity at Khulna station, most noticeable after 2013 (Figure S2). We exclude these years from our analysis but note the likelihood of non-negligible interannual variability in other years and the fact that the only available calibration data for five of the HEC-RAS boundaries came from 2014, a year in which high salinity at Khulna station exceeded downstream salinity at Mongla station. Although this increase could have been influenced by a variety of factors, understanding the causes of this variability is nonessential and the most important factor for the purpose of estimating the effect of the treaty is that we have adequately reproduced the effect of streamflow at Hardinge bridge on salinity within the delta. In this regard, our analyses indicate a reasonable fit between the model and predictions (see Section S2) suggesting that we adequately capture the signal of the agreement even if we do not reproduce the observed variability of salinity. Lastly, there is a possibility that our need to attenuate salinity at the HEC-RAS domain boundaries could have resulted in an underestimate of the effect on salinity within the HEC-RAS model, meaning that the results for salinity could be slightly conservative.

4.3 Policy implications

The agreement alone is unlikely to resolve the issue of salinization because there is not enough streamflow in dry season to prevent salinity intrusion in both the Hooghly and Gorai distributaries, especially in dry years. The effect of the agreement is considerably smaller than natural variability in the system, meaning that excessive salinity in some years will be followed by tolerable salinity in other years. Nevertheless, the agreement plays an important but non-dominant role in reducing salinity in the delta despite the fact that the agreement was not specifically designed to minimize salinization.

In the context of the expiration of the current agreement in 2026, upcoming negotiations should carefully consider the effect of the agreement on future salinity within the delta. From a research perspective, it will be very important to understand (i) how the agreement affects flow availability in conjunction with variable and nonstationary streamflow at Farakka barrage, (ii) the relationship between flow allocation and downstream salinity, and (iii) the effect of variability of streamflow and other hydrologic processes on salinity levels in the delta. The lack of available data to evaluate such relationships presents multiple methodological challenges. The research herein presents a solution to some of these challenges, but further research on (i-iii) guided by our work should be conducted to provide information to negotiators as they prepare to extend or re-negotiate the current agreement.

5 Conclusions

We estimated the causal effect of a policy change (the Ganges water treaty) on streamflow and salinity in a data-limited environment through the use of a natural experiment and synthetic counterfactual scenarios. Our analyses produced three main findings pertaining to the Ganges water treaty: (i) the treaty increased streamflow in Bangladesh during Bangladesh allocation periods by 18%, (ii) the increased streamflow produced a 10% decline in channel salinity in the delta, and (iii) the salinity intrusion distance was slightly reduced but natural variability greatly exceeded the effect of the treaty. Future research should explore how the agreement could be modified to better achieve the objectives of both countries in terms of flow availability and water quality.

The use of natural experiments is common in other disciplines (e.g., economics and ecology) but has rarely been applied within hydrologic sciences. The value of this approach can be understood by comparison with more conventional approaches to causal assessment in hydrology. Indeed, before settling on this methodology, we considered multiple alternative approaches to generating a counterfactual and assessing the effect of the Ganges treaty on streamflow. For instance, we could have used a naive comparison of streamflow before and after the agreement but, as we show, such an approach would have greatly exaggerated the effect of the treaty on streamflow. We had hoped to utilize direct observations of flow upstream and downstream of Farakka barrage, but were unable to obtain the necessary upstream data. We also considered using a hydrologic model to simulate streamflow in the Ganges river, but we had concerns about using such a large scale model to simulate dry season streamflow in the absence of calibration data upstream of Farakka barrage. In other words, the natural experiment we describe for assessing streamflow provided a solution where other, more common approaches would have had considerable shortcomings.

Accelerating interdependencies between society and hydrology create an imperative to understand the relationships between humans and water, including causal mechanisms driving change within and across systems. The predominant methodological approaches in hydrology sometimes fall short because of limitations of hydrologic modeling or scarcity of data [*Clark et al.*, 2015], and new empirical approaches are needed to assess nonstationarity in complex human-water systems. Studies that emphasize the use of natural experiments show considerable promise and can be assimilated with commonly accepted approaches in hydrology [e.g., observation and simulation, *Müller et al.*, 2017] as well as other approaches to causal assessment [e.g., the method of multiple hypotheses, *Srinivasan et al.*, 2015; *Penny et al.*, 2018]. Paradigms from causal inference that emphasize empirical design and analysis present a useful alternative for hydrologists to evaluate and attribute change.

A: Generating synthetic streamflow ensembles

Generating synthetic ensembles of streamflow at Hardinge bridge required a model that captured properties of the variability and autocorrelation of streamflow. Synthetic log-transform daily flows were drawn from the following distribution:

$$\ln Q \sim \mathcal{N}(\ln \hat{Q}, \Sigma_{GLS}), \tag{A.1}$$

where $\mathcal{N}(\cdot,\cdot)$ designates the normal distribution and $\ln \hat{Q}$ are deterministic predictions of the linear model in Equation 1 with and without ($\alpha_{treat} \equiv 0$) the treaty. The variance-covariance matrix Σ_{GLS} captures the serial correlation of streamflow at the daily time scale but neglects correlations at the interannual time scale:

$$\Sigma_{ij} = \begin{cases} \sigma^2 \rho^{|i-j|} & \text{if } i \text{ and } j \text{ observed in same year} \\ 0 & \text{otherwise} \end{cases}$$

The variance σ^2, autocorrelation coefficient ρ, and regression coefficients α_i used to generate synthetic streamflow were jointly estimated using generalized least squares (GLS). The GLS-estimated coefficients were finally used to generate an ensemble of 1,000 instances of synthetic streamflow time series for the 1985–2016 period, and a corresponding counterfactual without the treaty ($\alpha_{treat} \equiv 0$) using the model in Equation A.1.

B: Extrapolation to HEC-RAS domain boundaries

We assumed that salinity C evolves linearly along stream reaches ($C = ax + b$, where x is a distance along the stream) to obtain salinity at the six HEC-RAS model boundaries. We smoothed salinity at Khulna and Mongla station using a 30-day moving average, and then using a first-order linear regression to calibrate the coefficients (a, b) for each day. The coefficients were used to translate daily salinity at Khulna station to Mongla station, attenuating high salinity to ensure no values exceeded ocean salinity (35 ppt). We then translated salinity at Mongla to each of the other five HEC-RAS boundaries following the same approach but with a fixed ratio (i.e., $b = 0$). The simulated boundary salinity datasets (treated and counterfactual) were then used as time-varying boundary conditions in the HEC-RAS model. We validated this approach by comparing the original synthetic salinity at Khulna with the HEC-RAS model output at Khulna, with sufficient agreement (R^2=0.71). We also attempted calibrating Mongla station using a fixed ratio ($C = ax$), but this led to considerable over-estimates of salinity. Allowing a non-zero intercept ($b \neq 0$) dampened the the effect of the agreement but more accurately reproduced observed salinity at Mongla station.

Acknowledgments

This research has been conducted under a collaborative project between Bangladesh University of Engineering and Technology (BUET) and the University of Notre Dame, USA. Funding from the Keough School of Global Affairs, the Notre Dame Environmental Change Initiative, and Notre Dame Research is acknowledged. Penny, Bolster, and Müller acknowledge support from National Science Foundation under Grant No. ICER 1824951. Streamflow and salinity data used in this study are available for purchase from the Bangladesh Water Development Board.

References

Angrist, J. D., and J.-S. Pischke (2008), *Mostly harmless econometrics: An empiricist's companion*, Princeton university press.

Beven, K. (2006), A manifesto for the equifinality thesis, *Journal of Hydrology*, *320*(1-2), 18–36, doi:10.1016/j.jhydrol.2005.07.007.

Beven, K. (2012), Causal models as multiple working hypotheses about environmental processes, *Comptes Rendus Geoscience*, *344*(2), 77–88, doi: https://doi.org/10.1016/j.crte.2012.01.005.

Brown, A. E., L. Zhang, T. A. T. A. McMahon, A. W. Western, and R. A. Vertessy (2005), A review of paired catchment studies for determining changes in water yield resulting from alterations in vegetation, *Journal of Hydrology*, *310*(1-4), 28–61, doi: 10.1016/j.jhydrol.2004.12.010.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2011), Robust inference with multiway clustering, *Journal of Business & Economic Statistics*, *29*(2), 238–249.

Ceola, S., F. Laio, and A. Montanari (2019), Global scale human pressure evolution imprints on sustainability of river systems, *Hydrology and Earth System Sciences Discussions*, (May), 1–20, doi:10.5194/hess-2019-227.

Chen, J., and V. Mueller (2018), Coastal climate change, soil salinity and human migration in Bangladesh, *Nature Climate Change*, *8*(11), 981–985, doi:10.1038/s41558-018-0313-8.

Clark, M. P., Y. Fan, D. M. Lawrence, J. C. Adam, D. Bolster, D. J. Gochis, R. P. Hooper, M. Kumar, L. R. Leung, D. S. Mackay, R. M. Maxwell, C. Shen, S. C. Swenson, and X. Zeng (2015), Improving the representation of hydrologic processes in Earth System Models, *Water Resources Research*, doi:10.1002/2015WR017096.

Dang, Q., and M. Konar (2018), Trade openness and domestic water use, *Water Resources Research*, *54*(1), 4–18.

Dang, Q., M. Konar, and P. Debaere (2018), Trade openness and the nutrient use of nations, *Environmental Research Letters*, doi:10.1088/1748-9326/aaebcb.

de Groot, J. K., and P. van Groen (2001), The Gorai: Re-excavation project, *Terra et Aqua*.

Deryugina, T., and M. Konar (2017), Impacts of crop insurance on water withdrawals for irrigation, *Advances in Water Resources*, doi:10.1016/j.advwatres.2017.03.013.

Ehret, U., H. V. Gupta, M. Sivapalan, S. V. Weijs, S. J. Schymanski, G. Blöschl, A. N. Gelfan, C. Harman, A. Kleidon, T. A. Bogaard, D. Wang, T. Wagener, U. Scherer, E. Zehe, M. F. P. Bierkens, G. Di Baldassarre, J. Parajka, L. P. H. Van Beek, A. Van Griensven, M. C. Westhoff, and H. C. Winsemius (2014), Advancing catchment hydrology to deal with predictions under change, *Hydrology and Earth System Sciences*, *18*(2), 649–671, doi:10.5194/hess-18-649-2014.

Ferraro, P. J., J. N. Sanchirico, and M. D. Smith (2019), Causal inference in coupled human and natural systems, *Proceedings of the National Academy of Sciences*, *116*(12), 5311–5318, doi:10.1073/pnas.1805563115.

Ganguly, D., and S. Roy (2018), Water-supply alarm in rising salinity of Hooghly.

Gleeson, T., L. W. Erlandsson, S. C. Zipper, M. Porkka, F. Jaramillo, D. Gerten, I. Fetzer, S. E. Cornell, L. Piemontese, L. Gordon, T. Oki, M. Sivapalan, Y. Wada, K. A. Brauman, M. F. P. Bierkens, B. Lehner, P. Keys, M. Kummu, S. Dadson, T. J. Troy, W. Steffen, M. Falkenmark, J. S. Famiglietti, O. Sciences, S. R. Centre, G. Science, U. States, S. P. Road, M. Building, and C. Gifford (2019), The water planetary boundary: a roadmap to illuminate water cycle modifications in the Anthropocene, *Water Resources Research*.

GOB, and GoI (1996), Treaty between the Government of the People's Republic of Bangladesh and the Government of the Republic of India on sharing of the Ganges water at Farakka, *Tech. rep.*, The Government of the People's Republic of Bangladesh and The Government of the Republic of India, Dhaka.

Gupta, H. V., C. Perrin, G. Blöschl, A. Montanari, R. Kumar, M. Clark, and V. Andréassian (2014), Large-sample hydrology: A need to balance depth with breadth, *Hydrology and Earth System Sciences*, *18*(2), 463–477, doi:10.5194/hess-18-463-2014.

Hossain, I. (1998), Bangladesh-India relations: The ganges water-sharing treaty and beyond, *Asian Affairs: An American Review*, *25*(3), 131–150, doi:10.1080/00927679809601449.

Hrachowitz, M., H. H. G. Savenije, G. Blöschl, J. J. McDonnell, M. Sivapalan, J. W. Pomeroy, B. Arheimer, T. Blume, M. P. Clark, U. Ehret, F. Fenicia, J. E. Freer, A. Gelfan, H. V. Gupta, D. A. Hughes, R. W. Hut, A. Montanari, S. Pande, D. Tetzlaf, P. A. Troch, S. Uhlenbrook, T. Wagener, H. C. Winsemius, R. A. Woods, E. Zehe, C. Cudennec, and Others (2013), A decade of Predictions in Ungauged Basins (PUB)—a review, *Hydrological sciences journal*, *58*(6), 1198–1255, doi:10.1080/02626667.2013.803183.

Keiser, D. A., and J. S. Shapiro (2019), Consequences of the clean water act and the demand for water quality, *Quarterly Journal of Economics*, doi:10.1093/qje/qjy019.

Levy, M., A. Lopes, A. Cohn, L. Larsen, and S. Thompson (2018), Land use change increases streamflow across the arc of deforestation in brazil, *Geophysical Research Letters*, *45*(8), 3520–3530.

Lewis, D. (1973), *Counterfactuals*, Harvard University Press, Cambridge, MA.

Lewis, D. (2004), Causation as influence, *The Journal of Philosophy*, *97*(4), 182–197.

Mirza, M. (2004), *The Ganges water diversion: Environmental effects and Implications*, 373 pp., Kluwer Academic Publishers, Dordrecht / Boston / London.

Mondal, S. M., M. T. Islam, D. Saha, M. S. S. Hossain, P. K. Das, and R. Rahman (2019), Agricultural Adaptation Practices to Climate Change Impacts in Coastal Bangladesh, in *Confronting Climate Change in Bangladesh*, pp. 7–21, Springer, Cham, doi:10.1007/978-3-030-05237-9_2.

Montanari, A., G. Young, H. Savenije, D. Hughes, T. Wagener, L. Ren, D. Koutsoyiannis, C. Cudennec, E. Toth, S. Grimaldi, G. Blöschl, M. Sivapalan, K. Beven, H. Gupta, M. Hipsey, B. Schaefli, B. Arheimer, E. Boegh, S. Schymanski, G. Di Baldassarre, B. Yu, P. Hubert, Y. Huang, A. Schumann, D. Post, V. Srinivasan, C. Harman, S. Thompson, M. Rogger, A. Viglione, H. McMillan, G. Characklis, Z. Pang, and V. Belyaev (2013), "Panta Rhei—Everything Flows": Change in hydrology and society—The IAHS Scientific Decade 2013–2022, *Hydrological Sciences Journal*, *58*(6), 1256–1275, doi: 10.1080/02626667.2013.809088.

Mukherjee, A., S. N. Bhanja, and Y. Wada (2018), Groundwater depletion causing reduction of baseflow triggering Ganges river summer drying, *Scientific Reports*, *8*(1), 12,049, doi: 10.1038/s41598-018-30246-7.

Müller, M., and S. Thompson (2019), A value-based model selection approach for environmental random variables, *Water Resources Research*, *55*(1), 270–283.

Müller, M. F., and M. C. Levy (2019), Complementary vantage points: integrating hydrology and economics for sociohydrologic knowledge generation, *Water Resources Research*, doi: 10.1029/2019WR024786.

Müller, M. F., J. Yoon, S. M. Gorelick, N. Avisse, and A. Tilmant (2016), Impact of the Syrian refugee crisis on land use and transboundary freshwater resources, *Proceedings of the National Academy of Sciences*, *113*(52), 14,932–14,937, doi:10.1073/pnas.1614342113.

Müller, M. F. M. F., M. C. M. C. Müller-Itten, and S. M. Gorelick (2017), How Jordan and Saudi Arabia are avoiding a tragedy of the commons over shared groundwater, *Water Resources Research*, *53*(7), 5451–5468, doi:10.1002/2016WR020261.

Pande, S., and M. Sivapalan (2016), Progress in socio-hydrology: a meta-analysis of challenges and opportunities, *Wiley Interdisciplinary Reviews: Water*, *4*(4), e1193, doi: 10.1002/wat2.1193.

Penny, G., and J. Goddard (2018), Resilience principles in socio-hydrology: A case-study review, *Water Security*, *4-5*, doi:10.1016/j.wasec.2018.11.003.

Penny, G., V. Srinivasan, I. Dronova, S. Lele, and S. Thompson (2018), Spatial characterization of long-term hydrological change in the Arkavathy watershed adjacent to Bangalore, India, *Hydrology and Earth System Sciences*, *22*, 595–610, doi:10.5194/hess-22-595-2018.

Rahman, M. M., G. Penny, M. S. Mondal, M. H. Zaman, A. Kryston, M. Salehin, Q. Nahar, M. S. Islam, D. Bolster, J. L. Tank, and M. F. Müller (2019), Salinization in large river deltas: Drivers, impacts and socio-hydrological feedbacks, *Water Security*, *6*(February), 100,024, doi:10.1016/j.wasec.2019.100024.

Ray, J. K. (1978), The Farakka Agreement, *International Studies*, *17*(2), 235–246, doi: 10.1177/002088177801700201.

Salehin, M., M. M. A. Chowdhury, D. Clarke, S. N. Shahjahan Mondal, M. Jahiruddin, and A. Haque (2018), Mechanisms and drivers of soil salinity in coastal Bangladesh, *Ecosystem Services for Well-Being in Deltas*, p. 333.

Savenije, H. H. G. (2009), HESS Opinions "The art of hydrology", *Hydrology and Earth System Sciences*, *13*(2), 157–161, doi:10.5194/hess-13-157-2009.

Sigman, H. (2005), Transboundary spillovers and decentralization of environmental policies, *Journal of Environmental Economics and Management*, doi:10.1016/j.jeem.2004.10.001.

Sivapalan, M., and G. Blöschl (2015), Time scale interactions and the coevolution of humans and water, *Water Resources Research*, *51*(9), 6988–7022.

Sivapalan, M., H. H. G. Savenije, and G. Blöschl (2012), Socio-hydrology: A new science of people and water, doi:10.1002/hyp.8426.

Sivapalan, M., M. Konar, V. Srinivasan, A. Chhatre, A. Wutich, C. A. Scott, J. L. Wescoat, and I. Rodriguez-Iturbe (2014), Socio-hydrology: Use-inspired water sustainability science for the Anthropocene, *Earth's Future*, *2*(4), 225–230, doi:10.1002/2013EF000164.

Srinivasan, V., S. Thompson, K. Madhyastha, G. Penny, K. Jeremiah, and S. Lele (2015), Why is the Arkavathy River drying? A multiple-hypothesis approach in a data-scarce region, *Hydrology and Earth System Sciences*, *19*(4), 1905–1917, doi:10.5194/hessd-12-25-2015.

Srinivasan, V., M. Sanderson, M. Garcia, M. Konar, G. Blöschl, and M. Sivapalan (2016), Prediction in a socio-hydrological world, *Hydrological Sciences Journal*, *62*(3), 338–345, doi:10.1080/02626667.2016.1253844.

Stedinger, J. R., and M. R. Taylor (1982), Synthetic streamflow generation: 1. model verification and validation, *Water resources research*, *18*(4), 909–918.

Stewart-Oaten, A., W. W. Murdoch, and K. R. Parker (1986), Environmental impact assessment:" pseudoreplication" in time?, *Ecology*, *67*(4), 929–940.

Talukder, M. R. R., S. Rutherford, D. Phung, M. Z. Islam, and C. Chu (2016), The effect of drinking water salinity on blood pressure in young adults of coastal Bangladesh, *Environmental pollution*, *214*, 248–254.

Thompson, S. E., M. Sivapalan, C. J. Harman, V. Srinivasan, M. R. Hipsey, P. Reed, A. Montanari, and G. Blöschl (2013), Developing predictive insight into changing water systems: use-inspired hydrologic science for the Anthropocene, *Hydrology and Earth System Sciences*, *17*(12), 5013–5039, doi:10.5194/hess-17-5013-2013.

Underwood, A. (1994), On beyond baci: sampling designs that might reliably detect environmental disturbances, *Ecological applications*, *4*(1), 3–15.

USACE (2016), HEC-RAS River Analysis System, User's Manual, Version 5.0, *Tech. rep.*, U.S. Army Corps of Engineers, Davis.

Vörösmarty, C. J., P. B. McIntyre, M. O. Gessner, D. Dudgeon, A. Prusevich, P. Green, S. Glidden, S. E. Bunn, C. A. Sullivan, C. R. Liermann, Others, and P. M. Davies (2010), Global threats to human water security and river biodiversity., *Nature*, *467*(7315), 555–561, doi:10.1038/nature09549.

Wagener, T., M. Sivapalan, P. a. Troch, B. L. McGlynn, C. J. Harman, H. V. Gupta, P. Kumar, P. S. C. Rao, N. B. Basu, and J. S. Wilson (2010), The future of hydrology: An evolving science for a changing world, *Water Resources Research*, *46*(5), W05,301, doi: 10.1029/2009WR008906.

Williams, R. L. (2000), A note on robust variance estimation for cluster-correlated data, *Biometrics*, *56*(2), 645–646.

Wine, M. L. (2018), Under non-stationarity securitization contributes to uncertainty and Tragedy of the Commons, *Journal of Hydrology*, *568*(July 2018), 716–721, doi: 10.1016/j.jhydrol.2018.11.044.