The Association between Rainfall, Temperature and Reported Drinking Water Source: A Multi-Country Analysis

Andrea Buchwald[1], Evan Thomas[2], Kristopher B Karnauskas[2], Elise Grover[1], Karen L Kotloff[3], and Elizabeth J Carlton[1]

[1]Colorado School of Public Health
[2]University of Colorado Boulder
[3]Department of Pediatrics, University of Maryland School of Medicine

November 26, 2022

Abstract

Climate change may alter access to safe drinking water, with important implications for health. We assessed the relationship between temperature and rainfall and utilization of basic drinking water (BDW) in The Gambia, Mozambique, Pakistan, and Kenya. The outcomes of interest were 1) whether the reported drinking water source used in the past two weeks met the World Health Organization definition of BDW and 2) use of a BDW source that was always available. Temperature and precipitation data were compiled from weather stations and satellite data and summarized to account for long- and short-term weather patterns and lags. We utilized random forests and logistic regression to identify key weather variables that predicted outcomes by site and the association between important weather variables and BDW use. Higher temperatures were associated with decreased BDW use at three of four sites and decreased use of BDW that is always available at all four sites. Rainfall, both in the long- and short-term, was associated with increased BDW use in three sites. We found evidence for interactions between household wealth and weather variables at two sites, suggesting lower wealth populations may be more sensitive to weather-driven changes in water access. Changes in temperature and precipitation can alter safe water use in low-resource settings – investigating drivers for these relationships can inform efforts to build climate resilience.

Table 1. Description of drinking water use and general characteristics of the GEMS study sites. Countries with sufficient variability (>10% and <90% of observations with Outcome 1 or Outcome 2) in the primary and secondary outcome to be included in analysis, are indicated in bold.

	Gambia	Mali	**Mozambique**	**K**
Number of Participants	**2,598**	4,097	**1,976**	**3**
Study Site Characteristics				
Rural/Urban	**Rural**	Urban	**Rural**	**R**
Population at risk	**29,076**	31,768	**15,380**	**2**
Area (km^2)	**1,084**	16	**500**	**5**
Outcomes				
Main source of water is an improved water source*	**85.0**	99.9	**82.6**	**6**
More than 30 minutes wait time for main source of water	**8.6**	2.6	**15.9**	**1**
Main source of water is always available	**54.2**	94.3	**59.7**	**9**
Main source of water is basic drinking water** (Outcome 1)	**77.4**	97.2	**68.9**	**5**
Main source of water is basic drinking water that is always available (Outcome 2)	**35.0**	92.3	**41.5**	**4**
Included in analysis	**Yes**	No	**Yes**	**Y**

***Improved water sources include: Piped water, boreholes or tubewells, protected dug wells, protected springs, rainwater, and packaged or delivered water.**

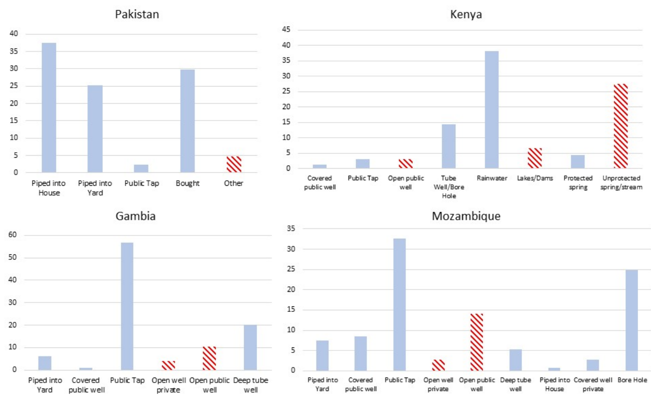
****Basic drinking water is defined as drinking water from an improved source, where collection time is not more than 30 minutes.**

*****Sites were included in our analysis if they had sufficient variability in the outcomes of interest (>10% and<90% of observations with Outcome 1 or Outcome 2).**

Table 2. Variables included in Random forests models

Variables	Variable name in RF plot	Variable form
Rainfall variables		
Mean two-week precipitation	biweekly_p	Continuous
Mean four-week precipitation	fourweekly_p	Continuous
Mean eight-week precipitation	eightweekly_p	Continuous
Mean two-week precipitation, lagged one week	biweekp_lag1	Continuous
Mean two-week precipitation, lagged two weeks	biweekp_lag2	Continuous
Mean four-week precipitation, lagged one week	fourweekp_lag1	Continuous
Mean four-week precipitation, lagged two weeks	fourweekp_lag2	Continuous
Days since previous rainfall	preraindays	Continuous
Maximum one-day rainfall in previous two weeks	max_2	Continuous
Maximum one-day rainfall in previous four weeks	max_4	Continuous
Number of high precipitation days (over 95^{th} percentile) in previous two weeks	sum_high_p	Continuous
Temperature Variables		
Mean two-week temperature	biweekly_t	Continuous
Mean four-week temperature	fourweekly_t	Continuous
Mean two-week temperature, lagged one week	biweekt_lag1	Continuous
Mean two-week temperature, lagged two weeks	biweekt_lag2	Continuous
Mean four-week temperature, lagged one week	fourweekt_lag1	Continuous
Mean four-week temperature, lagged two weeks	fourweekt_lag2	Continuous
Number of high temperature days (over 95^{th} percentile) in previous two weeks	sum_high	Continuous
Number of low temperature days (below 5^{th} percentile) in previous two weeks	sum_low	Continuous
Other Variables		
Case/Control status	Type	Dichotomous
Maternal education level	educat	Categorical
Socio-economic index	wealth	Continuous
Month and year of observation	monthyear	Continuous

Figure 1. Main sources of water use by site. Sources in blue indicate those categorized as an "improved water source" by the WHO.



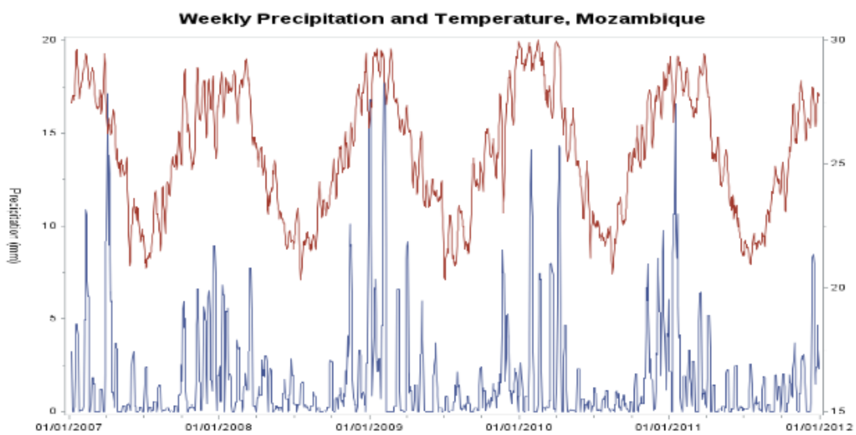


Figure 2. Weekly temperature (red) and precipitation (blue) over the study period, by site

Table 3. Top ten most important predictor variables of basic drinking water use (outcome 1) and using basic drinking water that is always available (outcome 2) identified using random forests models and model parameters by site and outcome, dark red = 1st most important, light yellow = 10th most important.

Variable	The Gambia	The Gambia	Mozambique	Mozan
Outcome Modeled	1	2	1	2
Demographic Variables	**Demographic Variables**	**Demographic Variables**	**Demographic Variables**	**Demo**
Wealth	1	1	1	1
Maternal education	2	7	2	
Case/Control				2
Temperature Variables	**Temperature Variables**	**Temperature Variables**	**Temperature Variables**	**Temp**
Two-week temp	5	2	8	8
Two-week temp, lag1	3	3	9	3
Two-week temp, lag2	4	4		4
Four-week temp		9		
Four-week temp, lag1	6	10	10	

Variable	The Gambia	The Gambia	Mozambique	Mozan
Four-week temp, lag2	7	5	7	5
Precipitation Variables	**Precipitation Variables**	**Precipitation Variables**	**Precipitation Variables**	**Precip**
Days since rainfall	8	8	4	
Two-week precip				
Two-week precip, lag1	9	6	3	6
Two-week precip, lag2	10		5	
Four-week precipitation				10
Four-week precip, lag1			6	7
Four-week precip, lag2				9
Eight-week precip				
Model Parameters	**Model Parameters**	**Model Parameters**	**Model Parameters**	**Mode**
Number of trees	250	250	500	250
Number of variables tried	15	15	14	11
OOB error (%)	4.55	11.86	8.68	10.21
Validation error (%)	5.00	13.50	9.00	9.95

Table 4. Magnitude and direction of associations between most important variables from random forest analysis and basic drinking water use (Outcome 1), adjusted for wealth, by site. Associations are odds ratios and 95% confidence intervals comparing the highest quartile/category to the lowest quartile/category of each variable. When, in tests for linearity, no difference was seen between adjacent categories, quartiles were collapsed and we provide ORs comparing the highest to lowest category (detailed descriptions of variable specification are provided in supplemental tables 2 - 5). Colors indicate direction and strength of association: red = decreased basic drinking water use; blue = increased basic drinking water use. Grey indicates association untested because the variable was not identified as an important predictor in random forests. White cells indicate the association was tested but was not statistically significant.

Variable	The Gambia	Mozambique	Kenya
Demographic Variables	**Demographic Variables**	**Demographic Variables**	**Demographic Varia**
Increasing wealth	3.38 (2.43, 4.71)	1.61 (1.28, 2.04)	1.19 (1.01, 1.39)
Maternal education*		1.67 (1.26, 2.21)	2.90 (2.12, 3.97)
Case (vs. control)			0.74 (0.64, 0.85)
Temperature Variables	**Temperature Variables**	**Temperature Variables**	**Temperature Varia**
Two-week temperature		0.77 (0.62, 0.96)	0.49 (0.40, 0.59)
Two-week temperature with 1-week lag			0.65 (0.53, 0.79)
Two-week temperature with 2-week lag	1.51 (1.05, 2.19)		0.70 (0.58, 0.85)
Four-week temperature			
Four-week temperature with 1-week lag			0.66 (0.57, 0.76)
Four-week temperature with 2-week lag			0.71 (0.61, 0.81)
Precipitation Variables	**Precipitation Variables**	**Precipitation Variables**	**Precipitation Varia**
Previous days since rain			
Two-week precipitation			
Two-week precipitation with 1-week lag			
Two-week precipitation with 2-week lag		1.36 (1.10, 1.68)	2.59 (2.12, 3.15)
Four-week precipitation			
Four-week precipitation with 1-week lag		1.28 (1.03, 1.60)	
Four-week precipitation with 2-week lag,			

Variable	The Gambia	Mozambique	Kenya
Eight-week precipitation			3.95 (3.21, 4.86)

*Maternal education was categorized, based on relevant schooling in each region, in three categories in the Gambia, Mozambique and Kenya, and four categories in Pakistan. Estimates compare the highest to lowest maternal education group for each region. Details are provided in Supplemental Tables 2-5.

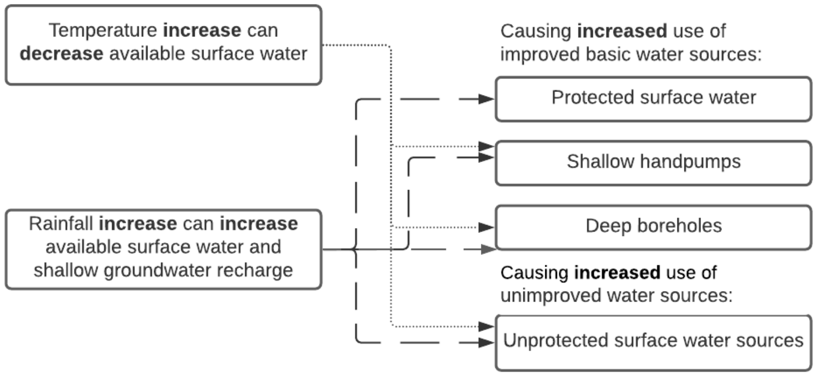
Table 5. Magnitude and direction of associations between most important variables from random forest analysis and using basic drinking water which is always available (Outcome 2), adjusted for wealth, by site. Associations are odds ratios and 95% confidence intervals comparing the highest quartile/category to the lowest quartile/category of each variable. When, in tests for linearity, no difference was seen between adjacent categories, quartiles were collapsed and we provide ORs comparing the highest to lowest category (detailed descriptions of variable specification are provided in supplemental tables 6 - 9). Colors indicate direction and strength of association: red = decreased basic drinking water use; blue = increased basic drinking water use. Grey indicates association untested because the variable was not identified as an important predictor in random forests. White cells indicate the association was tested but was not statistically significant.

Variable	Gambia	Mozambique	Kenya
Demographic Variables	**Demographic Variables**	**Demographic Variables**	**Demographic Varia**
Increasing wealth	0.46 (0.35, 0.62)	2.24 (1.73, 2.89)	1.24 (1.02, 1.50)
Increasing education levels	1.36 (1.05, 1.77)		2.48 (1.85, 3.32)
Case (vs. control)		0.44 (0.36, 0.54)	0.76 (0.66, 0.87)
Temperature Variables	**Temperature Variables**	**Temperature Variables**	**Temperature Varial**
Two-week temperature	0.51 (0.36, 0.71)	0.77 (0.60, 0.99)	0.51 (0.42, 0.62)
Two-week temperature with 1-week lag	0.73 (0.52, 1.03)	0.78 (0.60, 1.00)	0.67 (0.59, 0.78)
Two-week temperature with 2-week lag	0.66 (0.47, 0.92)	0.81 (0.68, 0.97)	0.73 (0.60, 0.88)
Four-week temperature	0.76 (0.62, 0.93)		
Four-week temperature with 1-week lag	0.79 (0.65, 0.96)		
Four-week temperature with 2-week lag			0.68 (0.60, 0.79)
Precipitation Variables	**Precipitation Variables**	**Precipitation Variables**	**Precipitation Varial**
Previous weeks since rain			
Two-week precipitation			2.77 (2.27, 3.37)
Two-week precipitation with 1-week lag	0.75 (0.64, 0.88)		
Two-week precipitation with 2-week lag			2.06 (1.70, 2.51)
Four-week precipitation		0.74 (0.60, 0.91)	
Four-week precipitation with 1-week lag			
Four-week precipitation with 2-week lag		0.79 (0.64, 0.97)	2.29 (1.88, 2.79)
Eight-week precipitation			

* Grey indicates association untested because the variable was not identified as an important predictor. White cells indicate the association was tested but was not statistically significant.

** Previous weeks since rain recalculated from previous days since rainfall used in RF model.

Figure 3. Potential pathways by which temperature or rainfall could impact availability and use of water sources



Hosted file

`gems wsh weather supplemental material.docx` available at https://authorea.com/users/555135/articles/605897-the-association-between-rainfall-temperature-and-reported-drinking-water-source-a-multi-country-analysis

The Association between Rainfall, Temperature and Reported Drinking Water Source: A Multi-Country Analysis

Andrea G. Buchwald[1*], Evan Thomas[2], Kristopher B. Karnauskas[3,1], Elise Grover[1], Karen Kotloff[4], Elizabeth J. Carlton[1]

[1] Colorado School of Public Health, Department of Environmental and Occupational Health, Aurora, Colorado, USA

[2] University of Colorado Boulder, Mortenson Center in Global Engineering, Boulder, Colorado, USA

[3] University of Colorado Boulder, Department of Atmospheric and Oceanic Sciences and Cooperative Institute for Research in Environmental Sciences

[4] University of Maryland School of Medicine, Center for Vaccine Development, Baltimore, Maryland, USA

Corresponding Author Andrea G. Buchwald (andreabuchwald@umaryland.edu)

[*]University of Maryland School of Medicine, 655 W. Baltimore Street, Baltimore, MD, 21201

Key Points:

Access to and reported use of basic drinking water is dependent on rainfall and temperature in The Gambia, Mozambique, Pakistan, and Kenya

Higher temperatures are associated with decreased access to and use of basic drinking water

Climate change threatens access to safe drinking water in settings where infrastructure is vulnerable to rainfall and temperature

Abstract

Climate change may alter access to safe drinking water, with important implications for health. We assessed the relationship between temperature and rainfall and utilization of basic drinking water (BDW) in The Gambia, Mozambique, Pakistan, and Kenya. The outcomes of interest were 1) whether the reported drinking water source used in the past two weeks met the World Health Organization definition of BDW and 2) use of a BDW source that was always available. Temperature and precipitation data were compiled from weather stations and satellite data and summarized to account for long- and short-term weather patterns and lags. We utilized random forests and logistic regression to identify key weather variables that predicted outcomes by site and the association between important weather variables and BDW use. Higher temperatures were associated with decreased BDW use at three of four sites and decreased use of BDW that is always available at all four sites. Rainfall, both in the long- and short-term, was associated with increased BDW use in three sites. We found evidence for interactions between household wealth and weather variables at two sites, suggesting lower wealth populations may be more sensitive to weather-driven changes in water access. Changes in temperature and precipitation can alter safe water use in low-resource settings – investigating drivers for these relationships can inform efforts to build climate resilience.

Keywords:

WASH, climate resilience, water infrastructure, safe water, weather, drinking water

1 INTRODUCTION

Reducing the climate impacts on diarrheal diseases is important, as the burden of diarrheal diseases is high: in 2019, 1.5 million people died from diarrheal disease, with the greatest burden of deaths occurring among children under five years of age (1). Investments in water, sanitation and hygiene (WASH) have been promoted as a way to build resilience to climate variability and change, based on the idea that provision of reliable and safe drinking water sources will reduce vulnerability to enteric diseases in a future with more extremes of rainfall, high temperature and drought (2). This is grounded in two evidence streams. First, there is strong evidence that high temperature, rainfall, and drought increase the risk of diarrheal diseases (3-5). Growing evidence suggests that rainfall, in particular, may impact diarrheal illness via exposures to pathogens in drinking water (6, 7). Second, it is well established that access to safe drinking water can reduce diarrheal diseases: a recent analysis identified unsafe drinking water as the leading environmental risk factor for diarrheal diseases, with approximately 75% of diarrhea-related deaths attributed to use of unsafe drinking water (8). WASH interventions, including providing improved drinking water systems, are associated with significant improvements in early childhood health, including decreases in diarrheal diseases (9).

However, the ways in which temperature and rainfall impact the use and availability of safe drinking water are poorly characterized. We hypothesize that meteorological conditions, such as periods of low rainfall, can impact the availability and use of different drinking water sources. Understanding this relationship is important because while there is considerable evidence that rainfall can compromise water quality through fecal contamination (10, 11), less is known about how different weather conditions alter the use of different types of drinking water sources. If people are using more or less safe water sources under different weather conditions, this can alter the impacts of WASH investments on health and climate vulnerability.

Prior work has found evidence of seasonal patterns in drinking water use, but the results are inconsistent. Qualitative research into WASH uptake has frequently identified seasonal factors including temperature, rain, flooding, water scarcity, and seasonal field-work as influencing WASH uptake, desirability and feasibility (12-20). A number of these studies have found evidence that seasonality directly influences water-source choice. For example in India, treated water is preferred in the rainy season due to perception of decreased water quality following rain (11).

There is also evidence that seasonality influences water availability. In some sites in Ghana, Kenya, and Zambia, less safe water sources were used in the rainy season due to failure of solar-powered pumps (21). Quantitative studies are limited. Several studies indicate preference for surface water sources during the rainy season or after heavy rainfalls, even when groundwater sources were available (23-25). Surface water sources are often more convenient and available free of cost but are vulnerable to fecal contamination (25). Rainy season was associated with increased rainwater use in the Pacific (26) and increased surface water usage in East Africa (27, 28). Drought, which is happening with greater frequency and severity, can lead to limited water availability (30, 32), and increased contamination (33). Additionally, season is known to impact the ability of communities to maintain water sources and latrines, with stressors in both rainy and dry seasons (21, 34, 35). Despite substantial qualitative evidence supporting seasonal changes in water source selection, there is limited quantitative research on how changing meteorological conditions affect water source use and access.

In this study, we aim to evaluate how meteorological conditions including high temperature and drought impact the use and availability of drinking water sources across four diverse locations in Asia and Africa. Because research on this topic has been limited and the evidence to date is inconsistent, we adopted an analytical framework that allowed us to consider a large set of candidate predictors, describing long- and short-term rainfall and precipitation patterns. This approach is designed to be hypothesis generating, facilitating identification of key predictors for investigation in future studies, while avoiding the perils of multiple hypothesis testing. We used a standard World Health Organization definition of basic drinking water and water that is always available to ensure the generalizability of our findings to global safe water standards. Given the well-recognized role of socio-economic status in access to safe drinking water we included this as a predictor and tested for evidence that socio-economic status modifies the relationship between climate variables and basic water use.

2 METHODS

This analysis utilizes data from the Global Enteric Multicenter Study (GEMS) of moderate-to-severe diarrheal disease (MSD) in infants and young children in developing countries (36) as well as in situ and modeled meteorological data to assess the relationship between weather and the utilization of and access to improved water sources.

2.1 Study Population:

Household drinking water use behaviors were drawn from GEMS. GEMS was conducted in seven countries (Kenya, Mali, Mozambique, The Gambia, Bangladesh, India, and Pakistan) with moderate-to-high under-five child mortality to study enteric disease epidemiology and has been described at length (36).

In brief, GEMS was a 3 year (Dec 2007 to March 2011), prospective, age-stratified, matched case-control study of MSD among children 0–59 months of age belonging to a geographically-defined censused population that varied in size from 10 km^2 to 1,084 km^2 (Table 1). Cases were systematically enrolled from those meeting the case definition of MSD and seeking care at hospitals and health centers. For each case, one to three controls were randomly selected from a demographic surveillance system to serve as controls. Controls were enrolled within 14 days of the index case and matched to cases by age, gender, and location. Upon enrollment, parents or primary caretakers of cases and controls were administered a detailed survey to assess demographics, household wealth indicators, and water usage. At a follow-up visit 50-90 days after enrollment, water usage questions were asked again but water collection time was collected only at enrollment.

Because cases and controls were enrolled year-round over a 36-month period and asked about water sources, availability and fetching times over the past two weeks, this presents a unique opportunity to assess temporal variation in drinking water source use. This analysis includes data from all participating households. While GEMS participants may not represent a true random sample of the population, both cases and controls were evenly sampled throughout the year and selected based on the date of case-illness and thus any selection bias related to weather variables is assumed to be uniform between cases and controls.

2.2 Basic water use:

The primary outcome of interest was whether a household's reported main source of drinking water used in the past two weeks meets the WHO definition of "Basic Drinking Water" (BDW) (37). BDW is defined as drinking water from an improved source, provided collection time is not more than 30 minutes, with improved sources including piped water, boreholes, tubewells, protected dug wells, protected springs, rainwater, and packaged water.

Water source type was assessed with the question "During the last two weeks, what was the main source of drinking water for the members of your household?" at enrollment. Only one answer was allowed. Water collection time was collected with the question, "How long does it take to go there [main source of drinking water], get water, and come back?".

As a secondary outcome, we examined the availability of BDW. Water availability was determined from the question, "In the last two weeks, how often has this water been available from this main source?" For this outcome, a household was classified as using BDW that is always available if their main water source met the above criteria for BDW and they reported the source was always available. Notably, this definition does not include any measure of drinking water quality. BDW sources have been known to be contaminated at the point of collection and/or point of use with fecal bacteria: a recent meta-analysis indicated that 10% of improved sources may contain over 100 *E. coli* or TTC per 100 ml, well above safe drinking water standards (38). Therefore, this data set cannot identify water that is free of unsafe contamination.

We first examined the distribution of each outcome at each site. We restricted our analysis to sites with sufficient variability in both of the outcomes of interest (defined as having between 10% and 90% using BDW and between 10% and 90% using BDW that is always available). Only four sites (The Gambia, Mozambique, Kenya, and Pakistan) met this criterion and were included in this paper (Table 1).

Table 1. Description of drinking water use and general characteristics of the GEMS study sites. Countries with sufficient variability ($\geq$10% and $\leq$90% of observations with Outcome 1 or Outcome 2) in the primary and secondary outcome to be included in analysis, are indicated in bold.

	Gambia	Mali	**Mozambique**	**Kenya**	India	Bangladesh	**Pakistan**
Number of Participants	**2,598**	4,097	**1,976**	**3,359**	3,582	3,859	**3,096**
Study Site Characteristics							
Rural/Urban	**Rural**	Urban	**Rural**	**Rural**	Urban	Rural	**Urban**
Population at risk	**29,076**	31,768	**15,380**	**21,603**	13,416	25,560	**25,659**
Area (km^2)	**1,084**	16	**500**	**500**	10.5	374	**10**
Outcomes							
Main source of water is an improved water source*	**85.0**	99.9	**82.6**	**62.6**	98.6	99.8	**95.2**
More than 30 minutes wait time	**8.6**	2.6	**15.9**	**19.7**	8.4	0.1	**19.3**

for main source of water							
Main source of water is always available	**54.2**	94.3	**59.7**	**90.3**	1.0	99.9	**62.4**
Main source of water is basic drinking water** (Outcome 1)	**77.4**	97.2	**68.9**	**55.0**	90.5	99.6	**76.4**
Main source of water is basic drinking water that is always available (Outcome 2)	**35.0**	92.3	**41.5**	**46.9**	1.0	99.6	**45.7**
Included in analysis	**Yes**	No	**Yes**	**Yes**	No	No	**Yes**

***Improved water sources include: Piped water, boreholes or tubewells, protected dug wells, protected springs, rainwater, and packaged or delivered water.**

****Basic drinking water is defined as drinking water from an improved source, where collection time is not more than 30 minutes.**

*****Sites were included in our analysis if they had sufficient variability in the outcomes of interest (≥10% and ≤90% of observations with Outcome 1 or Outcome 2).**

2.3 Meteorological Data:

Precipitation data come from a gridded product that combines satellite measurements and rain gauges: the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) (39). Daily precipitation (mm) at a resolution of 0.05 degree (~5 km) was acquired for the years 2007 through 2011. A daily precipitation record for each study site was calculated by taking the mean spatial mean across a rectangular area encompassing the northmost, southmost, eastmost, and westmost points of the study site.

Temperature data were compiled from weather stations nearest to each study site. NOAA had available weather station data for three of the four study sites (40). For the fourth site, Kenya, the nearest weather station with NOAA data available was ~100 Km from the study, so the Kenya Medical Research Institute daily temperature records were used. To account for missing data in the weather station temperature records, data were infilled with temperature data from the 0.25 degree forcing dataset for version 1 of the Global Land Data Assimilation System (GLDAS) (41). The Mozambique site was missing 34% of observations for temperature from the weather stations. GLDAS data was highly correlated with observed temperatures from NOAA (R^2 = 0.86) and linearly transformed (temp = 1.03*(GLDASvalue C°)+1.45) to fill in missing observed temperatures. Only 4% of the temperature observations from Kenya were missing, and GLDAS data was linearly transformed with the following equation: temp = 0.48*(GLDASvalue mm) + 13.3 to infill the missing data points. Pakistan and Gambia had excellent observational coverage (<1% of days missing), and so were not infilled.

We posited that BDW use may be impacted both by seasonal rainfall and temperature patterns as well as by short-term meteorological events. For example, months-long dry periods may reduce surface water availability and prolonged heat may favor evaporative processes over groundwater recharge. Recent rainfall may favor use of surface water and replenished shallow ground water sources. We also posited that there are likely lags between rainfall, temperature and water use, given the time required for recharge of improved water sources. For this reason, we defined a set of meteorological variables that capture potential long-term and short-term conditions defining temperature and rainfall conditions over two, four and eight-week periods, and also considered lags of zero, one and two weeks (Table 2).

Table 2. Variables included in Random forests models

Variables	Variable name in RF plot	Variable format	Lag
Rainfall variables			
Mean two-week precipitation	biweekly_p	Continuous	0
Mean four-week precipitation	fourweekly_p	Continuous	0
Mean eight-week precipitation	eightweekly_p	Continuous	0
Mean two-week precipitation, lagged one week	biweekp_lag1	Continuous	1 week
Mean two-week precipitation, lagged two weeks	biweekp_lag2	Continuous	2 weeks
Mean four-week precipitation, lagged one week	fourweekp_lag1	Continuous	1 week
Mean four-week precipitation, lagged two weeks	fourweekp_lag2	Continuous	2 weeks
Days since previous rainfall	preraindays	Continuous	0
Maximum one-day rainfall in previous two weeks	max_2	Continuous	0
Maximum one-day rainfall in previous four weeks	max_4	Continuous	0
Number of high precipitation days (over 95th percentile) in previous two weeks	sum_high_p	Continuous	0
Temperature Variables			
Mean two-week temperature	biweekly_t	Continuous	0
Mean four-week temperature	fourweekly_t	Continuous	0
Mean two-week temperature, lagged one week	biweekt_lag1	Continuous	1 week
Mean two-week temperature, lagged two weeks	biweekt_lag2	Continuous	2 weeks
Mean four-week temperature, lagged one week	fourweekt_lag1	Continuous	1 week
Mean four-week temperature, lagged two weeks	fourweekt_lag2	Continuous	2 weeks

Number of high temperature days (over 95th percentile) in previous two weeks	sum_high	Continuous	0
Number of low temperature days (below 5th percentile) in previous two weeks	sum_low	Continuous	0
Other Variables			
Case/Control status	Type	Dichotomous	N/A
Maternal education level	educat	Categorical	N/A
Socio-economic index	wealth	Continuous	N/A
Month and year of observation	monthyear	Continuous	N/A

2.4 Demographic data*:*

Household socioeconomic status (SES) and maternal education were included as potential predictors of BDW use and access as SES has previously been found to be an important predictor of water access (42, 43). An asset-based SES index was calculated for each site using PCA incorporating standard economic indicator variables (44) including household assets, and household population. Distribution of indicators varied substantially between sites, thus some indicators were excluded for some sites due to a lack of variability (either no ownership, or complete saturation of the indicator) at the given site (Supplemental Table 1). For each site we utilized the first principal component which explained the greatest percentage of variance across the population as the wealth index. Maternal education level was collected in the survey as a 7-level categorical variable with categories: No formal schooling, less than primary, completed primary, completed secondary, post-secondary, religious education only, or unknown. Maternal education level was categorized based on the education distribution by site, these categories were not the same between sites due to differences in the distribution of education level between sites.

Date of survey was included in models in order to account for other time-dependent changes in water use not captured by weather variables (i.e. political or infrastructural changes that may take place over time). SES, maternal education, and case-status were all examined as potential predictors.

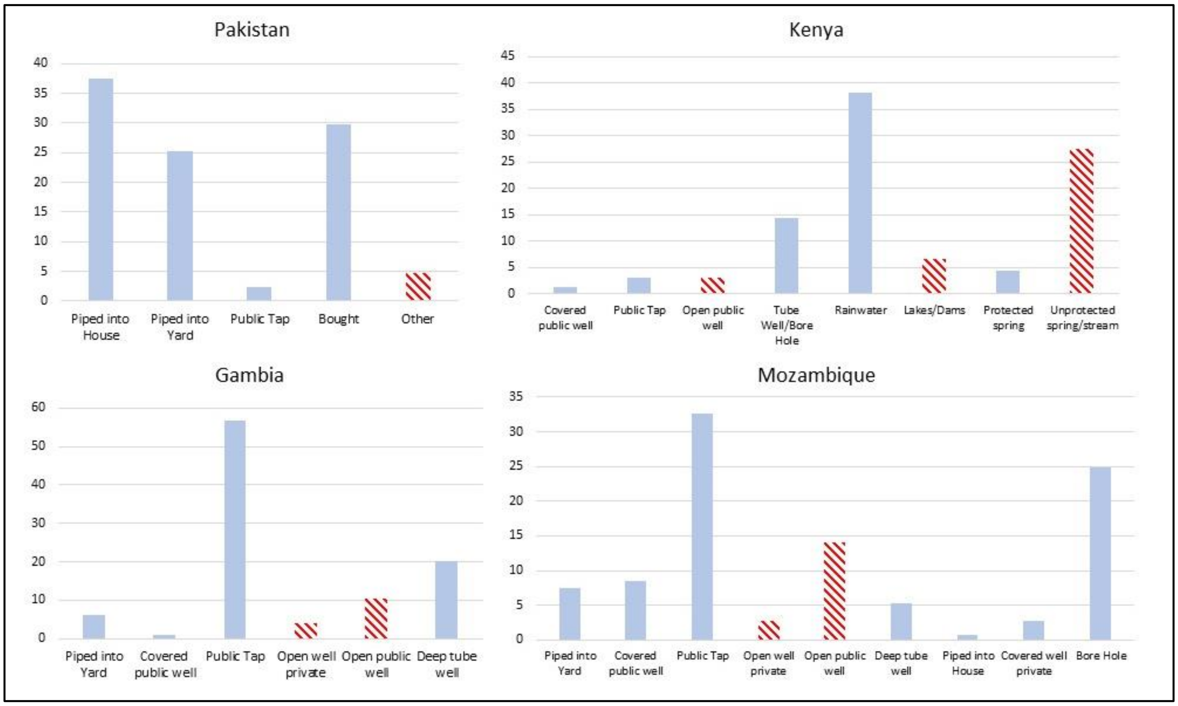
2.5 Analysis:

Given the large number of potential predictor variables, and the limited research to date on this topic, we opted to employ an analytical approach to identify key predictors and assess the magnitude and direction of the association between key predictors and the outcome of interest. This has the advantage of allowing us to consider a wide array of candidate predictors, avoids the perils of multiple hypothesis testing, and is intended to narrow the list of key meteorological conditions that could be pursued with more focused causal models in subsequent studies. To this end, we conducted random forests (RF) machine learning to identify the most important rainfall and temperature variables for predicting the use of 1) BDW or 2) BDW always available by site. RF requires data to be balanced in respect to the outcome (i.e. approximately equal proportions using BDW as not), and as only a fraction of the population at each site reported using BDW, data was weighted and resampled for each site to achieve balanced datasets for RF. RF models included all rainfall and temperature variables, as well as SES, maternal education level, date, and case/control status. Data was split 70%/30% into training and validation sets: models were constructed using the training datasets and tuned by varying the number of trees created and the number of variables randomly sampled at each stage. Final RF models were selected based on out of bag error rate using the validation dataset, and models with the lowest error rate were used to identify most important variables. We then selected the ten variables with highest mean decrease in accuracy values from the final RF models to examine for direction of association with BDW use in logistic regression.

We constructed logistic regression models, incorporating the ten most important variables from RF for each outcome, in order to assess the direction and magnitude of the associations between the key predictors and our outcome of interest. We first generated unadjusted estimates of associations between BDW use and each important independent variable using logistic regression, dividing all continuous independent variables into quartiles, based on the spread of the data. This approach avoids the assumption of linearity and allows us to identify more complex relationships between variables (e.g., thresholds). We modeled the independent variable as continuous when linear relationships were evident. Additionally, when no difference was seen between adjacent categories, we collapsed quartiles into fewer categories. SES was categorized into high (top 25% of population), middle (middle 50% of population), low (lower 25% of population) and modeled as linear when justified. Education level was categorized into three or four groups based on the differences in the types of schooling between sites. Because SES was the most important predictor of BDW use at all sites, we adjusted all estimates for SES. Estimates were generated separately for each of the top ten variables at each site.

We constructed multivariate logistic regression models, in order to identify relationships between exposure variables and BDW use, independent of other important variables. Effect modification was tested by including interaction terms in models. Variables that were statistically significant at $p \le 0.1$ in the SES-adjusted models were tested for inclusion in a final multivariate model. Variables were excluded from the model in order of least significance/effect on other variables, and then retested for inclusion in the final model. If two variables were collinear (variance inflation factor >5), the variable with the greater statistical significance was included, and the other was excluded. As a sensitivity analysis, we used the final model to test for evidence of effect modification by SES, case status, and education level. We repeated this process for Outcome 2, BDW that is always available.

Figure 1. Main sources of water use by site. Sources in blue indicate those categorized as an "improved water source" by the WHO.



3 RESULTS

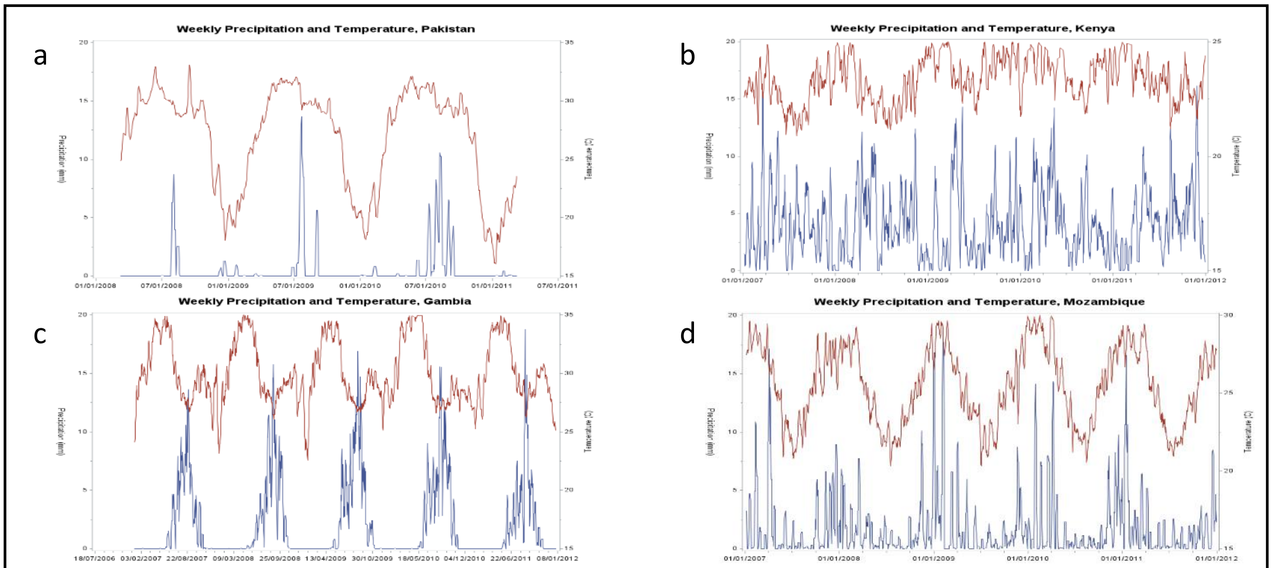
<u>3.1 BDW use:</u>

Participants reported high use of improved drinking water sources at all sites, ranging from 63% in Kenya to 95% in Pakistan (Figure 1, Table 1). Main sources of drinking water varied by site (Figure 1); participants in Pakistan primarily used improved water sources that were piped (water was piped in from Karachi). In The Gambia, over half of households reported using a public tap for drinking water. Kenya and Mozambique had a wide range of reported water sources, including various wells and taps. Kenya was the only site with significant surface water and rainwater use. The percent of households using BDW (Outcome 1) ranged from 55% in Kenya to 77% in Gambia. Having a main source of water that was always available (Outcome 2) was lowest in The Gambia, with only 35% of participants reporting water was always available, and highest in Kenya (47%).

<u>3.2 Distribution of rainfall and temperature variables:</u>

Daily precipitation and temperature over the study period by site are shown in Figure 2. Temperature variability was lowest in Kenya (Figure 2b), with highest temperature variability seen in Pakistan (Figure 2a). Pakistan had very little rainfall compared to the other sites.

Figure 2. Weekly temperature (red) and precipitation (blue) over the study period, by site



Precipitation data obtained from Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS). Temperature data obtained from NOAA weather stations and missing temperature data was infilled from Global Land Data Assimilation System (GLDAS) for Kenya and Mozambique.

3.3 Outcome 1: BDW use:

The best fitting RF model for rainfall and temperature-predictors of BDW use varied widely between sites (Table 3). Models were least predictive of water use outcomes in Kenya, with error rates as high as 19.2% in Kenya. Model fit was best for Outcome 1 in Gambia with 95% of observations in the validation dataset predicted correctly.

Table 3. Top ten most important predictor variables of basic drinking water use (outcome 1) and using basic drinking water that is always available (outcome 2) identified using random forests models and model parameters by site and outcome, dark red = 1st most important, light yellow = 10th most important.

Variable	The Gambia	The Gambia	Mozambique	Mozambique	Kenya	Kenya	Pakistan	Pakistan
Outcome Modeled	1	2	1	2	1	2	1	2
Demographic Variables								
Wealth	1	1	1	1	1	1	1	1
Maternal education	2	7	2		2	2	2	2
Case/Control				2	4	3	9	8
Temperature Variables								
Two-week temp	5	2	8	8	6	7	4	3
Two-week temp, lag1	3	3	9	3	8	9	5	4
Two-week temp, lag2	4	4		4	7	10	6	6
Four-week temp		9					8	7
Four-week temp, lag1	6	10	10		9		10	10

Four-week temp, lag2	7	5	7	5	10	6	7	9
Precipitation Variables								
Days since rainfall	8	8	4				3	5
Two-week precip						5		
Two-week precip, lag1	9	6	3	6				
Two-week precip, lag2	10		5		5	4		
Four-week precipitation				10				
Four-week precip, lag1			6	7				
Four-week precip, lag2				9		8		
Eight-week precip					3			
Model Parameters								
Number of trees	250	250	500	250	500	500	250	250
Number of variables tried	15	15	14	11	14	12	15	17
OOB error (%)	4.55	11.86	8.68	10.21	16.93	19.18	7.29	17.82
Validation error (%)	5.00	13.50	9.00	9.95	18.00	21.10	6.80	18.30

Most important variables from all RF models are summarized in Table 3. Wealth was the top predictor of BDW use for all four sites, followed by maternal education. Both temperature and precipitation variables ranked in the top ten predictors for all four sites. Mean two-week temperature over the previous two weeks with no, one- or two- week lags was important in all RF models, as was mean four-week temperature with a one- and two-week lag. The most frequently selected precipitation measure was the number of days since the last rainfall and mean two-week rainfall lagged by two weeks (both selected in models for three of the four sites). Variables describing maximum temperature or high precipitation events including the number of high precipitation days, high temperature days, low temperature days, and the maximum two-and four-week precipitation were not in the top ten most important variables for any site. The variable for date was also not in the top ten most important variables for any site.

Estimates of the strength and direction of the association between important variables and BDW use, based on logistic regression analysis, are shown in Table 4. Increasing household wealth was associated with increased use of BDW at all sites. Even after adjusting for household wealth, increasing maternal education was associated with increased BDW use in Mozambique, Kenya, and Pakistan.

Table 4. Magnitude and direction of associations between most important variables from random forest analysis and basic drinking water use (Outcome 1), adjusted for wealth, by site. **Associations are odds ratios and 95% confidence intervals comparing the highest quartile/category to the lowest quartile/category of each variable. When, in tests for linearity, no difference was seen between adjacent categories, quartiles were collapsed and we provide ORs comparing the highest to lowest category (detailed descriptions of variable specification are provided in supplemental tables 2 - 5). Colors indicate direction and strength of association: red = decreased basic drinking water use; blue = increased basic drinking water use. Grey indicates association untested because the variable was not identified as an important predictor in random forests. White cells indicate the association was tested but was not statistically significant.**

Variable	The Gambia	Mozambique	Kenya	Pakistan
Demographic Variables				
Increasing wealth	3.38 (2.43, 4.71)	1.61 (1.28, 2.04)	1.19 (1.01, 1.39)	2.62 (2.03, 3.38)
Maternal education*		1.67 (1.26, 2.21)	2.90 (2.12, 3.97)	1.68 (1.27, 2.23)
Case (vs. control)			0.74 (0.64, 0.85)	
Temperature Variables				
Two-week temperature		0.77 (0.62, 0.96)	0.49 (0.40, 0.59)	0.72 (0.59, 0.89)
Two-week temperature with 1-week lag			0.65 (0.53, 0.79)	0.66 (0.53, 0.81)
Two-week temperature with 2-week lag	1.51 (1.05, 2.19)		0.70 (0.58, 0.85)	0.67 (0.54, 0.82)
Four-week temperature				0.68 (0.55, 0.83)
Four-week temperature with 1-week lag			0.66 (0.57, 0.76)	0.61 (0.49, 0.76)
Four-week temperature with 2-week lag			0.71 (0.61, 0.81)	0.64 (0.51, 0.79)
Precipitation Variables				
Previous days since rain				0.71 (0.56, 0.90)
Two-week precipitation				
Two-week precipitation with 1-week lag				
Two-week precipitation with 2-week lag		1.36 (1.10, 1.68)	2.59 (2.12, 3.15)	
Four-week precipitation				
Four-week precipitation with 1-week lag		1.28 (1.03, 1.60)		
Four-week precipitation with 2-week lag,				
Eight-week precipitation			3.95 (3.21, 4.86)	

*Maternal education was categorized, based on relevant schooling in each region, in three categories in the Gambia, Mozambique and Kenya, and four categories in Pakistan. Estimates compare the highest to lowest maternal education group for each region. Details are provided in Supplemental Tables 2-5.

Increasing rainfall, both in the long- and short-term, was associated with increased use of BDW in Mozambique and Kenya and longer dry periods were associated with decreased use of BDW in Pakistan. Increasing temperatures were associated with decreased use of BDW in Mozambique, Kenya, and Pakistan. However, in The Gambia, BDW use increased when mean two-week temperature with a two-week lag was above 26.6 degrees C (the 25th percentile value) and no precipitation measure was associated with BDW use, adjusting for wealth. Estimates generated using linear exposure variables (when appropriate) were generally consistent with these findings (Supplemental Tables 2 – 5).

Adjustment for other statistically significant weather and demographic variables had minimal effect on estimates of association in Mozambique or Pakistan (Supplemental Tables 3 & 5). In Kenya, after adjustment, education was included in the final model predicting BDW use, and SES was not. In Kenya, adjustment did lead to a change in the estimate of association for biweekly temperature with a two-week lag, but this is assumed to be a result of collinearity between that variable and biweekly temperature with no lag (Supplemental Table 4). In Gambia, biweekly temperature with a two-week lag was the only variable with a strong association with BDW use, so an adjusted model was not constructed (Supplemental Table 2). There was no evidence of effect modification by SES, wealth or case status on the relationship between weather variables and BDW use at any of the four sites.

3.4 Outcome 2: Use of BDW that is always available:

As with Outcome 1, wealth was the top predictor of using BDW that is always available, however maternal education was no longer the 2nd most important in Gambia and Mozambique. The same temperature and precipitation variables that were important for Outcome 1 were usually important for predicting Outcome 2, but Kenya and Mozambique both had long-term precipitation variables that were important for Outcome 2 which had not been important at any site for Outcome 1 (Table 3).

Estimates of the strength and direction of the association between weather variables and use of BDW that is always available are shown in Tables 5. Increasing wealth was positively associated with Outcome 2 in three of the four sites; a negative association was seen in Gambia. Increasing education level was associated with increased use in Gambia, Kenya, and Pakistan. Households with moderate to severe diarrhea cases were significantly less likely to use BDW which was always available in three sites: Mozambique, Kenya, and Pakistan. Increasing temperature, on both a long and short scale, was consistently associated with decreased use of BDW that was always available at all study sites. The association between precipitation and Outcome 2 varied by site. Increasing long-term (four- and eight-week) precipitation was associated with increased use of always available BDW in Kenya, and longer dry periods were associated with decreased use of always available BDW in Pakistan, however, in contrast to Outcome 1, increasing precipitation was associated with decreased use of BDW that is always available in Gambia and Mozambique.

Adjustment for other important variables had minimal effect on the association between weather variables and use of BDW that was always available in Kenya, Pakistan, and Gambia (Supplemental Tables 6, 8, 9). In Mozambique, there was evidence for qualitative interaction between case-status and biweekly temperature with a two-week lag (Supplemental Table 7a), such that the decreased use of always available BDW at higher temperatures was only seen among controls. In Pakistan, there was evidence of interaction between SES and case-status and moderate evidence that the association between the number of previous weeks since rain and use of BDW that is always available was most pronounced in the lowest SES group. Among those in the lowest SES category, high severity of drought (>6 weeks since rainfall) is associated with an OR = 0.47 (95% CI: 0.30, 0.74) for use of BDW that is always available, compared to having rainfall in the past week. Gambia similarly had evidence of interaction between education and SES, and the association between mean two-week temperature and Outcome 2 was most pronounced in those without any formal education, OR = 0.40 (95%CI: 0.23, 0.70). Among those with any formal education, the OR = 1.28 (95%CI: 0.41, 3.96).

Table 5. Magnitude and direction of associations between most important variables from random forest analysis and using basic drinking water which is always available (Outcome 2), adjusted for

wealth, by site. Associations are odds ratios and 95% confidence intervals comparing the highest quartile/category to the lowest quartile/category of each variable. When, in tests for linearity, no difference was seen between adjacent categories, quartiles were collapsed and we provide ORs comparing the highest to lowest category (detailed descriptions of variable specification are provided in supplemental tables 6 - 9). Colors indicate direction and strength of association: red = decreased basic drinking water use; blue = increased basic drinking water use. Grey indicates association untested because the variable was not identified as an important predictor in random forests. White cells indicate the association was tested but was not statistically significant.

Variable	Gambia	Mozambique	Kenya	Pakistan
Demographic Variables				
Increasing wealth	0.46 (0.35, 0.62)	2.24 (1.73, 2.89)	1.24 (1.02, 1.50)	1.84 (1.50, 2.26)
Increasing education levels	1.36 (1.05, 1.77)		2.48 (1.85, 3.32)	1.38 (1.13, 1.70)
Case (vs. control)		0.44 (0.36, 0.54)	0.76 (0.66, 0.87)	0.75 (0.65, 0.87)
Temperature Variables				
Two-week temperature	0.51 (0.36, 0.71)	0.77 (0.60, 0.99)	0.51 (0.42, 0.62)	0.87 (0.73, 1.04)
Two-week temperature with 1-week lag	0.73 (0.52, 1.03)	0.78 (0.60, 1.00)	0.67 (0.59, 0.78)	0.73 (0.63, 0.85)
Two-week temperature with 2-week lag	0.66 (0.47, 0.92)	0.81 (0.68, 0.97)	0.73 (0.60, 0.88)	0.67 (0.54, 0.82)
Four-week temperature	0.76 (0.62, 0.93)			0.73 (0.63, 0.85)
Four-week temperature with 1-week lag	0.79 (0.65, 0.96)			0.69 (0.69, 0.80)
Four-week temperature with 2-week lag			0.68 (0.60, 0.79)	0.68 (0.59, 0.79)
Precipitation Variables				
Previous weeks since rain				0.72 (0.59, 0.88)
Two-week precipitation			2.77 (2.27, 3.37)	
Two-week precipitation with 1-week lag	0.75 (0.64, 0.88)			
Two-week precipitation with 2-week lag			2.06 (1.70, 2.51)	
Four-week precipitation		0.74 (0.60, 0.91)		
Four-week precipitation with 1-week lag				
Four-week precipitation with 2-week lag		0.79 (0.64, 0.97)	2.29 (1.88, 2.79)	
Eight-week precipitation				

* Grey indicates association untested because the variable was not identified as an important predictor. White cells indicate the association was tested but was not statistically significant.

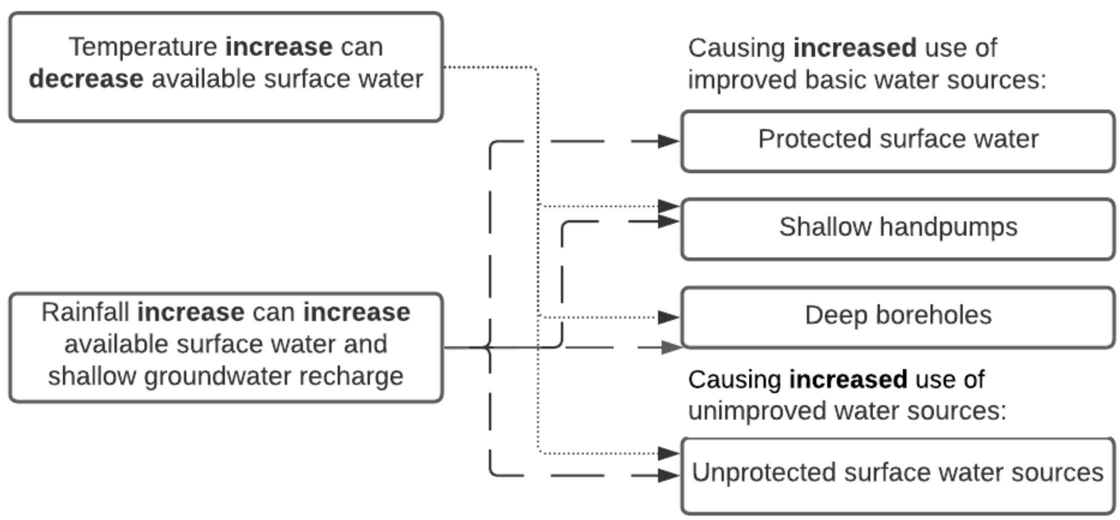
** Previous weeks since rain recalculated from previous days since rainfall used in RF model.

4 DISCUSSION

By combining weather data with a large population-based study of diarrheal disease in four countries, we found temperature and precipitation were significantly associated with the availability and use of BDW, however with different directions of association depending on the context. This study capitalized on a large population-level longitudinal dataset with thousands of observations per country, capturing a wide temporal and spatial range. Patterns in the availability and use of different water sources may be influenced by seasonality and short- and long-term rainfall variability.

In this study, we had four key findings. 1) Across all countries, household socioeconomic status was by far the most important predictor of increased use of BDW, followed closely in importance by education status. Beyond predicting BDW overall, individuals with low SES were more vulnerable to prolonged dry periods (in Pakistan) or high temperatures (in The Gambia). In three of four locations studied, 2) as temperature increases, BDW use, and use of BDW that is always available decreases and 3) increasing rainfall increased BDW use but did not always increase availability of BDW. Lastly, 4) in The Gambia the association between weather and BDW use did not follow the same patterns in most analyses – suggesting some water systems may be less impacted by weather than others. Notably, The Gambia had the highest BDW use (77%) of sites in our study, was the only location where >50% of the population reported using public tap, and had the lowest spatial resolution. As a result, it is unclear if the unique patterns seen in The Gambia are due to imprecision of our weather estimates or increased resilience to extreme weather.

Figure 3. Potential pathways by which temperature or rainfall could impact availability and use of water sources



There are numerous pathways by which climate change may lead to changes in use of and access to BDW sources (Figure 3). In some contexts, increasing temperatures may correlate with decreased surface water retention or shallow groundwater and motivate users toward less safe groundwater sources or less convenient water sources, alternately increasing temperatures may decrease motivation for seeking out safer sources and prompt fallback to more convenient sources of water including groundwater or open wells. Likewise, decreasing rainfall may result in surface water sources drying up, motivating use of other water sources that may or may not be protected. The evidence to date has shown that seasonal and long-term changes in temperature and rainfall can change the mix and convenience of available water sources for communities. A study in Ethiopia identified that water collection times increase during the dry season (27), and a qualitative study of water users and managers in Ghana, Kenya and Zambia reported less time collecting water in the rainy seasons (21). This is consistent with our findings that in three of four sites, BDW use decreased during hot periods and increased during wet period. However, both our study and the work of others suggest the impact of weather on BDW use and access is context specific. A study in South Africa found some households switch from more contaminated surface water to safer municipal water sources during the dry season (45). In several recent studies, researchers have examined patterns in use of groundwater boreholes in arid regions of Kenya and Ethiopia and compared these patterns to rainfall trends in the region. In these studies, an inverse relationship between use of electrical borehole pumps as well as handpumps and recent rainfall was observed, as well as overall seasonal trends in decreased groundwater pump use during rainy seasons (23, 24). These trends appear to reflect behavioral choices to use surface water sources when available, and do not, generally, reflect an intrinsic hydrologic relationship between rainfall and aquifer recharge. Notably, this behavior has been observed as a risk to professional drinking water services as users may be less willing to pay for improved water sources when unimproved surface water sources are seasonally available (46). In this study, rainfall was not predictive of BDW at the site with the greatest baseline access to BDW, underscoring the importance of work to understand how the mix of available water sources impacts climate vulnerability.

4.1 Limitations:

This study has several important limitations. Although we tried to standardize our definition of water sources using the World Health Organization categories of improved water source, the categories are not perfect and do not distinguish between “improved” drinking water sources and water that is free of unsafe contamination. We were unable to measure contamination directly and the water sources were not observed by study staff. Unfortunately, given the data set available, verifying the nature of these water sources was beyond the scope of our analysis. Future studies examining these questions would benefit from testing and observing the water sources. Further, these improved BDW sources include protected surface water sources, and shallow and deep groundwater sources, which have different hydrological and climatic response profiles as well as contamination risk.

There was substantial variation in the size and population density of the study sites we examined by design, including both rural and urban locations, ranging from 10 Km^2 in Pakistan to over 1,000 Km^2 in The Gambia. Weather data – rainfall, in particular, frequently varies over small spatial scales (47). By averaging weather variables over the study sites, we may have introduced misclassification of the weather variables, particularly in the larger sites. Similarly, acute weather events may play a large role in access to and decisions about water-source use. By asking about water-use over the last two weeks, we are unable to capture the day to day shifts that may occur as a result of acute events. Lastly, we were unable to capture a change in water source use within individual households as a result of weather, as water collection time was only measured at baseline. Comparing main water sources, 90% of households that reported using improved water sources at enrollment reported improved water sources as their main water source at follow-up, and 70% of households reporting using unimproved water sources at enrollment continued to use unimproved sources at follow-up. Future analysis of this dataset could be used to examine within-household changes in drinking water source use among households with multiple observations.

4.2 Conclusions

Despite these limitations, we found strong associations between weather patterns and drinking water source use. These associations have plausible drivers given the intrinsic relationships between the climate variables examined and water availability as well as user preferences for more convenient and/or free water sources. Given the geographic and cultural disparity between the study sites, it is not surprising that there is some diversity in the direction of associations—the conclusion that water use and availability *do* depend on climate is important and lays the groundwork for further studies of mechanisms and implications.

Climate change is anticipated to bring about greater variability in both temperature and rainfall, and low-resource settings are particularly vulnerable to these changes (48, 49). The impact of these changes on WASH uptake are expected to be diverse and vary by setting. Increasing prevalence and severity of drought will have obvious consequences in terms of water scarcity and availability and may lead to selection of less-safe water sources, as we saw in Pakistan, but may also lead to increased willingness to utilize improved water sources (23, 24). Therefore, any future interventions intended to increase access to and use of safe drinking water should consider the potential impacts of climate on WASH use and availability, and develop infrastructure with these potential mechanisms in mind.

Acknowledgements:

Conflicts of Interest: N/A

Funding: N/A

REFERENCES

1. Vos T, Lim SS, Abbafati C, Abbas KM, Abbasi M, Abbasifard M, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. The Lancet. 2020;396(10258):1204-22.
2. IPCC. Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty; 2018.
3. Levy K, Woster AP, Goldstein RS, Carlton EJ. Untangling the impacts of climate change on waterborne diseases: a systematic review of relationships between diarrheal diseases and temperature, rainfall, flooding, and drought. Environmental science & technology. 2016;50(10):4905-22.
4. Mertens A, Balakrishnan K, Ramaswamy P, Rajkumar P, Ramaprabha P, Durairaj N, et al. Associations between high temperature, heavy rainfall, and diarrhea among young children in rural Tamil Nadu, India: a prospective cohort study. Environmental health perspectives. 2019;127(04):047004.
5. Carlton EJ, Woster AP, DeWitt P, Goldstein RS, Levy K. A systematic review and meta-analysis of ambient temperature and diarrhoeal diseases. Int J Epidemiol. 2016 Feb;45(1):117-30.
6. Kraay ANM, Man O, Levy MC, Levy K, Ionides E, Eisenberg JNS. Understanding the Impact of Rainfall on Diarrhea: Testing the Concentration-Dilution Hypothesis Using a Systematic Review and Meta-Analysis. Environ Health Perspect. 2020 Dec;128(12):126001.
7. Jagai JS, Li Q, Wang S, Messier KP, Wade TJ, Hilborn ED. Extreme Precipitation and Emergency Room Visits for Gastrointestinal Illness in Areas with and without Combined Sewer Systems: An Analysis of Massachusetts Data, 2003-2007. Environ Health Perspect. 2015 Sep;123(9):873-9.
8. Murray CJ, Aravkin AY, Zheng P, Abbafati C, Abbas KM, Abbasi-Kangevari M, et al. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. The Lancet. 2020;396(10258):1223-49.
9. Fewtrell L, Kaufmann RB, Kay D, Enanoria W, Haller L, Colford Jr JM. Water, sanitation, and hygiene interventions to reduce diarrhoea in less developed countries: a systematic review and meta-analysis. The Lancet infectious diseases. 2005;5(1):42-52.
10. Kostyla C, Bain R, Cronk R, Bartram J. Seasonal variation of fecal contamination in drinking water sources in developing countries: a systematic review. Science of the Total Environment. 2015;514:333-43.
11. Poulin C, Peletz R, Ercumen A, Pickering AJ, Marshall K, Boehm AB, et al. What Environmental Factors Influence the Concentration of Fecal Indicator Bacteria in Groundwater? Insights from Explanatory Modeling in Uganda and Bangladesh. Environmental Science & Technology. 2020;54(21):13566-78.
12. Hoque BA, Hoque M, Ahmed T, Islam S, Azad A, Ali N, et al. Demand-based water options for arsenic mitigation: an experience from rural Bangladesh. Public health. 2004;118(1):70-7.
13. Wood S, Foster J, Kols A. Understanding why women adopt and sustain home water treatment: insights from the Malawi antenatal care program. Social science & medicine. 2012;75(4):634-42.
14. Simms VM, Makalo P, Bailey RL, Emerson PM. Sustainability and acceptability of latrine provision in The Gambia. Transactions of the Royal Society of Tropical Medicine and Hygiene. 2005;99(8):631-7.
15. Banda K, Sarkar R, Gopal S, Govindarajan J, Harijan BB, Jeyakumar MB, et al. Water handling, sanitation and defecation practices in rural southern India: a knowledge, attitudes and practices study. Transactions of the royal society of tropical medicine and hygiene. 2007;101(11):1124-30.

16. Francis MR, Nagarajan G, Sarkar R, Mohan VR, Kang G, Balraj V. Perception of drinking water safety and factors influencing acceptance and sustainability of a water quality intervention in rural southern India. BMC public health. 2015;15(1):731.
17. Hoat LN, Rheinländer T, Dalsgaard A, Konradsen F. Sanitation behavior among schoolchildren in a multi-ethnic area of Northern rural Vietnam. BMC public health. 2012;12(1):140.
18. Halvorson SJ, Williams AL, Ba S, Dunkel FV. Water quality and waterborne disease in the Niger River Inland Delta, Mali: a study of local knowledge and response. Health & place. 2011;17(2):449-57.
19. Gomes MCRL, Moura EAF, Pedro B, Paulo J, Bezerra MM, Brito OS. Sustainability of a sanitation program in flooded areas of the Brazilian Amazon. Journal of Water, Sanitation and Hygiene for Development. 2015;5(2):261-70.
20. Anthonj C, Fleming L, Godfrey S, Ambelu A, Bevan J, Cronk R, et al. Health risk perceptions are associated with domestic use of basic water and sanitation services—Evidence from rural Ethiopia. International journal of environmental research and public health. 2018;15(10):2112.
21. Kelly E, Shields KF, Cronk R, Lee K, Behnke N, Klug T, et al. Seasonality, water use and community management of water systems in rural settings: Qualitative evidence from Ghana, Kenya, and Zambia. Science of the Total Environment. 2018;628:715-21.
22. Dunston C, McAfee D, Kaiser R, Rakotoarison D, Rambeloson L, Hoang AT, et al. Collaboration, cholera, and cyclones: a project to improve point-of-use water quality in Madagascar. American journal of public health. 2001;91(10):1574-6.
23. Thomson P, Bradley D, Katilu A, Katuva J, Lanzoni M, Koehler J, et al. Rainfall and groundwater use in rural Kenya. Science of the Total Environment. 2019;649:722-30.
24. Thomas EA, Needoba J, Kaberia D, Butterworth J, Adams EC, Oduor P, et al. Quantifying increased groundwater demand from prolonged drought in the East African Rift Valley. Science of the total environment. 2019;666:1265-72.
25. Kelly E, Shields KF, Cronk R, Lee K, Behnke N, Klug T, et al. Seasonality, water use and community management of water systems in rural settings: Qualitative evidence from Ghana, Kenya, and Zambia. Sci Total Environ. 2018 Jul 1;628-629:715-21.
26. Elliott M, MacDonald MC, Chan T, Kearton A, Shields KF, Bartram JK, et al. Multiple household water sources and their use in remote communities with evidence from Pacific Island countries. Water Resources Research. 2017;53(11):9106-17.
27. Tucker J, MacDonald A, Coulter L, Calow RC. Household water use, poverty and seasonality: Wealth effects, labour constraints, and minimal consumption in Ethiopia. Water Resources and Rural Development. 2014;3:27-47.
28. Pearson AL, Zwickle A, Namanya J, Rzotkiewicz A, Mwita E. Seasonal shifts in primary water source type: a comparison of largely pastoral communities in Uganda and Tanzania. International journal of environmental research and public health. 2016;13(2):169.
29. Howard G, Bartram J, Organization WH. Vision 2030: the resilience of water supply and sanitation in the face of climate change: technical report. 2010.
30. Howard G, Calow R, Macdonald A, Bartram J. Climate change and water and sanitation: likely impacts and emerging trends for action. Annual Review of Environment and Resources. 2016;41:253-76.
31. Eisenhauer IF, Hoover CM, Remais JV, Monaghan A, Celada M, Carlton EJ. Estimating the risk of domestic water source contamination following precipitation events. The American journal of tropical medicine and hygiene. 2016;94(6):1403-6.
32. Watts N, Amann M, Arnell N, Ayeb-Karlsson S, Beagley J, Belesova K, et al. The 2020 report of The Lancet Countdown on health and climate change: responding to converging crises. Lancet. 2021 Jan 9;397(10269):129-70.
33. Lal A, Konings P. Beyond reasonable drought: hotspots reveal a link between the ‘Big Dry’and cryptosporidiosis in Australia's Murray Darling Basin. Journal of water and health. 2018;16(6):1033-7.

34. Foster T. Predictors of sustainability for community-managed handpumps in sub-Saharan Africa: evidence from Liberia, Sierra Leone, and Uganda. Environmental science & technology. 2013;47(21):12037-46.
35. Whaley L, Webster J. The effectiveness and sustainability of two demand-driven sanitation and hygiene approaches in Zimbabwe. Journal of Water Sanitation and Hygiene for Development. 2011;1(1):20-36.
36. Kotloff KL, Nataro JP, Blackwelder WC, Nasrin D, Farag TH, Panchalingam S, et al. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. The Lancet. 2013;382(9888):209-22.
37. Organization WH. Safely managed drinking water: thematic report on drinking water 2017. 2017.
38. Bain R, Cronk R, Hossain R, Bonjour S, Onda K, Wright J, et al. Global assessment of exposure to faecal contamination through drinking water based on a systematic review. Tropical Medicine & International Health. 2014;19(8):917-27.
39. Dinku T, Funk C, Peterson P, Maidment R, Tadesse T, Gadain H, et al. Validation of the CHIRPS satellite rainfall estimates over eastern Africa. Quarterly Journal of the Royal Meteorological Society. 2018;144:292-312.
40. Menne MJ, Williams CN, Gleason BE, Rennie JJ, Lawrimore JH. The global historical climatology network monthly temperature dataset, version 4. Journal of Climate. 2018;31(24):9835-54.
41. Beaudoing MRaHK. GLDAS CLM Land Surface Model. In: NASA/GSFC/HSL, editor. Greenbelt, MD, USA: NASA Goddard Earth Sciences Data and Information Services Center (GES DISC); 2019.
42. Raihan MJ, Farzana FD, Sultana S, Haque MA, Rahman AS, Waid JL, et al. Examining the relationship between socio-economic status, WASH practices and wasting. PLOS ONE. 2017;12(3):e0172134.
43. Gomez M, Perdiguero J, Sanz A. Socioeconomic factors affecting water access in rural areas of low and middle income countries. Water. 2019;11(2):202.
44. Vyas S, Kumaranayake L. Constructing socio-economic status indices: how to use principal components analysis. Health policy and planning. 2006;21(6):459-68.
45. Nguyen KH, Operario DJ, Nyathi ME, Hill CL, Smith JA, Guerrant RL, et al. Seasonality of drinking water sources and the impact of drinking water source on enteric infections among children in Limpopo, South Africa. International Journal of Hygiene and Environmental Health. 2021;231:113640.
46. Armstrong A, Hope R, Munday C. Monitoring socio-climatic interactions to prioritise drinking water interventions in rural Africa. npj Clean Water. 2021;4(1):1-3.
47. Levy MC, Collender PA, Carlton EJ, Chang HH, Strickland MJ, Eisenberg JN, et al. Spatiotemporal error in rainfall data: consequences for epidemiologic analysis of waterborne diseases. American journal of epidemiology. 2019;188(5):950-9.
48. Watts N, Amann M, Arnell N, Ayeb-Karlsson S, Beagley J, Belesova K, et al. The 2020 report of The Lancet Countdown on health and climate change: responding to converging crises. The Lancet. 2020.
49. IPCC. Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. In Press; 2021.