

Final Project Report

Ella Bartt, 8 Dec 2025

Introduction

Water and food are valuable yet oftentimes inaccessible resources for communities around the world. However, when looking for data on water access in developing countries, many data sources directly compare GDP and percentage of clean water access, leading to an often oversimplified model. For smallholder communities (agricultural regions smaller than farms) in rural Nicaragua, the poverty rate is above 55.9% (Bacon et al. 2021, 4), necessitating a more complex assessment of livelihood risks than those previously employed to assess food and water usage. In the paper “Towards smallholder food and water security: Climate variability in the context of multiple livelihood hazards in Nicaragua,” Bacon et al. propose an interdimensional framework for analyzing these risks. They state, “Like food security, our strategies to assess water insecurity focused on the temporality of access, coping mechanisms due to water-related stress, and perceptions related to quality” (6). The data highlights climate factors like agricultural crop diseases and rainfall as lead drivers for livelihood risk in these rural communities, which I hope to replicate with the available data.

Methodology

The authors created the dataset and paper based on surveys across smallholder communities in Nicaragua between 2014 and 2017. They have a total of 698 household observations, with 311 households in common between the two years. The data consists of three datasets: one for survey household data, the second for food prices (both buying and selling), and the third for precipitation by month and year across the region. I maintained separate datasets for most of the analyses, as that is what the authors did. The survey dataset originally had 698 observations with 258 variables. I subsetted the dataset to have only the columns I was interested in for the analyses, and added `log_income` as another variable. Key means are as follows:

	year	mean_income	mean_food_mo	mean_water_mo	mean_food_score
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1	2014	1278.	3.11	3.59	16.7
2	2017	1309.	2.65	3.08	4.48

	yy	bean_price	maize_price	other_price
	<dbl>	<dbl>	<dbl>	<dbl>
1	2014	32.3	6.28	200.
2	2017	18.9	5.57	151.

Mean_income is in USD and mean_food_mo and mean_water_mo are the means of the number of months households were food or water scarce. The prices are in Nicaraguan reales. There is an overall drop in livelihood risk between 2014 and 2017, largely in part due to the end of a drought, trade and economic policies, and political risk. Overall, most households reported a

nonzero number of food or water lean months, and most coffee farmers saw a decrease in production due to crop disease (Bacon et al., 8).

Visualizations

In order to better understand the data, we look to visualizations and regressions. The following two reproductions can be found on page 8 of the Bacon et al. paper:

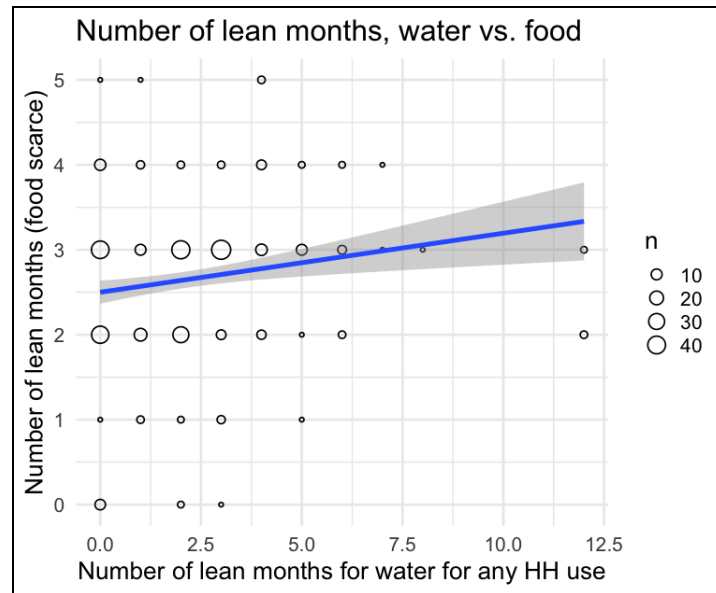


Fig 1a. Number of lean months, water (any household use) and food for 2017 households only. $Y = 2.500 + 0.069x$, p -value = .0028, adjusted R squared = .0237, AIC = 869.20. $N = 334$, size of points is the number of observations.

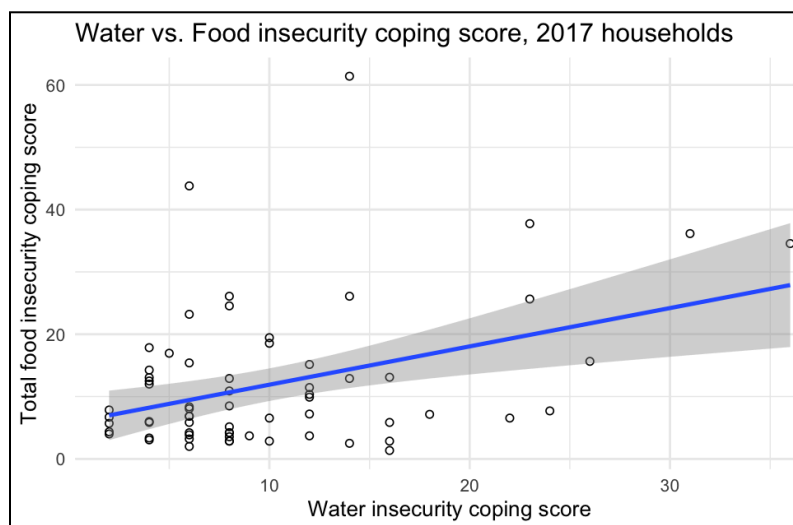


Fig 1b. Insecurity score, water (any household use) and food for 2017 households only. $Y = 5.757 + 0.615x$, Adjusted R-squared = .136, p -value = .00148, AIC = 495. $N = 65$, na values removed.

Figures 1a and b show a statistically significant positive correlation between water and food insecurity measures, both in the number of insecure months and the aggregated coping scores. These both show that water and food insecurity are very closely linked in the study.

Figures 2 and 3 are models from the paper with original visualizations. Both show strong, statistically significant, correlations between 14-month precipitation and log income and water insecurity. These figures also have relatively narrow confidence intervals.

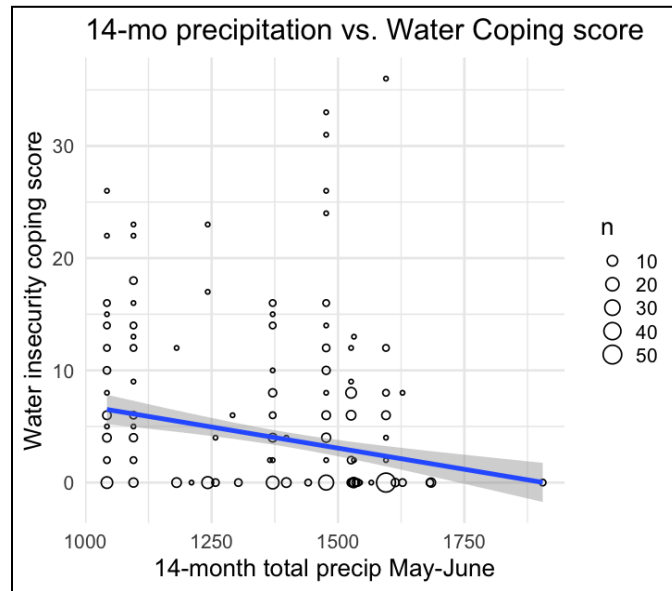


Fig 2. 14-month precipitation and water coping score for 2017 households. $Y = 14.374 - 0.0075x$, adjusted R-squared = .0581, p-value = 4.994e-06, AIC = 2152. N = 334, na values removed.

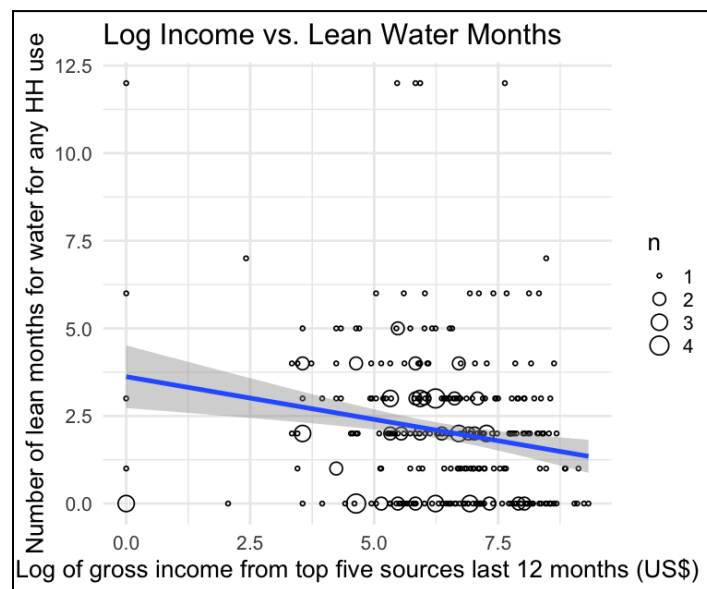


Fig 3. Log of gross income vs. number of lean months for any household use. $Y = 5.739 - 0.421x$, adjusted R-squared = .04223, p-value = .0001076, AIC = 1688.

When logincome, precipitation, and food_coping_score are aggregated into the same linear model, the p-value stays low (2.635e-09), the adjusted R-squared increases to 0.1136, and the equation is:

$$\text{water_coping_score} = 16.299 - 0.0068(\text{precipitation}) - 0.5658(\text{logincome}) + 0.1246(\text{food_coping_score})$$

The above linear model represents the highest adjusted R-squared out of all the models. However, when adding even more predictors (such as number of household members), the added predictors are not statistically significant (high p-values). Farm size was positively correlated with income, as shown in the paper, and thus negatively correlated with food insecurity (Bacon et al., 10). Bacon et al. show very similar, if not identical, results, stating that most other factors were not as statistically significant as those shown above.

Analysis

The statistical insignificance of other household factors is significant in itself. When looking to address food and water insecurity in smallholder communities, one must address the correlated factors for income and precipitation along with country specific data, beyond simply money. The authors of the original paper state:

“In the longer term, addressing the chronic problems of seasonal hunger and malnutrition and securing sustained access to safe drinking water will require access to relevant technologies and broader political economic changes, including transformations of the adverse terms of resource access (e.g., prices received by farmers, land tenure policies, drinking water delivery systems) that smallholders in Nicaragua and elsewhere have faced for decades” (14).

The data from the above models, especially given the statistical insignificance of certain factors, illustrate the shortcomings of simply aggregating any and all variables together into one water or food insecurity model.

Conclusion

The data collected and displayed in the original paper, “Towards Smallholder Food and Water Security: Climate Variability in the Context of Multiple Livelihood Hazards in Nicaragua,” proposes a new framework for addressing food and water insecurity in smallholder communities. This framework goes much further beyond GDP per capita vs. water insecurity, as often shown in studies about water insecurity. The linear regressions that can then be extracted from the replication data support the hypothesis that income, farm size, and rainfall drive livelihood risk, along with political and economic policies for farming trade. The variability in these livelihood risks is sure to increase with political disruptions and climate change, necessitating a multifaceted, unique way to address these issues specific to each community and country.

Works Cited

- Bacon, Christopher M., William A. Sundstrom, Iris T. Stewart, Ed Maurer, and Lisa C. Kelley. 2021. "Towards Smallholder Food and Water Security: Climate Variability in the Context of Multiple Livelihood Hazards in Nicaragua." *World Development* 143 (July): 105468. <https://doi.org/10.1016/j.worlddev.2021.105468>.
- Bacon, Christopher M., William A. Sundstrom, Iris Stewart-Frey, Ed Maurer, and Lisa C. Kelley. 2020. "Replication Data for: Towards Smallholder Food and Water Security: Climate Variability in the Context of Multiple Livelihood Hazards in Nicaragua." Harvard Dataverse. <https://doi.org/10.7910/DVN/ZGPTNI>.