# Supporting Information: A Systems Approach to Climate, Water and Diarrhea in Hubli-Dharwad, India

Jonathan Mellor,*,† Emily Kumpel,‡ Ayse Ercumen,¶ and Julie Zimmerman§

†*Department of Civil and Environmental Engineering, University of Connecticut, Storrs, CT 06269*

‡*Aquaya Institute, Nairobi, Kenya*

¶*Division of Epidemiology, University of California at Berkeley, Berkeley, CA 94720*

§*Department of Chemical and Environmental Engineering, Yale University, New Haven, CT 06511*

E-mail: mellor@engineer.uconn.edu

Phone: +1 (860) 486-0548. Fax: +1 (860) 486-2298

This Supporting Information is 12 pages in length, contains ten figures and six tables.

# Supporting Information

An identical analysis was conducted using *E. coli* as the indicator bacteria instead of total coliform. The model validation step compared weekly modeled versus actual weekly diarrhea prevalence. Results were consistent with field results on 56% of weeks (Figure S7). All-cause diarrhea prevalence rates as a function of temperature departures below normal exhibit little linear trend ($R^2$= 0.5918 p = 0.074), while temperature departures above normal show a near linear trend ($R^2$= 0.9598 p < 0.001) of 8.3% per degree Celsius (Figure S8). The overall all-cause diarrhea prevalence results indicate that diarrhea increases by 4.1% (Range: -3.7-11.2%) during the 2011-2030 time period rising to a 17.5% increase (Range: 6.4-32.2%) by the 2046-2065 period, and to 28.0% (Range: 12.5-43.2%) by 2080-2099. Diarrhea prevalence attributable to rotavirus decreases of 6.3% by 2046-2065 and 11.5% by 2080-2099. Diarrhea cases attributable to *Cryptosporidium* increases by 9.1% by 2046-2065 and by 17.6% by 2080-2099 when using *E. coli* as the indicator. Diarrhea cases attributable to *E. coli* are predicted to rise by 2.4% by 2046-2065 and 5.4% by 2080-2099. When rainfall is not considered in the Water Quality Correlation Routines, the model predicts that diarrhea prevalence would be higher than is predicted when considering rainfall. Diarrhea prevalence is predicted to increase by 21.4% (Range: 14.4-28.3%) during the 2011-2030 time period rising to a 37.5% increase (Range: 22.8-48.1%) by the 2046-2065 period, and to 44.7% (Range: 34.0-51.5%) by 2080-2099. Both the SWS and CWF showed remarkable ability to reduce diarrhea rates in both the short and long term (Figure S10), with CWFs becoming increasingly preferred over SWS over time.

## Multiple Imputation

Although water quality samples were taken consistently over the course of 101 separate sampling days between November 2010 and November 2011, there were still more than 264 days over the course of the year without data collection. For modeling purposes, this necessitated

Table S1: Calibrated indicator to pathogen ratios for the three pathogen types used in this study along with the modeled baseline attributable diarrhea fractions.

| Pathogen | Lower Limit | | Upper Limit | | Baseline Diarrhea Fraction | |
|---|---|---|---|---|---|---|
| | *E. coli* | Total Coliform | *E. coli* | Total Coliform | *E. coli* | Total Coliform |
| Rotavirus | $10^{-5.85}$ | $10^{-7.0}$ | $10^{-4.5}$ | $10^{-6.10}$ | 52.8% | 43.3% |
| *E. Coli* | $10^{-0.3}$ | $10^{-1.73}$ | $10^{-0.04}$ | $10^{-1.54}$ | 34.7% | 28.2% |
| *Cryptosporidium* | $10^{-6.0}$ | $10^{-6.10}$ | $10^{-3.85}$ | $10^{-5.50}$ | 17.7% | 15.2% |

the need to fill in the incomplete total coliform measurements using a standard multiple imputation technique.[1] The predictive mean matching routine from Stata Statistics/Data Analysis software was used to impute the missing data (StataCorp, College Station, TX, USA). The imputed dataset has statistically equivalent means, variability and regression parameters as the original data. For consistency with the original field data, the imputation was done 10 different times to represent sampling scenarios whereby between at least one and to at least 10 samples were taken on each sampling day. This 1-10 sample per day range was chosen to represent typical data collection routines from the prior study.[2]

Table S2: Median, minimum and maximum bootstrapped $\alpha$, $N_{50}$ and $k$ parameters for the three pathogens used in equations 2 and 3

| *Pathogen* | $\alpha$ | | | $N_{50}$ | | | $k$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Median | Min | Max | Median | Min | Max | Median | Min | Max |
| *E. Coli* | 0.126 | 0.084 | 0.379 | 2,684,000 | 405,600 | 14,480,000 | NA | NA | NA |
| Rotavirus | 0.262 | 0.099 | 10.000 | 6.801 | 1.223 | 92.630 | NA | NA | NA |
| Cryptsporidium | NA | NA | NA | NA | NA | NA | 0.062 | 0.034 | 0.113 |

Table S3: Parameters for Ceramic Water Filters (CWFs) and Safe Water System (SWS) Log Reduction Values (LRVs) used in Equation 5.

| Pathogen | m | $\overline{LRV_0}$ | $\sigma_{LRV_0}$ |
|---|---|---|---|
| CWF ($t_{CWF-life} <= 182$ days) | -0.003857 | 2.75 | .47 |
| CWF ($t_{CWF-life} > 182$ days) | -0.000667 | 2.08 | .57 |
| SWS ($t_{SWS-life} <= 182$ days) | -0.005857 | 3.00 | .41 |
| SWS ($t_{SWS-life} > 182$ days) | 0.000556 | 1.65 | .54 |

Table S4: *Cryptosporidium* Temperature Modification (Delhi). Adapted from Ajjampur, S. S. R.; Liakath, F. B.; Kannan, A.; Rajendran, P.; Sarkar, R.; Moses, P. D.; Simon, A.; Agarwal, I.; Mathew, A.; O'Connor, R.; Ward, H.; Kang, G. Multisite study of cryptosporidiosis in children with diarrhea in India. Journal of clinical microbiology 2010, 48, 2075-2081. Copyright (2010) American Society for Microbiology.[3]

| Temperature Range (Celsius) | Prevalence Modification |
|---|---|
| $< 30$ | 0.0826 |
| $30 - 34$ | 1.0700 |
| $34 - 35$ | 2.2639 |
| $35 - 36$ | 1.1967 |
| $36 - 37$ | 1.5769 |
| $37 - 38$ | 0.6195 |
| $38 - 39$ | 2.1823 |
| $> 39$ | 1.8021 |

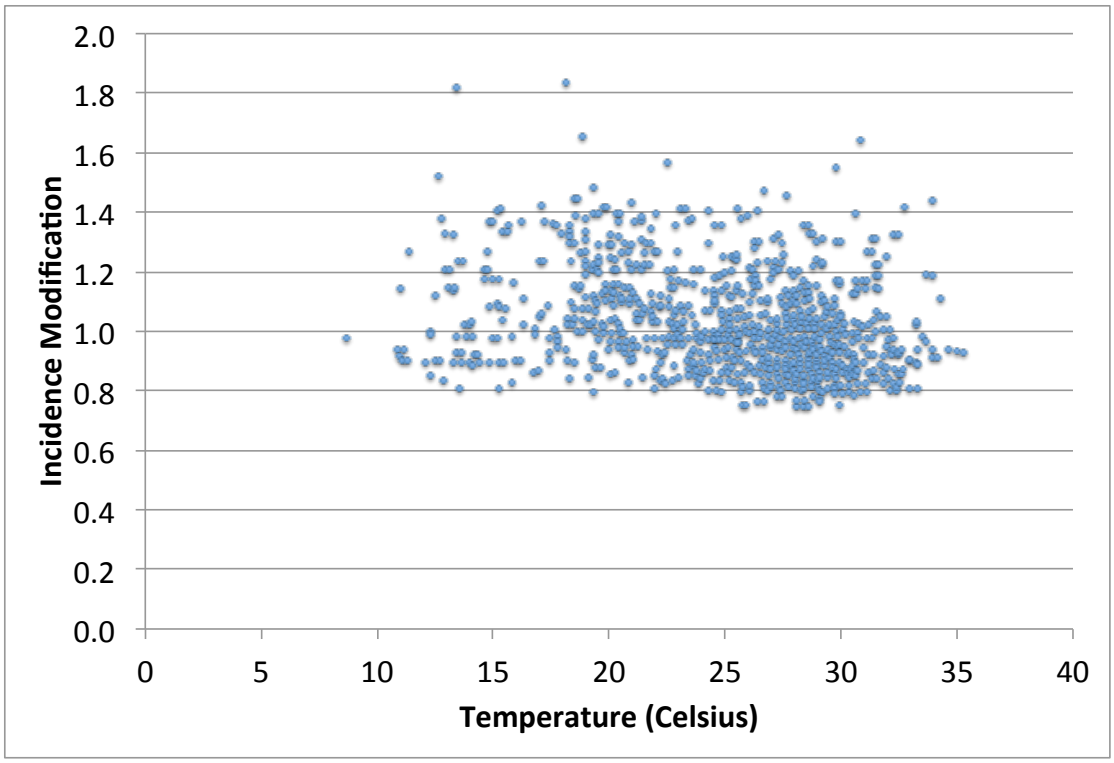


Figure S1: Rotavirus prevalence modification as a function of ambient temperature. Adapted from Jagai, J. S.; Sarkar, R.; Castronovo, D.; Kattula, D.; McEntee, J.; Ward, H.; Kang, G.; Naumova, E. N. Seasonality of rotavirus in South Asia: a meta-analysis approach assessing associations with temperature, precipitation, and vegetation index. PloS one 2012, 7. Copyright (2012) PLOS.

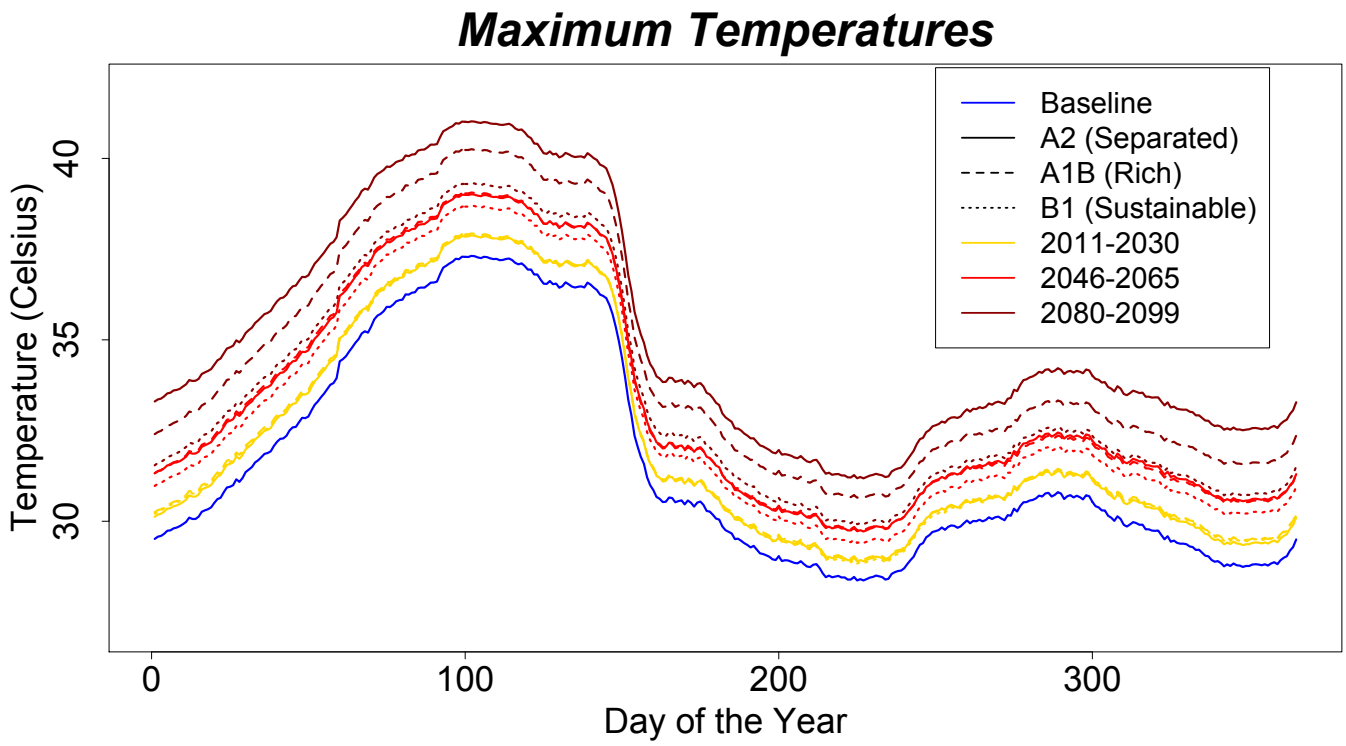


Figure S2: Output of LARS-WG for average daily maximum temperatures for Hubli-Dharwad for each day of the year for three Special Report on Emissions Scenarios (SRES)[5] climate scenarios and three timeframes with individual GCM runs averaged.

Table S5: *Cryptosporidium* Temperature Modification (Vellore). Adapted from Ajjampur, S. S. R.; Liakath, F. B.; Kannan, A.; Rajendran, P.; Sarkar, R.; Moses, P. D.; Simon, A.; Agarwal, I.; Mathew, A.; O'Connor, R.; Ward, H.; Kang, G. Multisite study of cryptosporidiosis in children with diarrhea in India. Journal of clinical microbiology 2010, 48, 2075-2081. Copyright (2010) American Society for Microbiology.[3]

| Temperature Range (Celsius) | Prevalence Modification |
|---|---|
| $< 29$ | 0 |
| $29 - 30$ | 1.5370 |
| $30 - 31$ | 2.0631 |
| $31 - 32$ | 0 |
| $32 - 33$ | 0.4757 |
| $33 - 34$ | 3.3111 |
| $34 - 35$ | 0.4111 |
| $35 - 36$ | 1.3267 |
| $36 - 37$ | 0 |
| $37 - 38$ | 1.9072 |
| $38 - 39$ | 0 |
| $> 39$ | 0.9170 |

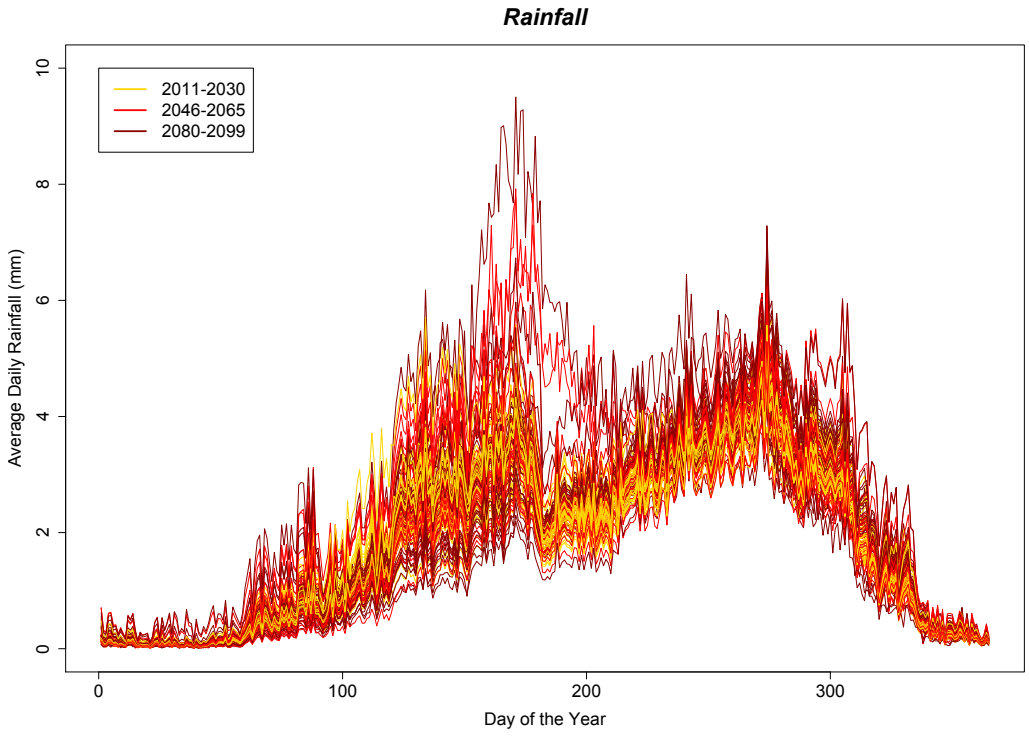


Figure S3: Output of LARS-WG for average daily precipitation for each day of the year for three timeframes. All model runs and all scenarios are shown.

Table S6: Incident rate ratios of *E. coli* incidence per degree Celsius. Adapted from Philipsborn, R.; Ahmed, S. M.; Brosi, B. J.; Levy, K. Climatic drivers of diarrheagenic Escherichia coli: A systematic review and meta-analysis. Journal of Infectious Diseases 2016, Copyright (2016) Oxford University Press.[4]

| Incident rate ratios |
|---|
| 1.11 |
| 1.11 |
| 1.03 |
| 0.99 |
| 1.02 |
| 1.06 |
| 1.22 |
| 1.09 |
| 1.05 |
| 1.12 |
| 1.03 |
| 1.13 |
| 0.80 |
| 1.32 |
| 0.72 |
| 1.38 |
| 1.14 |
| 1.12 |

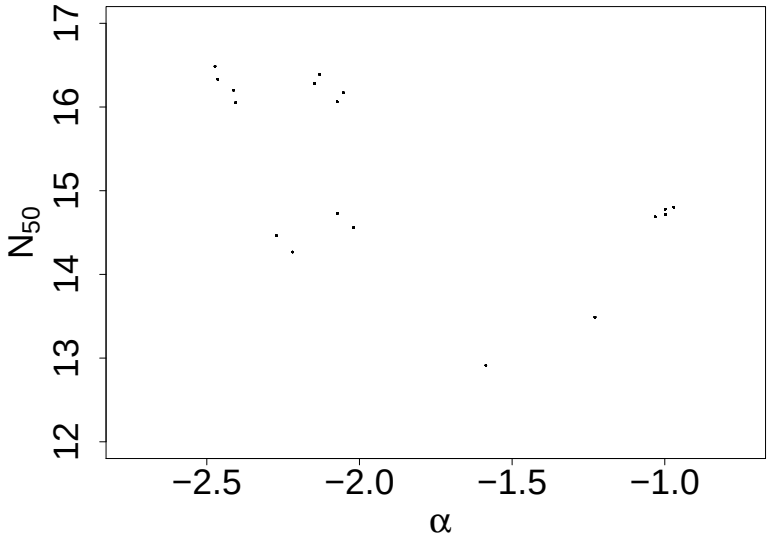


Figure S4: *E. coli* bootstrapped $N_{50}$ and $\alpha$ plotted on a log scale.

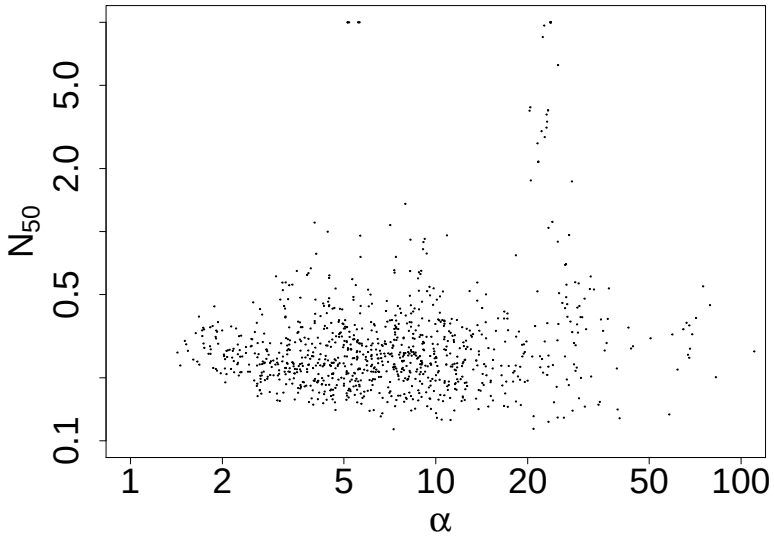


Figure S5: Rotavirus bootstrapped $N_{50}$ and $\alpha$ plotted on a log scale.

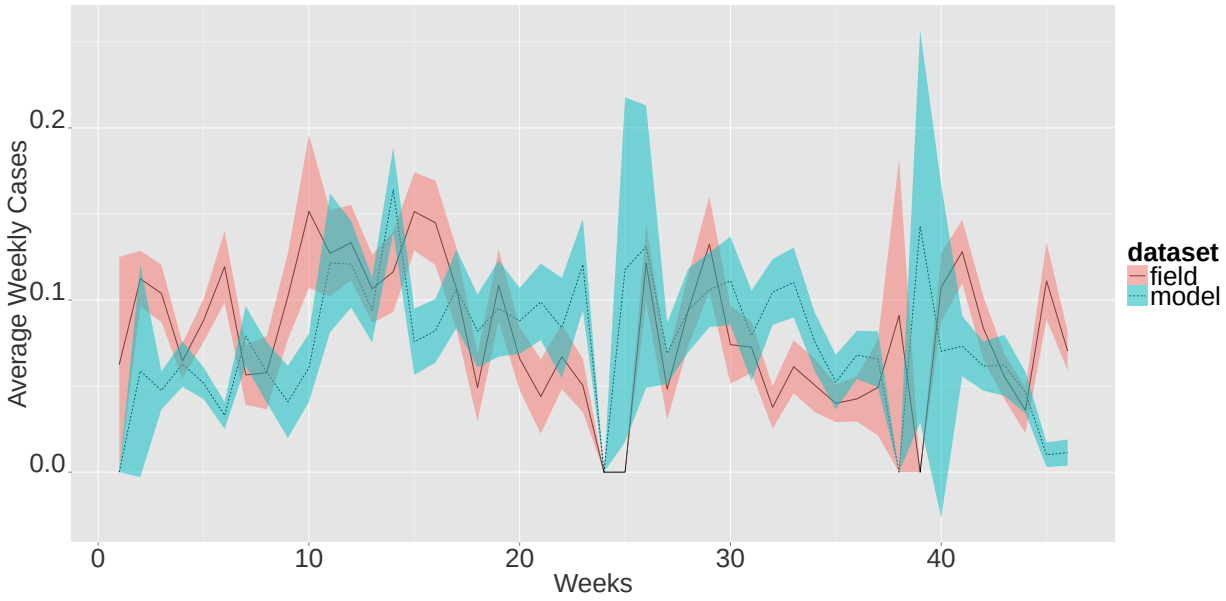


Figure S6: Weekly average diarrhea rates for the field and model results using total coliform as the indicator bacteria. The model is able to correctly predict diarrhea prevalence 85% of the time.

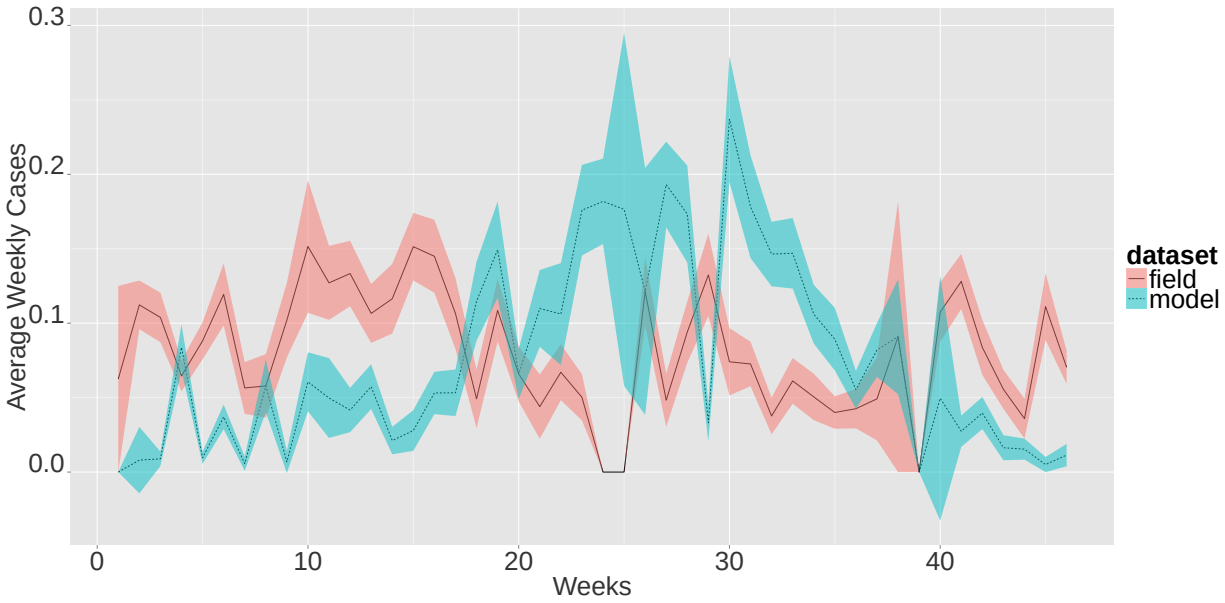


Figure S7: Weekly average diarrhea rates for the field and model results using *E. coli* as the indicator bacteria. The model is able to correctly predict diarrhea prevalence 56% of the time.

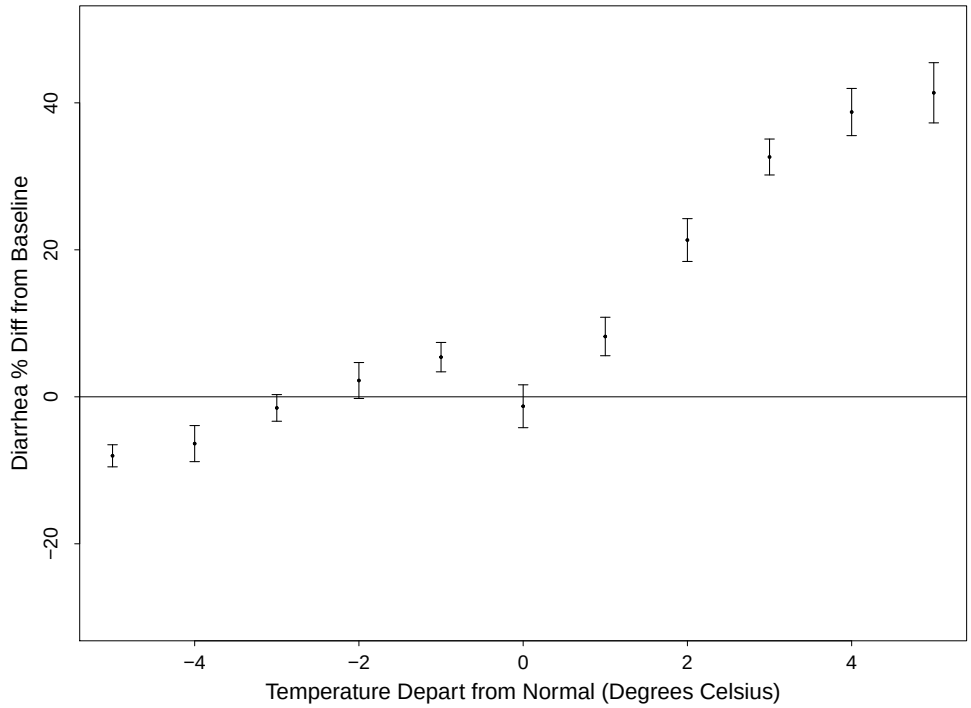


Figure S8: Changes in diarrhea prevalence as a function of temperature using *E. coli* as the indicator bacteria.

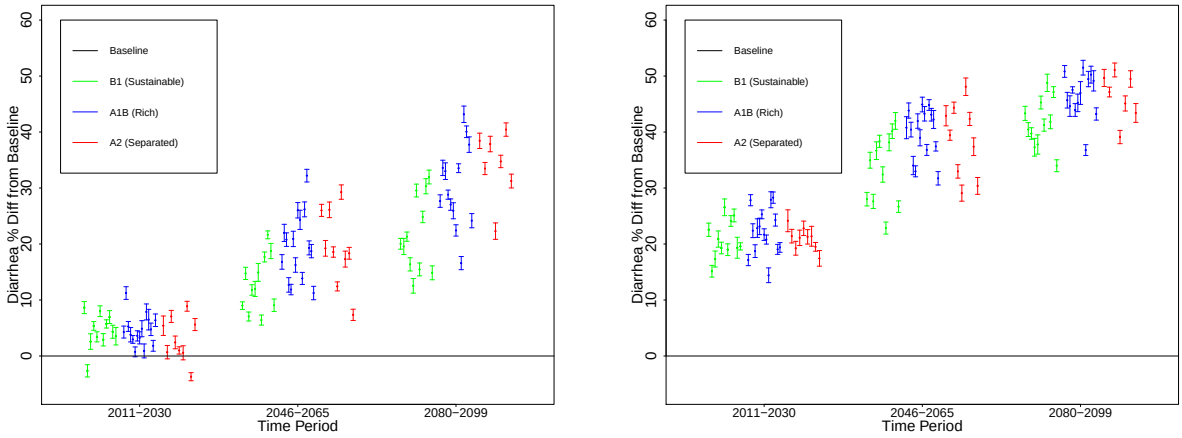


Figure S9: (a) Results from the full model using *E. coli* as the indicator bacteria. All-cause diarrhea prevalence increases with time. Each data point indicates a separate GCM model run that was incorporated into LARS-WG. (b) The full model results excluding rainfall.

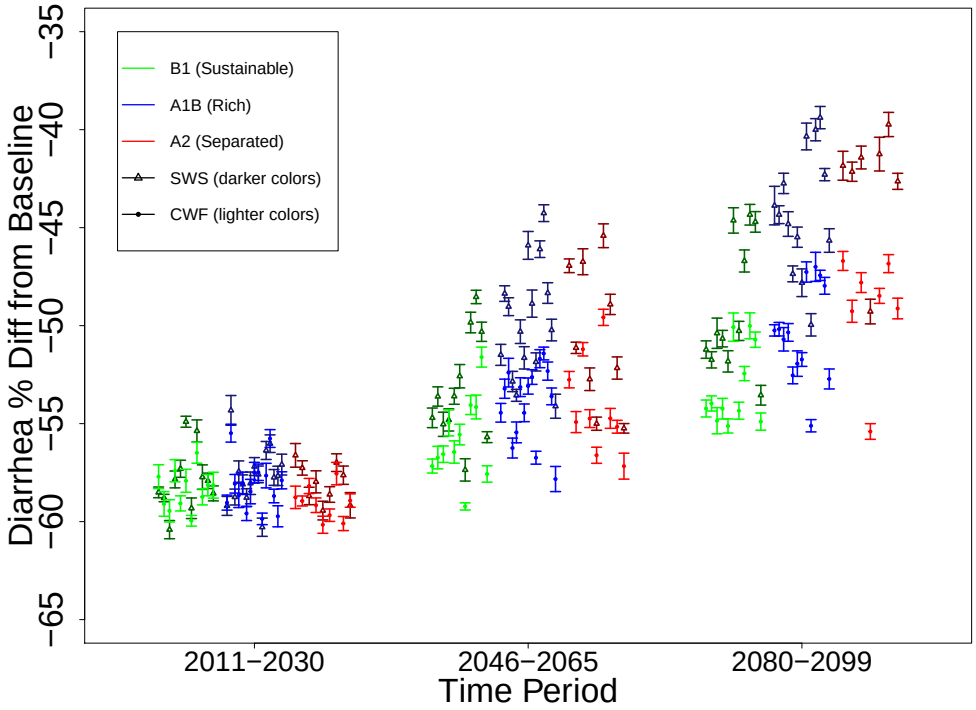


Figure S10: Diarrhea changes for all climate scenarios and time periods using *E. coli* as the indicator bacteria. Data indicate that both CWF and SWS can be effective diarrhea reducing tools, but that CWF are increasingly better towards the end of the century. Darker colors represent SWS and the lighter colors are CWFs.

# References


(1) Rubin, D. B. *Multiple Imputation for Nonresponse in Surveys*; Wiley: New York, 1978.

(2) Kumpel, E.; Nelson, K. L. Comparing microbial water quality in an intermittent and continuous piped water supply. *Water Research* **2013**, *47*, 5176–5188.

(3) Ajjampur, S. S. R.; Liakath, F. B.; Kannan, A.; Rajendran, P.; Sarkar, R.; Moses, P. D.; Simon, A.; Agarwal, I.; Mathew, A.; O'Connor, R.; Ward, H.; Kang, G. Multisite study of cryptosporidiosis in children with diarrhea in India. *Journal of clinical microbiology* **2010**, *48*, 2075–2081.

(4) Philipsborn, R.; Ahmed, S. M.; Brosi, B. J.; Levy, K. Climatic drivers of diarrheagenic Escherichia coli: A systematic review and meta-analysis. *Journal of Infectious Diseases* **2016**,

(5) Nakicenovic, N.; Swart, R. Special report on emissions scenarios. *Edited by Nebojsa Nakicenovic and Robert Swart, pp. 612. ISBN 0521804930. Cambridge, UK: Cambridge University Press, July 2000.* **2000**, *1*.