

Quantifying and Modeling Global Water Accessibility including Water Typology

Hichul Chung

Emily Kumpel

Jimi Oke

1-8-21

1 Introduction

In 2017, 29% of the global population (2.2 billion people) did not use a safely managed drinking - water service which is defined as water located on-premise, available, and free from contamination (Organization, 2020). If you meet about 100 people, you can imagine that 29 of the people you meet do not have access to improved drinking water. Another recognition is that many people spend a lot of time collecting water where their access to drinking water supplies located on-premises are not common, about 26.3 percent of the world's population (Cassivi et al., 2018). Where this raises concerns about inequality problems related to the task and time variation to collect water can be inferred in the urban and rural areas (Cassivi et al., 2018). Therefore, there needs to be a great motivation to bring improved drinking water sources to all people to improve their quality of life.

Globally, access to improved drinking water sources and sanitation is increasing. From 1990 to 2012, utilization of improved drinking water sources has increased globally from 76 percent to 89 percent, and utilization of improved sanitation has increased from 45 percent to 64 percent (Fuller et al., 2016). According to the MDG categorization, they consider improved drinking water sources as following: public tap, borehole, protected spring, rainwater collection, bottled water source, and piped water (Bartram et al., 2014). However, these variables do not truly explain if the households have speedy access to retrieve the water. Therefore, we will investigate further on what variables should be included to enhance the accessibility of improved drinking water.

In the next section we discuss relevant prior research on water accessibility quantification. We are interested in what type of modes of transportation are utilized to provide clean and drinking water to the community. Also, we desire to learn more about the time it takes to transport water to the community in a global modeling context. So that we can develop and quantify water accessibility in different countries to compare different trends. Also, we will be exploring the most efficient and effective modes of transportation of accessible water. Therefore, the new models will be able to influence and encourage policymakers in specific countries to improve water accessibility for all. Next we present the data and methods of approaching the quantification of the water accessibility.

2 Relevant work

Organized by chronology and the Categorical.

Onda, LoBuglio, et al. (2012) discusses the global access to safe water by accounting for water quality and the resulting impact on MDG progress. They altered the current Joint monitoring programme (JMP) estimation by including for microbial water quality and sanitary risk utilizing national water quality data. Onda, LoBuglio, et al. (2012) used principal components analysis to analyze the national environment and development indicators and then created models. Overall, our analysis can highlight the potential fatigue illness caused by people transporting the water if it is done manually. How the modes of transportation can affect one's health.

Sima et al. (2013) highlights the relation of the topic where it discovered that over 76 percent of the city's water is utilized by less than ten percent of household piped water. The study is focused in Kisumu Kenya and Sima interviewed 260 informal water business operators. This article revealed a new insight where the majority of the population in Kisumu relied on purchased water from kiosks (1.5 million m^3 per day) and used hand-drawn water-carts (0.75 million m^3 per day)(Sima et al., 2013). Also, we discovered that the water trucking industry utilized the most energy consumption and was high cost: water delivery trucks have the highest per cubic meter energy demand (35 MJ/m^3)(Sima et al., 2013). The article utilized SPSS Version 19.0 to analyze the data.

Ho et al. (2014) highlights a head-to-head comparison of such indicators with other possible distance and time metrics among rural 1,103 households in Nampula province, Mozambique. The researchers found out that Euclidean

line distance to be a good representative for route distance ($R^2 = 0.98$), while self-reported travel time is a poor representative ($R^2 = 0.12$). One of key insights was that, 15-minute decrease in one-way walk time to water source is related to a 41% reduction in diarrhea, improved children's nutrition, and an 11% reduction in under-five child mortality rate (Ho et al., 2014). This research can be a helpful tool to understand how long people are traveling in a rural area in Africa where many of our data focus. Also, understanding how people are traveling to receive the water can better understand the condition of the rural people and implement better policies that help them to thrive.

Bartram et al. (2014) highlights a background information about the Millennium Development Goals. How it has established global targets for drinking water and sanitation access (Bartram et al. 2014). The article indirectly gives much information about progress towards focused targets, directed by international monitoring, has reduced the global disease burden and improved the quality of life (Bartram et al., 2014). It utilizes a variety of sources such as JMP, DHS, MICS surveys to analyze the data of many different locations in the world, therefore, it is a broad scope of study of the progress of accessible and drinking water and sanitation globally. Our research can study further what modes of transportation of clean drinking water would affect the progress of improving drinking water and sanitation in one's community. Also, understanding how long it takes to attain the water can improve the quality of the services provided for the people.

Onda, Crocker, et al. (2014) discusses their new development of typology of country clusters pertain to water and sanitation sector based on congruence through multiple related indicators. Onda, Crocker, et al. (2014) used a hierarchical clustering method and a gap statistic analysis to cluster 156 countries which has 6.75 billion people into a relational clusters. Onda, Crocker, et al. (2014) suggests previous geography or income-based country clustering should be improved by using water and sanitation related indicators. Our research can improvise upon their finding and focus on identifying which water accessibility variables may improve water accessibility for people who are seeking clean drinking water.

Organization (2020) article gives an ambitious plan to reach universal drinking water by the year of 2030 that is affordable and safe followed by the Sustainable Development Goal six goals. The WHO/UNICEF Joint Monitoring Program for Water Supply and Sanitation (JMP) are monitoring progress on drinking water and sanitation since 1990 globally. The article outlines different plans with consistent traits to determine what is deemed to be safe and affordable drinking water and good sanitation. Our study can highlight which modes of transportation of safe and affordable drinking water can develop why one's community may excel in the improvement of safe and affordable drinking water and sanitation. Drinking, sanitation, and hand washing ladder explains the qualitative determination can indirectly help what is a good source of transporting safe and drinking water.

Fuller et al. (2016) highlights the global accessibility to safe drinking water and sanitation and concludes that it has been improving well during the Millennium Development Goal Period. This topic will benefit our research to understand what countries are doing for the mode of transportation of drinking water and sanitation which may have contributed to the improvements. Fuller et al. (2016) chose some countries and studied the overall countries progress which in his methodology revealed sigmoidal trends. Our study can reveal if certain methods of transportation of drinking water and sanitation may have influenced the effectiveness of the improved drinking water met by MDG standards.

Jia et al. (2016) highlights that there is inequality within safe sanitation. There is a great difference in wealth levels in many low-income countries. Therefore, this will hinder the improvement of MDG progress. Jia et al. (2016) discusses future interventions to be engaged with different wealth categories. Our study can assist in understanding improved sanitation increase by making inferences about how they receive water may give the community focusing on sanitation. Overall, income inequality affects who will have access to proper sanitation.

Cassivi et al. (2018) highlights about how there should be more information to the effect of including a 30-minute collection time to monitor access to drinking water. The lack of access to on-premises water sources can cause the use of greater distant alternative sources. Therefore causing the quantity of retrieved water to reduce as the water plateau phenomenon. Cassivi et al. (2018) hypothesizes that the households who must travel further are using unimproved sources. For our research, we can discover what modes of transportation people are traveling to retrieve the water and in what quantity depending on their means of transportation. Therefore, learning how to decrease the travel time of people to retrieve the water because according to the article, it can be a great burden and does not increase the quality of one's life to retrieve water daily and far away from their homes.

Tortajada and Biswas (2018) highlights that sharing the immense improvement at the 2015 convention negatively affected many people because they actually have not acquired clean drinking water. While 2.6 billion people gained access to drinking water, they did not have constructive feedback on the quality of the water. Also, our research can fill in the blank on how did these people actually retrieve the water? Our research can measure the performance of how people retrieve water which may not be in their premises.

Amit and Sasidharan (2019) highlights their data analysis similar to what we will be doing and found out that for non-piped households, collection costs are 22% of the coping costs, while collection costs for piped households are less than 2% of the coping costs. The researchers also realize that there is economic inequality because the majority of the piped households are in households that are considered wealthy households and vice versa. Our research can pinpoint how households are traveling to collect water to improve potential policy-making for water accessibility.

Price et al. (2019) highlights the temporal dynamics of drinking water access and quality in urban slums which is home to about 1 billion people globally. Price et al. (2019) discusses the temporal changes in drinking water access and quality in urban slums may influence on health risks. Price et al. (2019) believes that monitoring temporal dynamics should be considered over three interlinked time scales:

short-term medium term long-term

Price et al. (2019) recommends future research to explore the short-term water access and quality monitoring. Also we should learn to recognize the role of socio-cultural factors that may impact the temporal dynamics of safe water access. Our project can pinpoint additional water accessibility variables that can be used to focus to recognize the time variations as well.

Rawas et al. (2020) highlights the need for different approaches of intermittent water supplies (IWS) to monitor piped water supply perpetually in Peru. This relates and informs one mode of transportation of clean and accessible water. Our research can make inferences about if the customers receive clean water elsewhere, such as water tankers, kiosks, vendors, etc. Overall, the article found that the percent of households with (IWS) was a few percentage points greater than of the reported utilities.

Deshpande et al. (2020) discusses their findings of improved overall water availability and piped water accessibility for the globe (40% to 50.3%). Deshpande et al. (2020) also found that it was the lowest increase in sub-Saharan Africa, where the accessibility was mostly concentrated in urban regions. Deshpande et al. (2020) focuses mostly on understanding trends in diarrhea burden, where the needs are greatest for the improvement of access to safe drinking water and sanitation facilities. Our study can benefit this research by implementing the best and effective modes of transportation of clean water to improve safe drinking water and sanitation. Even with the water pipe access increasing, they can't conclude if the people had clean water as they needed right away. Overall, the access to safe drinking water and sanitation improved globally between 2000 and 2017, but dis-proportionality undermines reaching the Sustainable Development Goals.

Goal 6 (2020) explains specific targets that have goals and indicators to track global water accessibility. The article quantifies and provides Graphs: safe drinking water, safe sanitation, and hygiene, safe sanitation and hygiene, ambient water quality, water use efficiency, levels of freshwater stress, integrated water management, trans-boundary water cooperation, protect and restore water-related ecosystems, water and sanitation support, local participation in sanitation management. They have many ambitious plans to achieve by 2020 and 2030, and our research can find a correlation to why achieving this may be possible.

Gaps and areas for improvement

We are interested in analyzing the impact of transportation modes (usage/ownership) on water accessibility. We also desire to see an implementation of this statistical research piloted in one of the countries to see if the water accessibility variables predictors can improve the water accessibility for its country.

Data

In this section, describe the data sources and the variables, along with data preparation procedures, variable selection criteria, etc. You should have a table summarizing the water accessibility variables and their definitions. Another table can summarize the explanatory variables (e.g. transportation-related, income-related, etc). Also discuss the number of countries, etc. I've started a table as an example. You can add other columns if needed.

Data Preparation

Data structure this study were collected from the Demographic Health Survey (DHS) program. We used the STATCompiler from the DHS Program to develop customized table based on focusing on water, transportation, and wealth indicators across 78 countries and from 1990 to 2019.

We preprocess the data by differentiating the different surveys collected by the DHS program between DHS and Malaria Indicators Survey (MIS) through excel. After, we used R to develop and simplify the table with the most recent surveys for the most up to date analysis for the 78 countries. There was two sets of data of population and households. Therefore, we decided to remove the population data from the table to simplify because population and households data were identical.

Next, there were many missing values from the data. Therefore, we had to do a simple missing value analysis. We determined to avoid using water, transportation, or wealth variables which had greater than 50% missing values. Ultimately, simplified data frame was 76 countries by 28 variables. After, two different tables are generated where one focused on the water accessibility variables and a explanatory variables. We selected 11 water accessibility variables based on the sources available such as location, time, and methodology (Price et al., 2019). The water accessibility variable was organized in the following table to allow for quick visualization for table 1. The other variables were used as an explanatory variables for determining the water accessibility predictors in Table 2.

| Variable | Description | Definition |
|----------|--|--|
| ptap | Households using a public tap/standpipe | Percentage of households whose main source of drinking water is a public tap/standpipe. |
| bore | Households using a tube well/borehole | Percentage of households whose main source of drinking water is a tube-well/borehole. |
| pspr | Households using a protected spring | Percentage of households whose main source of drinking water is a protected spring. |
| uspr | Households using an unprotected spring | Percentage of households whose main source of drinking water is an unprotected spring. |
| pwel | Households using a protected well | Percentage of households whose main source of drinking water is a protected well. |
| uwel | Households using an unprotected well water | Percentage of households whose main source of drinking water is an unprotected well. |
| surw | Households using surface water | Percentage of households whose main source of drinking water is surface water. |
| ties | Households with water 30 minutes or less away round trip | Percentage of households with water 30 minutes or less away round trip. |
| tmor | Households with water more than 30 minutes away round trip | Percentage of households with water more than 30 minutes away round trip. |
| rain | Households using rainwater | Percentage of households whose main source of drinking water is rainwater. |
| phom | Households using water piped into dwelling | Percentage of households whose main source of drinking water is water piped into the dwelling. |

TABLE 1 Summary of water accessibility variables (national averages at household level)

| Variable | Description | Definition |
|----------|---|---|
| imws | Households using an improved water source | Percentage of households whose main source of drinking water is an improved source. |
| pipy | Households using water piped into yard/plot | Percentage of households whose main source of drinking water is water piped into the yard or plot. |
| truc | Households using tanker truck | Percentage of households whose main source of drinking water is tanker truck. |
| ctan | Households using cart with tank | Percentage of households whose main source of drinking water is cart with tank. |
| bott | Households using bottled water | Percentage of households whose main source of drinking water is bottled water/demi john. |
| uiws | Households using an unimproved water source | Percentage of households whose main source of drinking water is an unimproved source. |
| othw | Households using other water source | Percentage of households whose main source of drinking water is other. |
| iwsp | Households with improved water source on the premises | Percentage of the de jure population living in households with an improved water source on the premises. |
| basw | Households with basic water service | Percentage of households with basic water service, defined as an improved water source with either water on the premises or round-trip collection time is 30 minutes or less. |
| liws | Households with limited water service | Percentage of households with limited water service, defined as an improved water source with round-trip collection time greater than 30 minutes. |
| prem | Households with water on the premises | Percentage of households with water on the premises. |
| bicy | Households possessing a bicycle | Percentage of households possessing a bicycle. |
| acar | Households possessing an animal drawn cart | Percentage of households possessing an animal drawn cart. |
| mcyc | Households possessing a motorcycle | Percentage of households possessing a motorcycle. |
| pcar | Households possessing a private car | Percentage of households possessing a private car. |
| boat | Households possessing a boat with a motor | Percentage of households possessing a boat with a motor. |
| wigc | Wealth index Gini coefficient | The Gini coefficient indicates the level of concentration of wealth in the country. |

TABLE 2 Summary of explanatory variables (national averages at household level)

Methods

Preamble

We analyzed data from the Demographic Health Survey (DHS) and the Multiple Indicator Cluster Survey (MICS) (1990-2020) to highlight water accessibility predictors in 78 countries throughout the world.

Cluster analysis

Classification model to explain water accessibility predictors

This is the second phase of the work, which you should hopefully start in January

Results and Discussion

Summary of variables

Here, you show and discuss the data distributions (e.g. histograms, correlation plots, etc)

Water accessibility typologies

Show the typology/cluster results and discuss.

FIGURE 1 Spider plots of water accessibility variables by typology

Model results

Conclusion

Limitation

References

- Amit, R. K. and Subash Sasidharan (Feb. 1, 2019). “Measuring Affordability of Access to Clean Water: A Coping Cost Approach”. In: *Resources, Conservation and Recycling* 141, pp. 410–417.
- Bartram, Jamie et al. (Aug. 2014). “Global Monitoring of Water Supply and Sanitation: History, Methods and Future Challenges”. In: *International Journal of Environmental Research and Public Health* 11.8, pp. 8137–8165.
- Cassivi, A. et al. (Aug. 1, 2018). “Access to Drinking Water: Time Matters”. In: *Journal of Water and Health* 16.4, pp. 661–666.
- Deshpande, Aniruddha et al. (Sept. 1, 2020). “Mapping Geographical Inequalities in Access to Drinking Water and Sanitation Facilities in Low-Income and Middle-Income Countries, 2000–17”. In: *The Lancet Global Health* 8.9, e1162–e1185.
- Fuller, James A. et al. (Jan. 15, 2016). “Tracking Progress towards Global Drinking Water and Sanitation Targets: A within and among Country Analysis”. In: *Science of The Total Environment* 541, pp. 857–864.
- Goal 6 (2020). *Goal 6: Clean Water and Sanitation - SDG Tracker*. Our World in Data. URL: <https://sdg-tracker.org/water-and-sanitation> (visited on 09/03/2020).
- Ho, Jeff C., Kory C. Russel, and Jennifer Davis (Mar. 1, 2014). “The Challenge of Global Water Access Monitoring: Evaluating Straight-Line Distance versus Self-Reported Travel Time among Rural Households in Mozambique”. In: *Journal of Water and Health* 12.1, pp. 173–183.
- Jia, Peng et al. (July 12, 2016). “High-Resolution Spatial Distribution and Estimation of Access to Improved Sanitation in Kenya”. In: *PLOS ONE* 11.7, e0158490.
- Onda, Kyle, Jonny Crocker, et al. (Mar. 1, 2014). “Country Clustering Applied to the Water and Sanitation Sector: A New Tool with Potential Applications in Research and Policy”. In: *International Journal of Hygiene and Environmental Health* 217.2, pp. 379–385.
- Onda, Kyle, Joe LoBuglio, and Jamie Bartram (Mar. 2012). “Global Access to Safe Water: Accounting for Water Quality and the Resulting Impact on MDG Progress”. In: *International Journal of Environmental Research and Public Health* 9.3 (3), pp. 880–894.
- Organization, World Health (2020). *Drinking-Water*. URL: <https://www.who.int/news-room/fact-sheets/detail/drinking-water> (visited on 09/30/2020).

- Price, Heather, Ellis Adams, and Richard S. Quilliam (June 25, 2019). “The Difference a Day Can Make: The Temporal Dynamics of Drinking Water Access and Quality in Urban Slums”. In: *Science of The Total Environment* 671, pp. 818–826.
- Rawas, Farah, Rob Bain, and Emily Kumpel (Mar. 12, 2020). “Comparing Utility-Reported Hours of Piped Water Supply to Households’ Experiences”. In: *npj Clean Water* 3.1 (1), pp. 1–9.
- Sima, Laura C. et al. (Mar. 1, 2013). “Water Flows, Energy Demand, and Market Analysis of the Informal Water Sector in Kisumu, Kenya”. In: *Ecological economics : the journal of the International Society for Ecological Economics* 87, pp. 137–144.
- Tortajada, Cecilia and Asit K Biswas (Oct. 1, 2018). “Achieving Universal Access to Clean Water and Sanitation in an Era of Water Scarcity: Strengthening Contributions from Academia”. In: *Current Opinion in Environmental Sustainability*. Sustainability Science 34, pp. 21–25.