# Predicting arsenic contamination in groundwater wells in the Bengal Basin



VANDERBILF UNIVERSITY Christopher M. Tasich, Jonathan Gilligan, and Steven L. Goodbred Department of Earth & Environmental Science, Vanderbilt University, Nashville, TN Corresponding Author: Chris Tasich (chris.tasich@vanderbilt.edu)



### **THE GANGES-BRAHMAPUTRA DELTA**

People living in Bangladesh and West Bengal have long been plagued by drinking water insecurity despite often being surrounded by significant volumes of water. Many of the regions inhabitants have relied on surface waters for drinking water. One of the most common forms of surface water is trapped monsoonal rainwater in small community ponds. Historically, this was one of the most dominant methods of obtaining freshwater. However, these small community ponds are often contaminated with pathogens that induce diarrheal disease. Throughout the 1970s, millions of groundwater wells were installed in an effort to combat high rates of infant mortality rates due to diarrheal disease. In the 1990s, high rates of arsenicosis and related cancers led to the discovery of high concentrations of arsenic in the groundwater (Nickson et al., 1998, Chowdhurry et al., 2000, McArthur et al., 2001, Harvey et al., 2002). Estimates suggest that 7 to 11 million drinking water wells are contaminated by high concentrations (>50 ppm) of naturally occurring arsenic.

## **STRATIGRAPHIC DATA**

We have collected ~15,000 stratigraphic sediment samples from ~10 transect and ~400 total cores throughout Bangladesh. Approximately 5,000 of these samples have been analyzed for grain size, magnetic susceptibility, chemical composition, and organic matter content. We will use these as predictors in our arsenic investigation. Additionally, we will apply advanced data science techniques (e.g. self organizing maps) to predict stratigraphy within the region. Below (Fig. 3), we have included the stratigraphic data analyzed by Patrick (2016). This transect will serve as our pilot study.







**Figure 1** | Map of the Ganges-Brahmaputra delta. The inset in the bottom right shows arsenic concentrations within Bangladesh based on a survey conducted by Kinniburgh and Smedley (2002).

## **ARSENIC IN THE GROUNDWATER**

\_\_\_\_\_ 200 km \_\_\_\_\_

The Ganges river is extremely active and has dissected large portions of previously deposited sediments, introducing significant subsurface heterogeneity and complicating the search for safe drinking waters. In some regions such as Araihazar, Bangladesh, this heterogeneity is extremely high making statistical prediction difficult (van Geen et al. 2002), though, this may not be the case for other parts of the delta. We will use the Bangladesh Arsenic Mitigation Water Supply Program (BAMSWP) dataset of ~4.5 million wells to train a predictive model for arsenic concentration in the groundwater. Below (Fig. 2), we have plotted a sampling of the BAMWSP data along a transect across both the tidal and fluvial portions of the delta. As the dataset is very large, we subset the data in our initial investigation to coincide with stratigraphic data collected by Patrick (2016). The BAMWSP contains well depth and arsenic concentrations. Most of these wells contain arsenic concentrations >25 ppm. A confounding issue with the BAMWSP data is that it is generated from Hach test strips which require human interpretation of the arsenic concentration leading to discretized data.

**Figure 3** | Stratigraphic data from Patrick (2016) showing generalized stratigraphy (left) with associated magnetic susceptibility (right, a) and bulk strontium concentrations (right, b).

## **POSTGRESQL DATABASE**

We have built a PostgreSQL database to store environmental and social data from Bangladesh. This structured data will allow us to connect disparate datasets and leverage high-performance computing techniques. Below (Fig. 4), we have a included a sample database schema for our proposed database.





Figure 4 | Sample database schema for connecting environmental and social data from Bangladesh.

#### **RESEARCH QUESTIONS**

#### Regarding stratigraphy...

How does stratigraphy vary throughout Bangladesh?
 Can we use advanced data science techniques (e.g. self organizing maps) to predict stratigraphic arrangements?



Figure 2 | A. Map of the sample BAMWSP data that coincides with the stratigraphic data from Patrick (2016).
B. Frequency plot of arsenic concentrations found within wells along the transect. C. Frequency plot of well depth along the transect. D. Arsenic concentrations of wells at a given depth.

#### Regarding arsenic concentrations in the groundwater...

3. How do arsenic concentrations in the groundwater vary with

depth in Bangladesh?

4. Can we predict arsenic concentrations using statistical learning techniques (e.g. splines)? Can we use hierarchical modeling?



- 1. Kinniburgh, D. G. & Smedley, P. L. Arsenic contamination of groundwater in Bangladesh. (2001). Available at: http://nora.nerc.ac.uk/id/eprint/11986/. (Accessed: 14th May 2019)
- 2. McArthur, J. M., Ravenscroft, P., Safiulla, S. & Thirlwall, M. F. Arsenic in groundwater: Testing pollution mechanisms for sedimentary aquifers in Bangladesh. *Water Resources Research* 37, 109–117 (2001).
- 3. Harvey, C. F. et al. Arsenic Mobility and Groundwater Extraction in Bangladesh. Science 298, 1602–1606 (2002).
- 4. Nickson, R. et al. Arsenic poisoning of Bangladesh groundwater. Nature 395, 338 (1998).
- 5. Chowdhury U K et al. Groundwater arsenic contamination in Bangladesh and West Bengal, India. Environmental Health Perspectives 108, 393–397 (2000).
- 6. Nickson, R. T., McArthur, J. M., Ravenscroft, P., Burgess, W. G. & Ahmed, K. M. Mechanism of arsenic release to groundwater, Bangladesh and West Bengal. Applied Geochemistry 15, 403-413 (2000).
- 7. van Geen, A. et al. Promotion of well-switching to mitigate the current arsenic crisis in Bangladesh. Bull World Health Organ 80, 732–737 (2002).
- 8. Patrick, M. G. Stratigraphic evolution of the Ganges-Brahmaputra lower delta plain and its relation to groundwater arsenic distributions. (2016).