The Pennsylvania State University

The Graduate School

Department of Geosciences

SPATIAL VARIATION OF CHEMICAL CONSTITUENTS IN NATURAL WATERS AND THEIR RELATION TO INCIDENCE OF BURULI ULCER IN GOLD-MINING REGIONS OF GHANA

A Thesis in

Geosciences

by

Julianne Hagarty

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

May 2012

The thesis of Julianne Hagarty was reviewed and approved* by the following:

Kamini Singha
Associate Professor of Geosciences
Thesis Advisor

Erica A.H. Smithwick
Assistant Professor of Geography and Ecology

Richard R. Parizek
Professor of Geosciences

Chris J. Marone
Professor of Geosciences
Associate Department Head of Graduate Programs

*Signatures are on file in the Graduate School

ABSTRACT

Buruli ulcer, an emerging bacterial disease caused by *Mycobacterium ulcerans*, largely affects poor rural populations in tropical and sub-tropical countries, predominantly in western Africa. Occurring in at least thirty-three countries worldwide, it is the third most common Mycobacterial disease after tuberculosis and leprosy. Buruli ulcer is characterized by ulcerative lesions and can lead to severe deformity if untreated. While methods of treatment for Buruli ulcer are well known and have a high rate of success, the mode of transmission of Buruli ulcer remains elusive. Multiple hypotheses have been put forward in the search for the vector for this disease. Studies of Buruli ulcer to date seem to conclude that water is, in some way, closely related to the transmission of this disease. In particular, changes in water quality due to changes in land use may contribute to the emergence of Buruli ulcer. I hypothesize that stagnant pools, especially those with low pH, high metals, nitrogen, and phosphorus concentrations, will provide a favorable environment for M. ulcerans growth and transmission.

To determine the relationship between water chemistry and Buruli ulcer incidence, water samples were collected from five communities within Ghana: four in the southern part of the country (three Buruli-endemic communities: Pokukrom, Betenase, and Ayanfuri, and one control: Kedadwen) and one non-endemic community (Nangruma) in the north. The southern control accounts for differences between endemic and non-endemic communities with similar land uses and geological setting. The northern community has experienced massive floods in recent years, and it is suspected that this may trigger the onset of Buruli ulcer in the community.

Data from the rainy season in June and July 2011 show distinct differences in chemistry between water bodies. Gold-mining pits and pools of stagnant water have a significantly different chemical signature than rivers and naturally-occurring swamps. Trace metals, elements thought to aid in the preferential growth of *M. ulcerans*, are present in much higher concentrations in mining pits and stagnant pools than in other water bodies. A pilot study conducted during the dry season in January and February 2011 using different methods yielded similar results, showing a difference between stagnant and flowing water bodies based predominantly on trace metal concentrations. This work serves as a first step toward characterizing the environmental niche for *M. ulcerans* growth and persistence in the environment.

TABLE OF CONTENTS

List of Figures vi

List of Tables vii

Acknowledgments viii

Introduction 1

- Description of Field Sites 6
  - Geology 7
  - Climate 8
  - Agriculture 9
  - Mining 10

Field Methods 12

Statistical Methods 14

- Analysis of Variance 14
- Cluster Analysis 15
- Principal Component Analysis 17

Results and Discussion 19

- Univariate Analysis 19
- Multivariate Analysis 28
- Data Complications 32

Conclusions and Future Work 34

References 36

- Appendix A GPS Coordinates of Sampling Locations 42
- Appendix B June-July 2011 Data 45
- Appendix C Scores and Loadings for Principal Components Analysis 49
  - Scores 49
  - Loadings 53
- Appendix D Sampling Protocols and Field Data Sheet 56
- Appendix E Rainfall Data Analysis 60
- Appendix F January 2011 Data Analysis 64
- Appendix G January 2011 Data 69
- Appendix H WATSAN Data Analysis 72
  - Cluster Analysis 73
  - Principal Components Analysis 78

List of Figures

Figure 1. Abandoned galamsey (artisanal gold-mining) pit, Pokukrom, Central region, Ghana. 4

Figure 2. Map of study areas....6

Figure 3. Average monthly rainfall in study areas, 1960-2010. Data from GMet (2010). Average rainfalls for Pokukrom and Ayanfuri are based on a gauging station in Dunkwa, Betenase on Sefwi Bekwai, Kedadwen on Tarkwa, and Nangruma on Navrongo....9

Figure 4. Box-and-whisker plots of pH organized by a) water body type, and b) community. Red indicates an endemic community. n values show the number of samples in each group. P-values represent the statistical significance of these factors from one-way ANOVA....20

Figure 5. Box-and-whisker plots of a) arsenic organized by water body type, b) arsenic organized by community, c) cadmium organized by water body type, d) cadmium organized by community, e) copper organized by water body type, f) copper organized by community, g) iron organized by water body type, h) iron organized by community, i) lead organized by water body type, j) lead organized by community, k) selenium organized by water body type, l) selenium organized by community, m) zinc organized by water body type, and n) zinc organized by community. Red indicates endemic communities. n values display the number of samples in each group. P-values represent the statistical significance of these factors from one-way ANOVA. 25

Figure 6. Box-and-whisker plots of a) nitrate organized by water body type, b) nitrate organized by community, c) phosphate organized by water body type, and d) phosphate organized by community. Red indicates endemic communities. Values of n display the number of samples in each group. P-values represent the statistical significance of these factors from one-way ANOVA. 27

Figure 7. Groundwater sources. a) borehole, and b) hand-dug well. 28

Figure 8. R-mode cluster analysis....29

Figure 9. PCA of June-July 2011 data, with a) principal components 1 and 2, coded by type of water body, b) principal components 1 and 2, coded by community, c) principal components 1 and 3, coded by type of water body, and d) principal components 1 and 3, coded by community. In plots coded by type of water body, dark blue represents boreholes, light blue wells, green rivers and streams, brown swamps, orange galamsey pits, and red BU hot spots. In plots coded by community red represents Pokukrom, orange Betenase, yellow Kedadwen, green Ayanfuri, and blue Nangruma. Principal Component 1 is dominated by Cd, Cu, Fe, Pb, and Se; Principal Component 2 is dominated by PO_4, NH_4, and F; Principal Component 3 is dominated by Cl, hardness, NO_2, and NO_3. 32

List of Tables

Table 1. Summary of Study Areas. Bold communities are those selected for water chemistry study. 7

Acknowledgments

There is a very long list of people who made this work possible, on personal and professional notes. First and foremost, my parents and my sisters for their unending love and support: thank you for talking me down from metaphorical ledges and for celebrating my small successes.

Working in Ghana presented some unique challenges for me. Thank you to everyone on the reBUild team for putting up with me. Particular thanks to Petra Tschakert for leading the team; to Edith Parker and David Cohen for making field work fun; to Erica Smithwick for expanding my understanding of statistics; to David Ferring, Lindsay Kromel, Saviour Mantey, Rose Sandow, Yakubu Iddrisu Goro, Charles Abbey, and Emmanuel Effah for their incredible amount of effort toward this project; to David Azanu and Bernadette Atosona for their work in the field and laboratory to make this thesis possible; and to Heidi Hausermann for being a wonderful travel companion. Thank you to Kamini Singha for giving me the opportunity to work on this project. It has been a tremendously fulfilling experience.

Thank you to the hydro group (Aaron Regberg, Michael Fitzgerald, Katy Gerecht, Ryan Swanson, Rachel Urban, Rachel Lauer, and Katie Olcott) for your support and assistance and for making me think about my research in new ways. And for looking at pictures of lesions once a semester; I know it wasn't easy. Thank you to Laura Liermann for being so accommodating in giving me access to the geochemistry laboratory for acid-washing my sample bottles, and to Mark Patzkowsky for taking my knowledge of multivariate statistics from zero to what's presented in this thesis. Thank you also to Gustav Osiakwan for sharing WATSAN borehole chemistry data.

To the current and former members of Central Pennsylvania Community Housing, the co-op, Houseasaurus: my darling Coopers, thank you for reminding me continuously of the importance of work-life balance. Your love has made these years the best of my life. Particular thanks to Tristan Cleveland for baking and for commiserating when work was overwhelming; and to Dan Kikkert, Matt Potako, Bina Peters, Justin Stomp, and Sean Wilson for hugs, porch-sitting, pool games, and assorted antics.

Lastly, I must acknowledge National Science Foundation CNH Award #0909447 to Drs. Petra Tschakert, Erica Smithwick, Kamini Singha, Joseph Oppong, Edith Parker, and Annmarie Ward, for providing the funding that made this work possible.

Introduction

The Millennium Development Goals are a set of eight goals proposed in 2000 by the United Nations in an effort to improve quality of life in developing countries by 2015 (United Nations, 2000). Included in this list is a goal to combat major diseases such as HIV/AIDS, malaria, and tuberculosis. Lost in this list, however, is a class of diseases known as neglected tropical diseases. Neglected tropical diseases share a common theme: while disproportionally affecting the world's poorest and most marginalized populations (as do malaria and tuberculosis), they receive much less funding for research and treatment than do the more major diseases. In its first-ever report on neglected tropical diseases, the World Health Organization (2010) states that

> working to overcome the impact of neglected tropical diseases represents a largely untapped development opportunity to alleviate the poverty of many populations and thereby make a direct impact on the achievement of the Millennium Development Goals as well as fulfilling WHO's mission: ensuring attainment of the highest standard of health as a fundamental human right of all peoples.

One such disease is Buruli ulcer (BU). BU is a necrotizing skin disease caused by *Mycobacterium ulcerans*; it is the third most common mycobacterial disease in the world after tuberculosis and leprosy (Merritt et al., 2005). While BU is typically non-fatal, it can result in severe deformity and medical complications if not promptly and properly treated. Cases of BU have been reported in over 30 tropical and subtropical countries, typically in poor rural communities, and most frequently in West Africa (WHO, 2007). In 2010, more than 1,000 cases of BU were reported in Ghana alone (WHO, 2011).

This disease is an enigma for scientists, as its mode of transmission is unknown.

Treatment is well-understood clinically, although in practice, it is expensive, long, and often poorly managed (WHO, 2012; Hausermann, in prep). Recently, many scientists have turned their attention to determining the mode of transmission for BU in an effort to prevent this disease.

An extensive body of research suggests that BU is prevalent in areas subject to rapid environmental modification such as logging, irrigation, agriculture, mining, or dam construction (Hayman, 1991; Veitch et al., 1997; Wagner et al., 2008; Merritt et al., 2005; Merritt et al, 2010). Still others have found that direct contact with aquatic environments is a major risk factor (Aiga et al., 2004; Raghunathan et al., 2005; Debacker et al., 2006; Sopoh et al., 2010); others suggest aquatic environments. *Mycobacterium ulcerans* has been detected in fish (Eddyani et al., 2004), aquatic snails (Marsollier et al., 2004), and aquatic insects (Portaels et al., 1999). Aquatic insects from the genera *Naucoris* and *Diplonychus*, tested by Portaels et al. (1999), are known to bite humans. Benbow et al. (2008), discount predatory aquatic insects as a potential vector for BU, however, based on the similarities in abundance and *M. ulcerans* positivity of these insects in BU-endemic and non-endemic sites. Merritt et al. (2005) proposed that "poor water quality influences biological communities, leading to increased growth and proliferation of *M. ulcerans* in aquatic habitats." While some research, e.g., Fyfe et al. (2010), has demonstrated a link between *M. ulcerans* and strictly terrestrial animals, the majority of *M. ulcerans* research to date implicates aquatic environments in the potential mode(s) of transmission for BU. Duker et al. (2004) suggested a connection between arsenic enrichment in soil and water and incidence of BU. Taken together, these studies suggest that some difference in environmental conditions, human behavior, or both must exist between endemic and non-endemic communities to explain the difference in BU morbidity.

Williamson et al. (2008) conducted a microbiological assay of aquatic environmental samples in both Buruli-endemic and non-Buruli-endemic areas of Ghana, finding that *M. ulcerans* is present equally often in samples from endemic and non-endemic communities. Vandelannoote et al. (2010) conducted a similar study to that of Williamson et al. (2008) but came to a notably different conclusion: they found *M. ulcerans* in only one out of 148 environmental samples. Vandelannoote et al. (2010) suggests that the higher frequency of *M. ulcerans* detection by Williamson et al. (2008) could represent a) contamination of samples prior to or during laboratory analysis or b) a real difference due to inconsistent sampling sites between the two studies.

Here, I present a study of water chemistry throughout endemic and non-endemic communities in an effort to characterize the chemical signature of potentially hazardous areas with respect to BU. I explore the relationships between land use and water quality in several Buruli-endemic and non-endemic areas of Ghana in an effort to determine the characteristics of a favorable environmental niche for the persistence of *M. ulcerans* in the environment and for its transmission to human hosts. This study focuses on mining regions and particularly on “galamsey” (gather-and-sell), or artisanal small-scale, gold mining areas. These areas are characterized by large amounts of localized disturbance, most notably pools of water associated with active or defunct ore-washing stations. Galamsey operations, which are poorly regulated and often unregistered and/or illegal, have no obligation to remediate spent mining areas (Figure 1), leading to long-term environmental degradation.

Figure 1. Abandoned galamsey (artisanal gold-mining) pit, Pokukrom, Central region, Ghana.

A study of mycobacteria in brook waters, conducted by Iivanainen et al. (1993), found that culturable counts of slow-growing mycobacteria were most negatively correlated to pH. Likewise, counts were most positively correlated to chemical oxygen demand and metals concentrations. Consequently, one might expect that *M. ulcerans,* a slow-growing mycobacterium, would be likely to thrive in similar environments to those bacteria studied by Iivanainen et al. (1993), if indeed *M. ulcerans* can exist outside a host. Using this assumption, *M. ulcerans* should be most prevalent in water with low pH, and that *M. ulcerans* may be seen most commonly in waters with high metals concentrations, particularly iron and heavy metals. This assumption is supported by many studies of BU incidence relative to land use, as well as known chemical trends associated with these land uses (e.g., increased nitrogen in agricultural areas). Areas with high nitrogen and phosphorus concentrations will likely be a preferred environment for the growth of *M. ulcerans*, as environmental nutrient enrichment has been linked to the emergence of other direct-transmission and vector-borne bacterial diseases (Johnson et al., 2010). The positive correlation of mycobacterial population with metals concentrations suggests that BU incidence may be higher near mining sites, as heavy metals are commonly associated with tailing waste from mining activity (Walker et al., 2001). While *M. ulcerans* is not explicitly measured in this study and the incidence of BU is instead used, detection of preferred environments for *M. ulcerans* growth and persistence may be useful in the quest for the Buruli ulcer vector(s).

In this work, I use multi-variate statistical methods to determine whether surface water chemistry is related to water body type and surrounding land use. If a strong correlation is seen, then these factors may drive water quality. These statistical methods are also used to determine the chemical characteristics of potentially dangerous environments with respect to Buruli ulcer. If statistical analyses show that water samples from endemic communities have high metal concentrations and low pH relative to non-endemic communities, then this work will serve as a first step to linking Buruli ulcer incidence to poor water quality. More broadly, this and future work toward identifying the environmental niche for *M. ulcerans* will serve to predict hazardous water bodies and thus prevent incidence of BU in at-risk communities.

Description of Field Sites

This study of water chemistry was conducted in five distinct communities in Ghana, West Africa (Pokukrom, Betenase, Kedadwen, Ayanfuri, and Nangruma; Figure 2), each of which is associated with a larger study area for other related work (GIS, social science, land use studies, etc.). These study areas comprise three endemic areas (Pokukrom, Betenase, and Ayanfuri; Study Areas 1, 2, and 4) and two non-endemic areas (Kedadwen and Nangruma; Study Areas 3 and 5, respectively). These areas were selected based on data from Ghana's National BU Control Programme (2008).

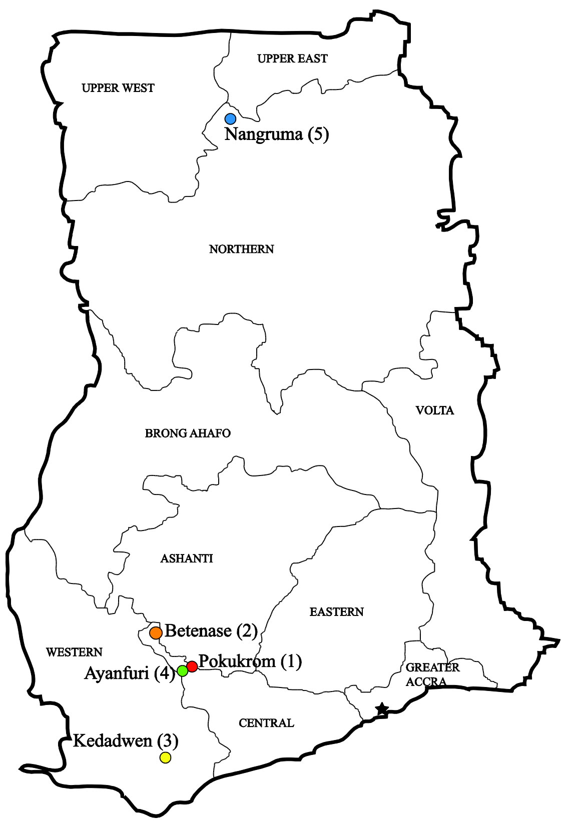


Figure 2. Map of study areas.

Most of these study areas are composed of more than one study community. The communities included in each study area are summarized in Table 1. Surveys of human behavior were conducted in every study community, with several participants from each community (Tschakert et al., in prep). Water samples were collected on a community scale, considering one community in each of the five study areas. These communities were selected for water sampling based on their status as either the only community or as the community with the highest number of recent Buruli ulcer cases within the respective study area. The number of recent Buruli ulcer cases was determined from visits to all communities in 2010.

Table 1. Summary of Study Areas. Bold communities are those selected for water chemistry study.

Study Area	Communities	Region	District
1	Obiaradaneden	Central	Upper Denkyira East
	Pokukrom		
	Powerline		
2	Ameyaw	Central	Upper Denkyira West
	Ampabena		
	Betenase		
	Subin		
3	**Kedadwen**	Western	Tarkwa-Nsuaem Municipal
4	**Ayanfuri**	Central	Upper Denkyira East
	Nkotumso		
5	**Nangruma**	Northern	West Mamprusi

GPS coordinates of all sampling sites in each community are shown in Appendix A.

Geology

All of these study communities are located in the Birimian series (Multilateral Investment Guarantee Agency, 2000), which is composed largely of volcanic rocks with gold-bearing quartz veins (Dzigbodi-Adjimah, 1993). These quartz veins contain carbonate minerals as well as metallic sulfides and arsenides. Gold-bearing quartz veins in the Birimian series were found to have high concentrations of arsenic, cadmium, copper, iron, lead, and zinc, while non-gold-bearing quartz veins showed low concentrations of these trace metals (Dzigbodi-Adjimah, 1993).

Climate

Ghana experiences two climatic seasons, rainy and dry. The rainy season occurs between May and October, while the dry season occurs between November and April. In the southern sites, annual rainfall averages 1600 millimeters; by contrast, Study Area 5 (Nangruma), in the northern region of Ghana, receives just 990 millimeters (GMet, 2010). Geographic differences also exist in the rainy season signature of these two regions: The south experiences a bimodal rainy season (major between May and July and minor between August and October), while the north sees a unimodal rainy season (Figure 3). Corresponding to these differences in annual rainfall, vegetation is different between these regions of Ghana: the south is characterized by thick deciduous forest, while vegetation in the north is much sparser, with shrubs and drought-resistant trees. See Appendix E for analysis of rainfall frequencies between study sites.

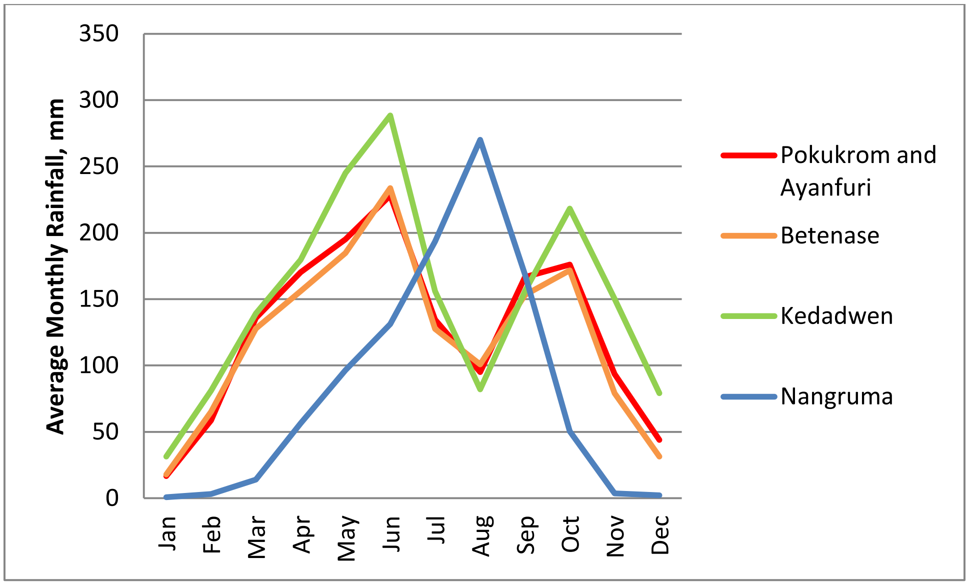


Figure 3. Average monthly rainfall in study areas, 1960-2010. Data from GMet (2010). Average rainfalls for Pokukrom and Ayanfuri are based on a gauging station in Dunkwa, Betenase on Sefwi Bekwai, Kedadwen on Tarkwa, and Nangruma on Navrongo.

Agriculture

In Study Areas 1, 2, and 4 (Pokukrom, Betenase, and Ayanfuri), the major crop is cocoa, grown in plantations. As Ghana is one of the world's leading cocoa producers, the national government subsidizes the crop and the fertilizers required to grow it effectively. Vegetables (tomatoes, eggplant, peppers, plantains, maize, etc.) are often grown in smaller polyculture plots near the cocoa crop, but some, like yams, are planted among the cocoa. In Study Area 3 (Kedadwen), a slightly more humid climate, rubber is the major commodity crop. Many of the same vegetables are farmed in Kedadwen as in the other southern communities, but cassava replaces yam as the major starch crop. Cassava is typically grown in monoculture fields. Study Area 5 (Nangruma), with its drier climate, has substantially different agriculture: there, yams, grown in monoculture fields in raised mounds, are the staple crop. Cattle herding is also common in Nangruma.

Mining

Gold mining is prevalent in all of these study communities, but the methods are different between northern and southern Ghana. In Study Areas 1, 2, and 4 (Pokukrom, Betenase, and Ayanfuri), mining is almost exclusively alluvial. Areas near rivers are cleared, and miners dig by hand or with excavators to expose buried river sediment deposits. These river bed sediments are then washed and panned to extract the gold; sediments are sometimes crushed, by hand or mechanically, prior to washing and panning. Men, women, and children alike are involved in the washing and panning processes. The alluvial mining process causes major soil disturbance. Alluvial mining is often associated with the blocking and/or rerouting of rivers to provide a water source for sediment washing or to expose river bed sediments. Left in the wake of illegal alluvial mining is a “moonscape” of water-filled pits surrounded by mounds of mine tailings.

Miners in Study Area 3 (Kedadwen) also operate by surface mining, though the method is different from that of the other southern communities. The side of a hill has been excavated by multiple gangs. These gangs, operating individually, mine a single large tract of land; they dig and crush buried river sediments and weathered rock. Washing and panning practices are consistent with the other southern communities.

By contrast, gold mining in Study Area 5 (Nangruma) is performed by hard-rock mining methods. Small shafts are dug as deep as 100 feet underground. Gold-bearing ore is mined by hand or using dynamite. Ore is carried to the surface and crushed at one of several mechanical crushers in the community. Because of the occupational hazards associated with this work, miners in Nangruma are typically adult males. The community’s river is pumped, in some cases in its entirety, to run the wet crushers and wash the crushed ore. Spent water then flows back to the river with an increased sediment load, and mine tailings are piled next to the crushers. While the landscape in Nangruma is dotted with mine shafts, the disturbance of soils is much less than in the south.

Field Methods

In each study community, water samples were collected from wells and boreholes, rivers and streams, galamsey mining pits, swamps, and potential "Buruli ulcer hot spots," where applicable, in the selected community. "Buruli ulcer hot spots" were identified by community members during participatory mapping exercises conducted in June 2010. Community members indicated areas that they felt posed a risk for contracting Buruli ulcer; these areas were consistently pools of stagnant water. Samples of rivers and streams were taken at popular crossing points, as these are contact points common to many community members. The samples presented here were collected in June and July of 2011. Protocols for sampling are outlined in Appendix D.

Each water sample was collected in three bottles:

- 500mL unpreserved sample for analysis of pH, major cations, major anions, and sulfate,
- 500mL preserved with H_2SO_4 for analysis of ammonia, nitrate, nitrite, and phosphate, and
- 100mL preserved with HNO_3 for analysis of trace metals.

Wells and boreholes were in active production when sampled. Sample bottles were filled directly from the water body when possible; inaccessible or hazardous sites were sampled using a bucket. Highly concentrated H_2SO_4 and HNO_3 were added, as necessary, to samples immediately upon collection to achieve a pH of 2. All samples were stored in insulated containers until they could be transported to the laboratory at Kwame Nkrumah University of Science and Technology (KNUST). Samples were refrigerated upon arrival at the laboratory. Analysis for major ions, nitrogen, and phosphorus occurred at KNUST, while trace metal samples were sent to SGS Laboratory Services, Ghana Ltd. for analysis by ICP-OES for cadmium, copper, iron, lead, and zinc and by gaseous hydride atomic absorption (APHA 3114B) for arsenic and selenium.

Statistical Methods

A series of univariate and multivariate analyses were used for analysis. Mathematical processes for these methods are described below.

Analysis of Variance

One-way analysis of variance (ANOVA) is a measure of the contribution of an independent factor to the variance of a single variable. A one-way ANOVA is calculated by the equation:

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij}-\bar{y})^2=\sum_{i=1}^{k}n_i(\bar{y_i}-\bar{y})^2+\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij}-\bar{y_i})^2 \quad (1)$$

where k is the number of unique groups in a factor, n_i is the number of samples in each group, y_{ij} is a sample, $\bar{y}$ is the overall mean of the variable, and $\bar{y_i}$ is the mean of samples within a group (NIST, 2012). This process separates the total sum of squares into the sums of squares between groups and for error within groups, respectively, such that the above equation simplifies to $SS_{total}=SS_{factor}+SS_{error}$. The sums of squares are then converted to mean squares; the mean square of the factor, or variability between groups, is calculated by $MS_{factor}=SS_{factor}/DF_{factor}$, where MS_{factor} is the mean square of the factor and DF_{factor} is the degrees of freedom in the factor, equal to k-1. The mean square of error, or variability within groups, is similarly calculated by $MS_{error}=SS_{error}/DF_{error}$, where DF_{error} is equal to N-k and N is the total number of samples. An F test statistic is then calculated by $F=MS_{factor}/MS_{error}$. This value is then compared to an F distribution table to determine statistical significance based on degrees of freedom.

Cluster Analysis

Cluster analysis is a preliminary analytical tool that can be used to illustrate similarity or dissimilarity among samples (Q-mode) or variables (R-mode) in a data set. A sample is linked to its most similar sample; this resulting group is then linked to its most similar sample or group. This hierarchical approach continues until all samples are linked to each other. The same approach applies for comparison of variables.

Cluster analysis is performed using a series of steps (Legendre and Legendre, 1998). First, the original data set, **A**, is standardized to have a mean of 0 and variance of 1, forming a matrix, **B**. This standardization occurs by the formula:

$$b_{ij} = \frac{a_{ij} - \bar{a}_j}{SD_j} \qquad (2)$$

where b_{ij} is the standardized value for sample i and variable j, a_{ij} is the original value for sample i and variable j, $\bar{a}_j$ is the mean value for variable j, and SD_j is the standard deviation of variable j. Pearson correlation coefficients, where 0 indicates no correlation and 1 and -1 indicate perfect positive or negative correlation, respectively, are then calculated from this standardized matrix. Pearson correlation coefficients are calculated by the equation:

$$r_{jk} = \frac{\sum_{i=1}^{n}\left(b_{ij} - \overline{b_j}\right)\left(b_{ik} - \overline{b_k}\right)}{(n-1)SD_j SD_k} \qquad (3)$$

where $\overline{b_j}$ and $\overline{b_k}$ are the mean values for columns j and k, n is the number of samples, and SD_j and SD_k are the standard deviations for columns j and k. These correlation coefficients are used as the similarity metric for cluster analysis. Each of these coefficients represents the degree of similarity between two columns. As such, calculations using matrix **B** will yield correlations between variables, while calculations with the transpose of matrix **B (**matrix **B')** will yield correlations between samples.

The basic combinatorial equation used to construct clusters from similarity data is:

$$S(\boldsymbol{hi},\boldsymbol{g}) = (1 - \alpha_h - \alpha_i - \beta) + \alpha_h S(\boldsymbol{h},\boldsymbol{g}) + \alpha_i S(\boldsymbol{i},\boldsymbol{g}) + \beta S(\boldsymbol{h},\boldsymbol{i}) - \gamma|S(\boldsymbol{h},\boldsymbol{g}) - S(\boldsymbol{i},\boldsymbol{g})| \quad (4)$$

where S represents a similarity measure, g, h, and i are clusters, hi is the cluster formed by clusters h and i, and α_h, α_i, β, and γ are combinatorial coefficients that control the order and conditions under which elements are combined into clusters. Clusters in this analysis were determined using Ward's method, which operates by minimizing increases in the error sum of squares as elements are combined. Ward's method was selected for this analysis because it is the most suitable method for clustering data using correlation distance (McCune and Grace, 2002). Ward's is also one of the preferred methods for analyzing standardized data. In Ward's method, $\alpha_h = \frac{n_h+n_g}{n_h+n_i+n_g}$, $\alpha_i = \frac{n_i+n_g}{n_h+n_i+n_g}$, $\beta = \frac{-n_g}{n_h+n_i+n_g}$, and $\gamma = 0$, where n_h represents the number of elements included in cluster h, n_i the number of elements included in cluster i, and n_g the number of elements in cluster g (Legendre and Legendre, 1998). The combinatorial equation is used repeatedly until all elements are included in the clusters.

Principal Component Analysis

Principal component analysis (PCA) reduces the number of variables in a data set to a much smaller number of synthetic variables (principal components) that represent most of the variance among samples. PCA calculates lines of best fit through multidimensional space, where each dimension represents a different variable. Lines of best fit are sequentially drawn perpendicular to each other until variance is completely explained. The number of best-fit lines drawn is typically less than the number of variables in the data set. Geochemical data lend themselves nicely to PCA due to linear relationships among variables and the low number of zeros in comparison to ecological data.

Principal components analysis is performed using a series of steps (McCune and Grace, 2002). First, the original data set, **A**, is standardized to have a mean of 0 and variance of 1, forming a standardized matrix, **B**, as was done with cluster analysis (Equation 1). This standardized matrix, **B**, is transformed to a correlation matrix, **S**, by $\boldsymbol{S} = \boldsymbol{B}'\boldsymbol{B}$ where **B'** is the transpose of the standardized matrix. This correlation matrix is then used to determine eigenvalues, each of which represents part of the variance in the data. Simply put, eigenvalues describe the variance explained by each principal component axis. These eigenvalues are calculated by solving the characteristic equation:

$$|\boldsymbol{S} - \lambda \boldsymbol{I}| = 0 \qquad (5)$$

where λ is an eigenvalue that solves the characteristic equation and **I** is the identity matrix. Expanding the characteristic equation produces a polynomial

$$a\lambda^p + b\lambda^{p-1} + c\lambda^{p-2} + \cdots + constant = 0 \qquad (6)$$

where p is the number of variables in the original data set; a, b, c, etc. are scalar coefficients, and λ represents all possible roots of the polynomial. After calculating eigenvalues, one can calculate eigenvectors, $\vec{y}$, such that each element of an eigenvector represents the contribution (loading) of a given variable to the principal component axis. As an eigenvector represents the contribution of each variable, the length of $\vec{y}$ should be equal to p. Eigenvectors should be normalized so that the sum of squares for each vector is equal to 1. This normalization occurs by applying a scaling factor, k_i, to each element of the vector. The scaling factor for each vector is calculated by

$$k_i = \frac{1}{\sqrt{\sum_{q=1}^{p} y_{qi}^2}} \quad (7)$$

where y_{qi} is the q^{th} element of eigenvector i. These normalized eigenvectors can be collated into a p-by-p matrix, **Y**, known as the loadings matrix. This normalized loadings matrix is used to calculate a matrix of scores, **X**, by **X**=**AY**. These scores represent the coordinates of data points in a rotated coordinate system.

Results and Discussion

Analyses presented here were performed using data shown in tabular format in Appendix B. All samples were collected during June and July 2011.

Univariate Analysis

All water bodies tend to be slightly acidic ($pH < 7$) (Figure 4a). Of the surface water bodies, BU hot spots and swamps are most consistently acidic, but wells are by far the most acidic waters ($p=2.207E-6$). This strong acidic signal in shallow groundwater is common to groundwater in Ghana and is likely related to the high acidity of soils in the region. Based on these results, pH may be an indicator of the ability of *M. ulcerans* to grow in certain water bodies. Endemic communities (Study Areas 1, 2, and 4: Pokukrom, Betenase, and Ayanfuri) are consistently acidic, but Study Area 3 (Kedadwen), a non-endemic community, displays similar pH values to these endemic communities (Figure 4b). Water samples from Study Area 5 (Nangruma) are typically neutral or slightly basic, but the number of samples from Nangruma for which pH data exist is very small ($n=3$). These results suggest more suitable conditions for *M. ulcerans* growth in southern Ghana than in northern Ghana, but the overlap of observed pH values between endemic and non-endemic communities in the south implies that pH is not a controlling factor for *M. ulcerans* growth in this region.

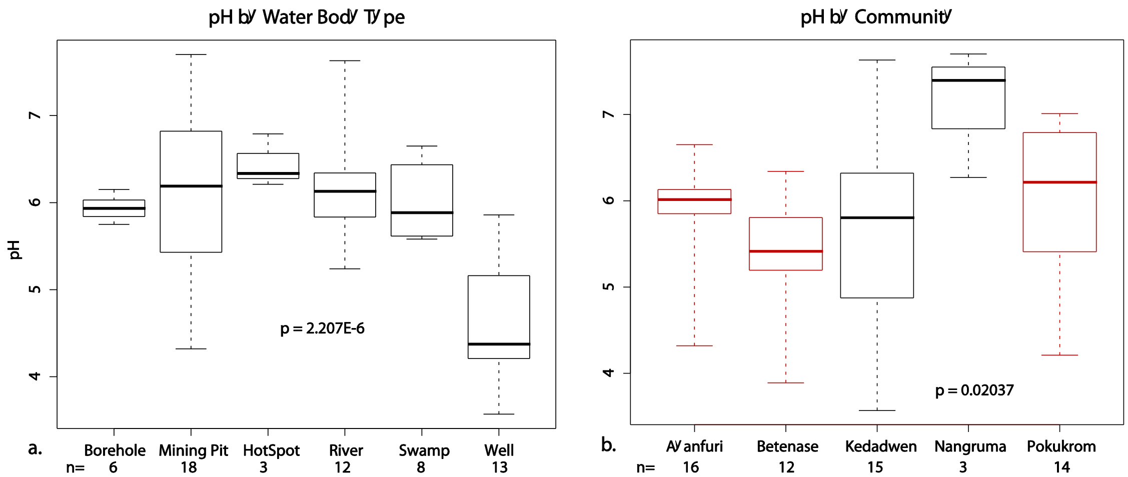


Figure 4. Box-and-whisker plots of pH organized by a) water body type, and b) community. Red indicates an endemic community. n values show the number of samples in each group. P-values represent the statistical significance of these factors from one-way ANOVA.

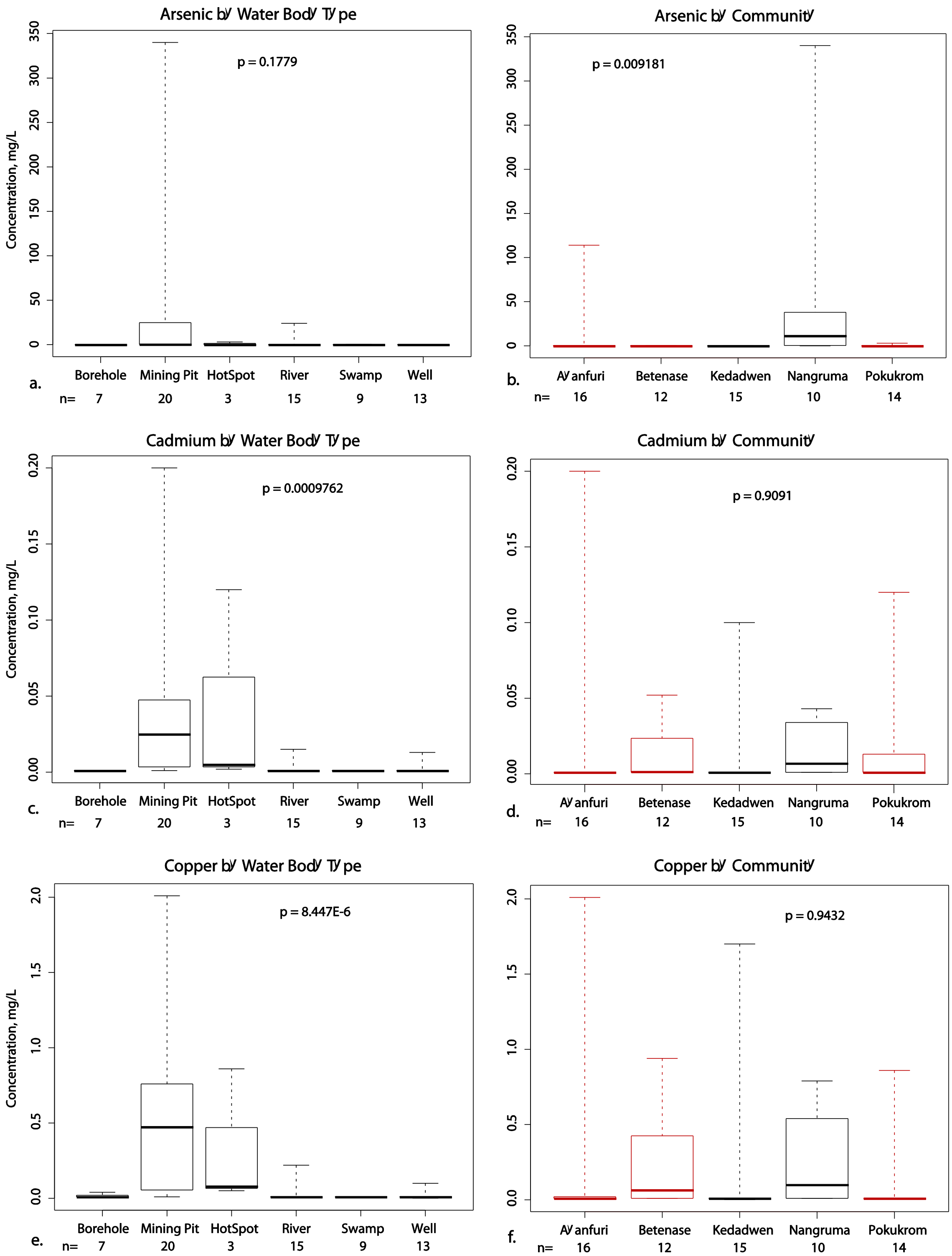
Elevated trace metal concentrations, postulated to be beneficial for *M. ulcerans* growth, are often associated with mining. As arsenopyrite is often associated with gold deposits, it is not surprising that high arsenic concentrations are seen in galamsey pits (Figure 5a). Arsenic concentrations are also high in BU hotspots and rivers relative to other water sources, though the differences between water bodies are not statistically significant (p=0.1779). Rivers are a common drinking water source, and arsenic in drinking water is known to have immunosuppressive properties (Banerjee et al., 2009). Arsenic could thus serve as a double threat for BU incidence: arsenic in water could support the growth of *M. ulcerans*, while arsenic in drinking water could suppress immune systems, making the population more susceptible to BU. Endemic communities (Study Areas 1, 2, and 4: Pokukrom, Betenase, and Ayanfuri) show some high arsenic concentrations, but the highest arsenic concentrations are seen in Study Area 5 (Nangruma) (Figure 5b). Within Nangruma, arsenic concentrations are highest in galamsey pits. This trend is likely caused by the differences in mining practices between these communities. In the southern communities (Study Areas 1, 2, 3, and 4: Pokukrom, Betenase, Kedadwen, and Ayanfuri), gold mining is largely alluvial: miners dig up and sift soil and weathered rock from river beds. Conversely, in Study Area 5 (Nangruma), the main mining practice is hard-rock: miners use dynamite to blast gold-bearing rock as much as 100 feet underground. When this rock is brought to the surface and crushed, arsenic may be released from veins of arsenopyrite within the ore in larger quantities than it would be washed from river sediments. This arsenic may wash from galamsey drainage pits to rivers and swamps, causing a general trend of elevated arsenic concentrations in Nangruma.

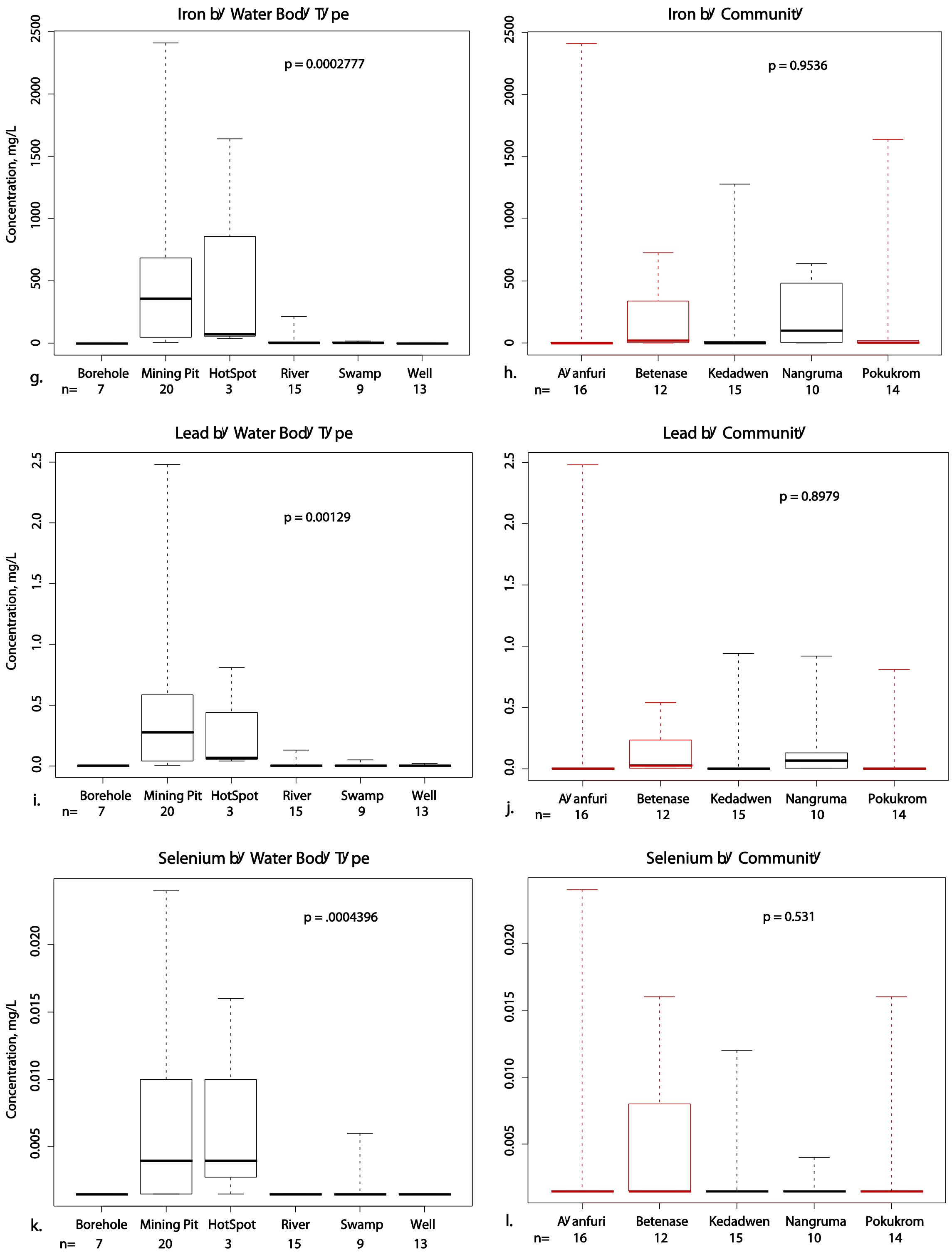
Cadmium concentrations are highest in galamsey pits and BU hot spots ($p=0.0009762$) (Figure 5c), so galamsey pits and stagnant water bodies may be risky locations for contracting BU. When organized by community, Study Areas 1, 2, and 5 (Pokukrom, Betenase, and Nangruma) show highest cadmium concentrations (Figure 5d), but the differences between communities are not statistically significant ($p=0.9091$). As such, cadmium concentrations do not relate to study area. Copper concentrations show similar trends to cadmium, in that galamsey pits and BU hot spots have consistently higher concentrations than other water bodies ($p=8.447E-6$) (Figure 5e). Trends for copper and cadmium are very different when organized by community (Figure 5f). Consistently high concentrations are seen in Betenase (Study Area 2), but concentrations are typically low in the other endemic communities and not significantly different from each other ($p=0.9432$). Nangruma (Study Area 5) also shows high copper concentrations. Iron exhibits much the same pattern as cadmium (Figure 5g and 5h). Concentrations are highest in galamsey pits and BU hot spots ($p=0.0002777$). Likewise, Betenase and

Nangruma have higher, though not statistically significant, iron concentrations than the other study communities. Lead follows similar trends to copper and iron (Figure 5i and 5j). Concentrations are highest in galamsey pits and BU hot spots, and somewhat elevated in Betenase and Nangruma over other communities. Lead, as a trace metal, may contribute to the favorable environment for *M. ulcerans* growth; however, there is no distinct difference between lead concentrations in endemic and non-endemic communities (p=0.8979).

Selenium concentrations are highest in galamsey pits and BU hot spots, but there are also elevated concentrations in some swamps (Figure 5k). Organized by community, Betenase (Study Area 2) has consistently highest concentrations of selenium (Figure 5l). Pokukrom, Betenase, and Ayanfuri (Study Areas 1, 2, and 4), the endemic communities, have the highest maximum selenium concentrations, though this distinction is not statistically significant (p=0.531).

Zinc shows patterns very similar to other trace metals considered in this study: concentrations are highest in galamsey pits and BU hot spots (Figure 5m). Again, concentrations are elevated in Study Areas 2 and 5 (Betenase and Nangruma) relative to other communities (Figure 5n).


Arsenic by Water Body Type
p = 0.1779
Concentration, mg/L
0 50 100 150 200 250 300 350
a.
Borehole Mining Pit HotSpot River Swamp Well
n= 7 20 3 15 9 13
Arsenic by Community
p = 0.009181
0 50 100 150 200 250 300 350
b.
Ayanfuri Betenase Kedadwen Nangruma Pokukrom
n= 16 12 15 10 14
Cadmium by Water Body Type
p = 0.0009762
Concentration, mg/L
0.00 0.05 0.10 0.15 0.20
c.
Borehole Mining Pit HotSpot River Swamp Well
n= 7 20 3 15 9 13
Cadmium by Community
p = 0.9091
0.00 0.05 0.10 0.15 0.20
d.
Ayanfuri Betenase Kedadwen Nangruma Pokukrom
n= 16 12 15 10 14
Copper by Water Body Type
p = 8.447E-6
Concentration, mg/L
0.0 0.5 1.0 1.5 2.0
e.
Borehole Mining Pit HotSpot River Swamp Well
n= 7 20 3 15 9 13
Copper by Community
p = 0.9432
0.0 0.5 1.0 1.5 2.0
f.
Ayanfuri Betenase Kedadwen Nangruma Pokukrom
n= 16 12 15 10 14


Iron b/ Water Bod/ T/ pe
p = 0.0002777
Concentration, mg/L
0
500
1000
1500
2000
2500
g.
Borehole Mining Pit HotSpot River Swamp Well
n= 7 20 3 15 9 13
Iron b/ Communit/
p = 0.9536
h.
A/ anfuri Betenase Kedadwen Nangruma Pokukrom
n= 16 12 15 10 14
Lead b/ Water Bod/ T/ pe
p = 0.00129
0.0
0.5
1.0
1.5
2.0
2.5
i.
Borehole Mining Pit HotSpot River Swamp Well
n= 7 20 3 15 9 13
Lead b/ Communit/
p = 0.8979
j.
A/ anfuri Betenase Kedadwen Nangruma Pokukrom
n= 16 12 15 10 14
Selenium b/ Water Bod/ T/ pe
p = .0004396
0.005
0.010
0.015
0.020
k.
Borehole Mining Pit HotSpot River Swamp Well
n= 7 20 3 15 9 13
Selenium b/ Communit/
p = 0.531
l.
A/ anfuri Betenase Kedadwen Nangruma Pokukrom
n= 16 12 15 10 14

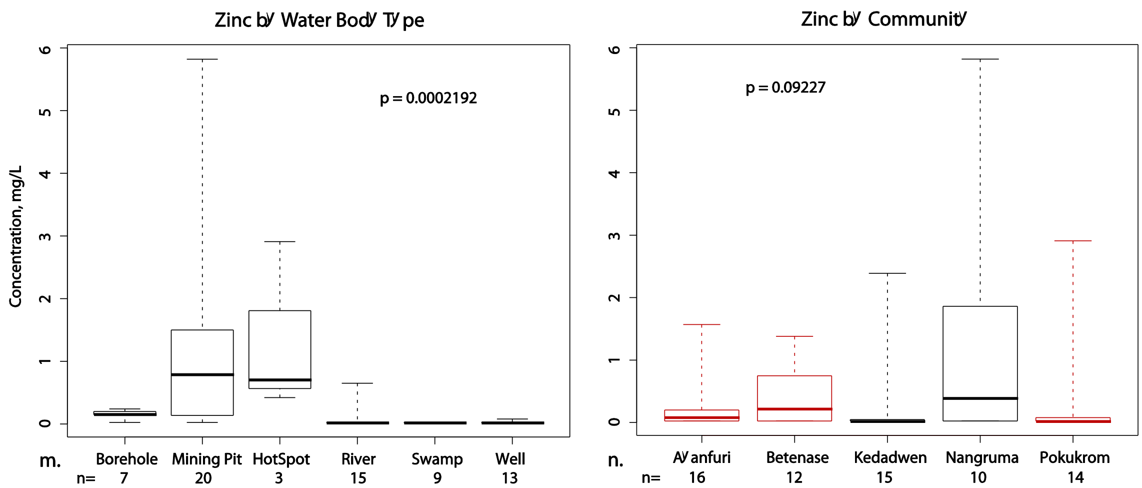
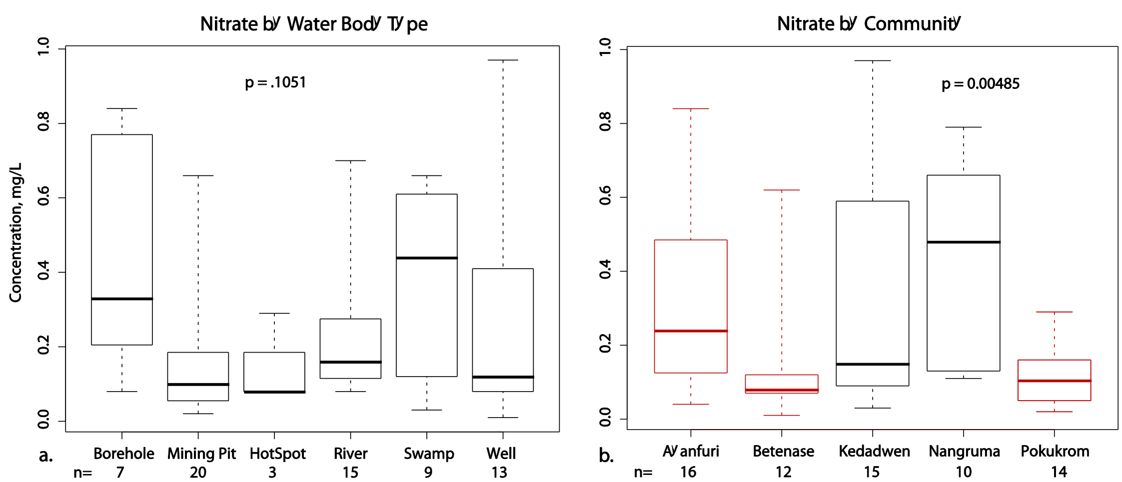


Figure 5. Box-and-whisker plots of a) arsenic organized by water body type, b) arsenic organized by community, c) cadmium organized by water body type, d) cadmium organized by community, e) copper organized by water body type, f) copper organized by community, g) iron organized by water body type, h) iron organized by community, i) lead organized by water body type, j) lead organized by community, k) selenium organized by water body type, l) selenium organized by community, m) zinc organized by water body type, and n) zinc organized by community. Red indicates endemic communities. n values display the number of samples in each group. P-values represent the statistical significance of these factors from one-way ANOVA.

M. ulcerans is also expected to thrive in environments with high nitrogen and phosphorus concentrations. Nitrate concentrations are highest in boreholes, swamps, and wells (Figure 6a). Elevated nitrate concentrations in boreholes are unexpected, as nitrate is typically sourced from the ground surface; the boreholes in this study are presumed to be drilled into deep groundwater and are installed with a pump and intact cement pad (Figure 7a). Conversely, the wells in this study are hand-dug and range in depth between 8 and 20 feet below ground surface. These wells are typically uncased, though some wells have cement pads (Figure 7b). A shallow water table could contribute to high nitrate concentrations as nitrate migrates into and through the subsurface from nearby agricultural fields (Chen et al., 2005). Considering nitrate concentrations by community, there is no significant difference between endemic and non-endemic communities (Figure 6b). Nitrate in samples from Nangruma is unexpected, as sampling in this community occurred prior to the growing season and thus prior to fertilizer application. Additionally, density of agricultural lands is much greater in the south; fertilizer use and corresponding nitrogen concentrations should also be greater in the southern communities. While nitrate generally supports growth of bacteria, it may not necessarily support the preferable growth of *M. ulcerans* over other bacteria.

Boreholes, representing deep groundwater, have the lowest phosphate concentrations, as phosphate is predominantly sourced from fertilizers (Figure 6c). Ranges of phosphate concentration are similar between endemic and non-endemic communities (Figure 6d). However, median concentrations are higher in endemic communities than in non-endemic communities. This may be notable, as phosphate could contribute to the growth of *M. ulcerans* in the environment.



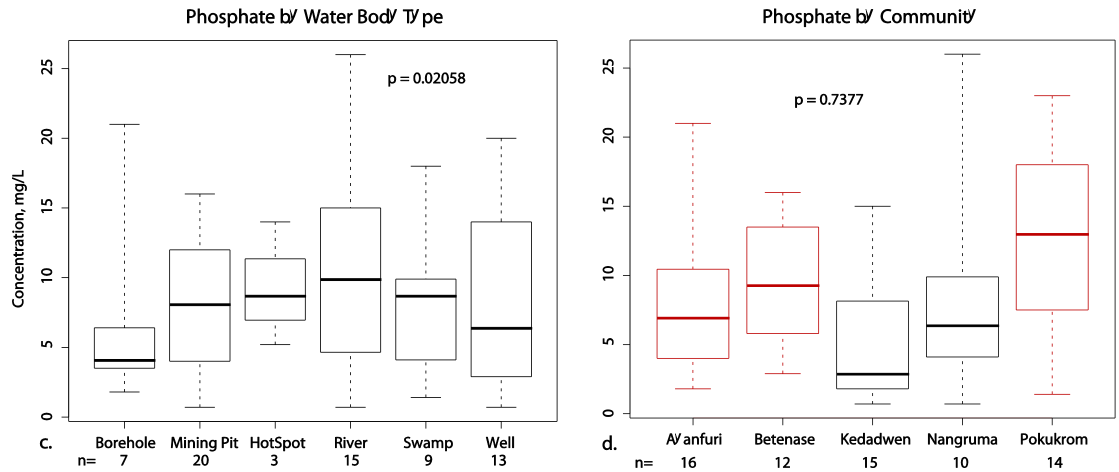


Figure 6. Box-and-whisker plots of a) nitrate organized by water body type, b) nitrate organized by community, c) phosphate organized by water body type, and d) phosphate organized by community. Red indicates endemic communities. Values of n display the number of samples in each group. P-values represent the statistical significance of these factors from one-way ANOVA.

a.

b. 

Figure 7. Groundwater sources. a) borehole, and b) hand-dug well.

Multivariate Analysis

An R-mode cluster analysis of this dataset reveals close groupings of certain chemical constituents (Figure 8). Height in this figure represents a measure of dissimilarity; two elements that combine at a low height are more similar to each other than those that combine at a greater height. Variables separate into two distinct groups: chloride, hardness, nitrate, nitrite, phosphate, ammonium, manganese, and fluoride represent major ions and agriculturally associated compounds, while alkalinity, sulfate, zinc, arsenic, cadmium, iron, lead, copper, and selenium represent the chemical components of gold-bearing rock in Ghana (Dzigbodi-Adjimah, 1993). Gold is most commonly found in carbonaceous fillings in quartz. Included in these carbonaceous fillings are sulfides of trace metals.

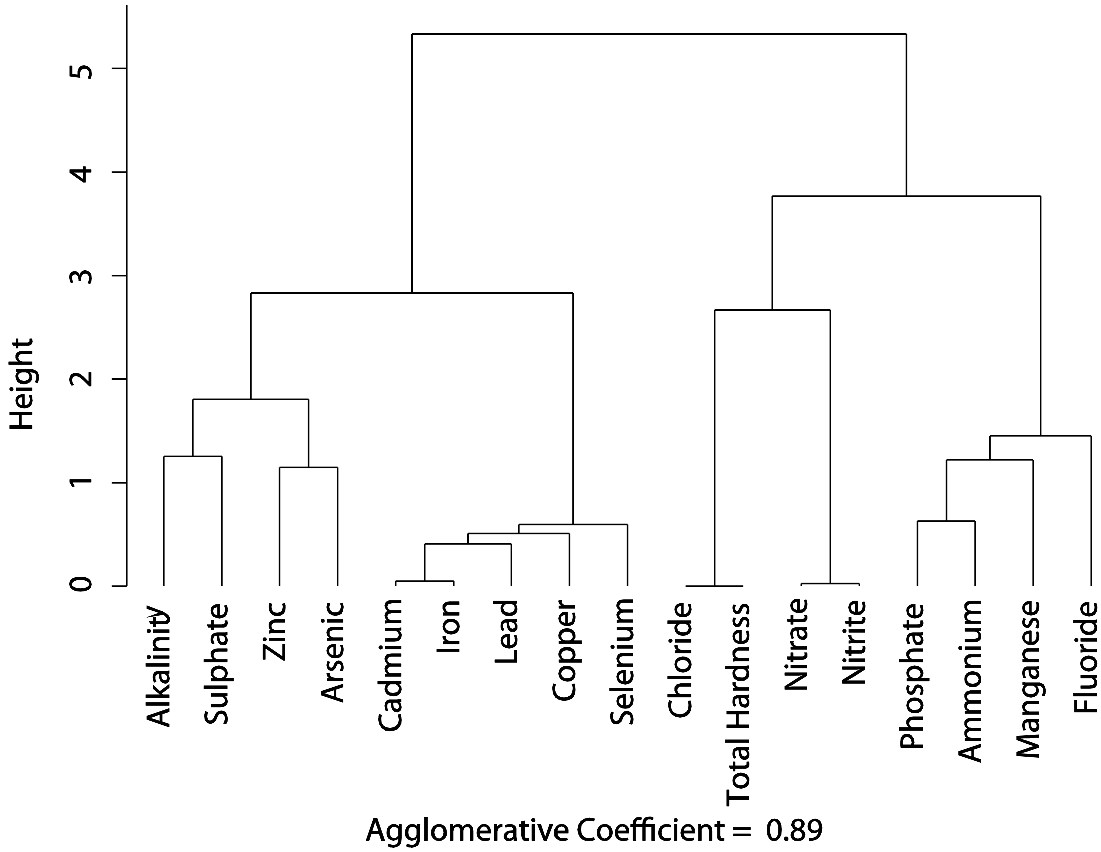


Figure 8. R-mode cluster analysis.

PCA of data from June-July 2011 shows 41.3 percent of variance explained by principal components 1 and 2. Scores and loadings for this analysis are presented for reference in Appendix C. Principal component 1 is dominated by trace metals (cadmium, copper, iron, lead, zinc, and selenium), while principal component 2 is primarily controlled by phosphate, ammonium, and fluoride. Principal component 1 explains 28.1 percent of the total variance in the data; principal component 2 explains an additional 13.2 percent. Principal component 3 explains an additional 12.7 percent of total variance and is controlled by chloride, total hardness, nitrite, and nitrate.

Figure 9a shows a plot of principal components 1 and 2, showing several interesting patterns. Galamsey pits plot distinctly from other water bodies, with the division driven by their higher concentrations of trace metals relative to rivers, swamps, and groundwater. BU hot spots, those areas identified by community members to be "risky," plotted more similarly with galamsey pits than with other water bodies. As high trace metal concentrations are thought to provide a favorable environment for *M. ulcerans* growth and persistence, and galamsey pits and stagnant water have significantly higher concentrations of these chemicals, it is likely that *M. ulcerans* could thrive in galamsey pits and pools of stagnant water. Phosphate concentration does not correlate to water body type, though high phosphate concentration could contribute to increased *M. ulcerans* growth. Despite differences in field and laboratory methods between sampling campaigns in June-July 2011 and January 2011 (see Appendix F, G), galamsey pits and stagnant water bodies consistently plot distinctly from other water bodies. In January 2011, however, arsenic, iron and alkalinity drove these differences.

In Figure 9b, there is little difference in chemical signature between study communities, an observation that is also true of surface water samples from January 2011 (Appendix F, G) and of groundwater samples from Ghana's Community Water and Sanitation Agency (Appendix H). The one exception to this statement is Nangruma (Study Area 5), shown in blue. This difference could be an artifact of relatively few data points relative to other communities or could represent an actual difference in water chemistry. Given the geographic distance between Nangruma and the other communities, it is likely that the variation seen here is a real difference in chemical signature.

Given that principal components 2 and 3 explain similar amounts of the total variance between samples, a consideration of principal component 3 is warranted. Figure

9c shows a plot of principal components 1 and 3, coded by water body type. Principal component 1 drives the differences between water bodies, while principal component 3 seems to have no effect on these differences. As such, nitrite, nitrate, chloride, and hardness are not major factors in water body chemistry differences. Nitrate and nitrite concentrations could control *M. ulcerans* growth, but these constituents are not significantly correlated to differences in water body type. As shown in Figure 9d, even less variation is detected between communities.

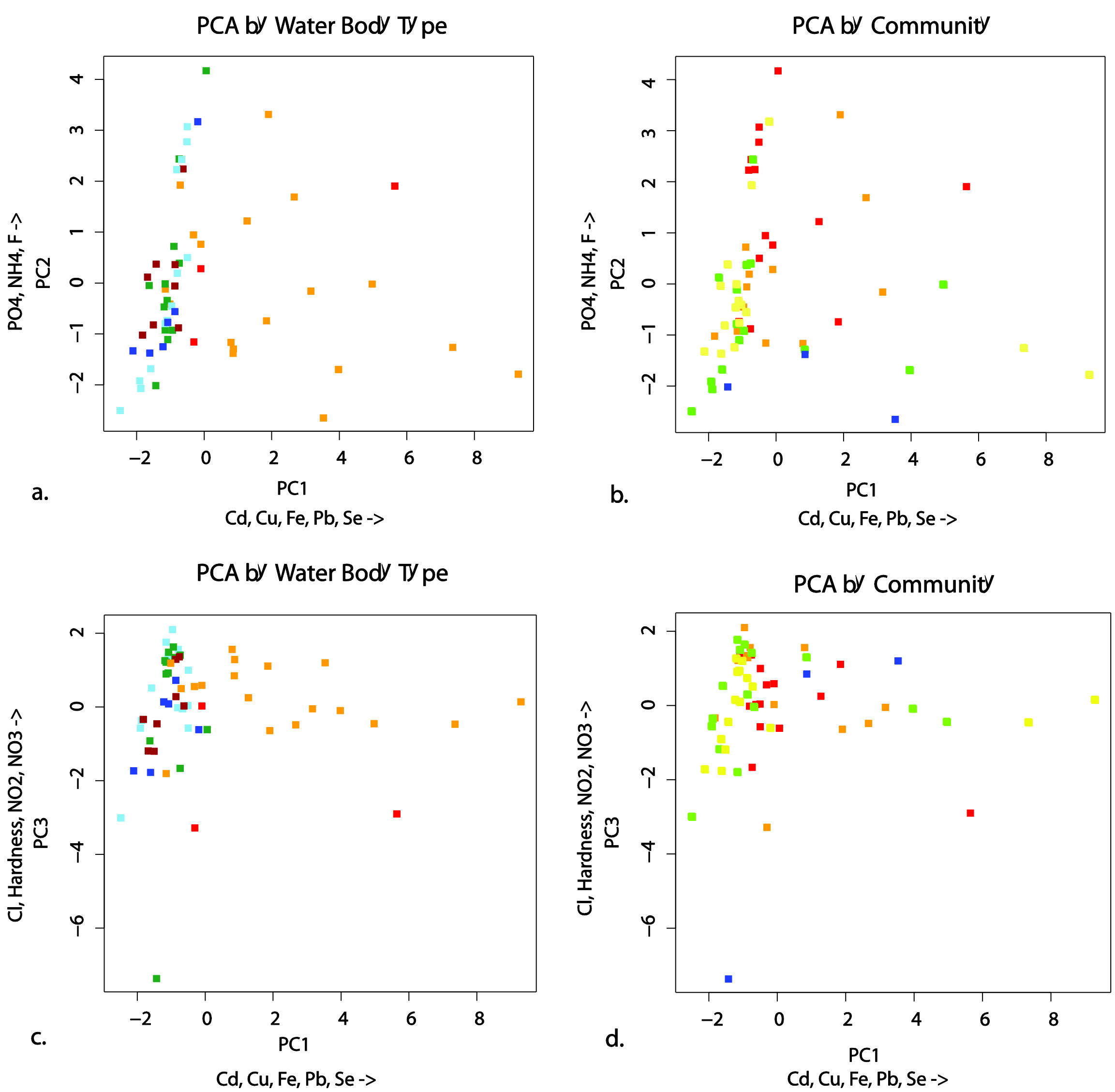


Figure 9. PCA of June-July 2011 data, with a) principal components 1 and 2, coded by type of water body, b) principal components 1 and 2, coded by community, c) principal components 1 and 3, coded by type of water body, and d) principal components 1 and 3, coded by community. In plots coded by type of water body, dark blue represents boreholes, light blue wells, green rivers and streams, brown swamps, orange galamsey pits, and red BU hot spots. In plots coded by community red represents Pokukrom, orange Betenase, yellow Kedadwen, green Ayanfuri, and blue Nangruma. Principal Component 1 is dominated by Cd, Cu, Fe, Pb, and Se; Principal Component 2 is dominated by PO_4, NH_4, and F; Principal Component 3 is dominated by Cl, hardness, NO_2, and NO_3.

Data Complications

Error in these data cannot be discounted due to several issues in sample collection and analysis, largely unavoidable given the difficulty of sampling in this rural field area without significant infrastructure. Best practices for sampling, based on EPA guidelines, suggest that water samples for major ions and sulfate should be kept at 4°C and analyzed within 28 days of collection. Similarly, samples for ammonia, nitrate, nitrite, and phosphorus should be acidified with highly concentrated sulfuric acid, kept at 4°C, and analyzed within 28 days of collection. Samples for trace metals are less restrictive, requiring acidification with nitric acid, but allowing six months for analysis. While in the field, samples were kept in insulated containers but were not stored at 4°C due to lack of access to ice or refrigeration. Because of the remote locations of our study communities, samples were held for up to two weeks before arrival at the laboratory in Kumasi. Such difficulties are common for sampling campaigns in developing countries (Silliman et al., 2007). Upon arrival at the lab, samples were refrigerated and held for as long as 25 days before analysis for alkalinity, total hardness, sulfate, sulfide, ammonium, nitrate, nitrite, phosphate, fluoride, and manganese. Samples for trace metals were analyzed five to six months after collection.

Two types of quality control standards were employed during sampling in an effort to quantify error. Due to a lack of reliable deionized (DI) water in the field, sample bottles were filled with DI water at a laboratory in the United States and were transported to Ghana alongside clean, empty sample bottles. These trip blanks represent a proxy for field blanks. Additionally, one field duplicate was taken at each study site.

Laboratory quality control (method blanks and laboratory duplicates) confirms accuracy and precision of trace metal analyses, with errors typically less than three percent. The accuracy and precision of all other analyses, however, is difficult to quantify. Samples were analyzed by titration, with results representing the average of two tests for each sample. Results from individual tests are not easily recoverable. The comparison of results between field duplicate samples is highly variable; some duplicates display identical concentrations, while others vary by an order of magnitude. Trip blank results were very precise with respect to trace metals, but all other analyses reveal substantial measurable concentrations for all constituents, a matter of concern in these analyses. Fortunately, the variation among samples is much larger than the sampling and analytical error, and so the error does not significantly affect the reliability of PCA (Faber et al. 1993).

Conclusions and Future Work

M. ulcerans is assumed to thrive in environments with low pH and high trace metal concentrations. The analyses presented here, however, find no significant differences in pH between water bodies or between communities. Lower pH values in southern Ghana suggest that if pH does affect the viability of *M. ulcerans*, then southern Ghana may be a more favorable growing environment than northern Ghana. Trace metals have consistently higher concentrations in galamsey pits than in any other water body; however, there is no significant difference in any trace metal's concentrations between communities. As such, some water bodies may be more suitable than others for *M. ulcerans* growth, but chemical evidence suggests that *M. ulcerans* could be capable of growth at all sites.

Statistical analyses show that, at these study sites in Ghana, differences in water chemistry are controlled by trace metal concentrations. It can also be seen that this signature is correlated to the type of water body. Galamsey pits and pools of persistent stagnant water (BU hot spots) are characterized by high trace metal concentrations relative to other water bodies. While they do not differ significantly in their nitrate and phosphate concentrations or in pH, in accordance with the assumed preferable environment for *M. ulcerans*, the high trace metal concentrations suggest that these water bodies may harbor and promote the growth of *M. ulcerans* in the environment. The results of this study suggest that much more work is necessary; the sample size for this study is very small. Samples were collected in only five communities, and sampling sites within those communities were limited to a small subset of the total water bodies in each community. It is possible and likely that the scale of data collection was not fine enough to resolve the variability of water body chemistry. The number of samples in this study prevented a significant analysis of different water bodies within each community. Given more time and money, a more spatially and temporally robust data set would yield more significant results. Weekly or monthly measurements of water chemistry would provide invaluable information about the temporal variability of water bodies in these study communities. Measurements of water chemistry in the wake of extreme rainfall events could also be useful; extreme rainfall and associated flooding may increase the number and size of water bodies that could harbor *M. ulcerans*. Consideration of a greater number of endemic and non-endemic communities, or a greater number of water bodies within the existing study communities, would provide a sense of spatial variability of water bodies; this could give a more robust description of surface water chemistry in and around endemic communities. A lack of resources, time, and expertise prevented the collection and analysis of samples for *M. ulcerans*. In future, water samples should be collected for chemistry and *M. ulcerans* in an effort toward identifying the environmental niche of *M. ulcerans*.

References

Aiga, H., Amano, T., Cairncross, S., Domako, J.A., Nanas, O.K., and Coleman, S. (2004). "Assessing water-related risk factors for Buruli ulcer: a case-control study in Ghana." *American Journal of Tropical Medicine and Hygiene*, **71**, 387-392.

Banerjee, N; Banerjee, S., Sen, R., Bandyopadhyay, A., Sarma, N., Majumder, P., Das, J.K., Chatterjee, M., Kabir, S.N., and Giri, A.K. (2009). "Chronic arsenic exposure impairs macrophage functions in the exposed individuals." *Journal of Clinical Immunology*, **29**, 582-594.

Benbow, M.E., Williamson, H., Kimbirauskas, R., McIntosh, M.D., Kolar, R., Quaye, C., Akpabey, F., Boakye, D., Small, P.L.C., and Merritt, R.W. (2008). "Aquatic invertebrates as unlikely vectors of Buruli ulcer disease." *Emerging Infectious Diseases*, **14**, 1247-1254.

Chen, J., Tang, C., Sakura, Y., Yu, J., and Fukushima, Y. (2005). "Nitrate pollution from agriculture in different hydrogeological zones of the regional groundwater flow system in the North China Plain." *Hydrogeology Journal*, **13**, 481-492.

Debacker, M., Portaels, P., Aguiar, J., Steunou, C., Zinsou, C., Meyers, W., and Dramaix, M. (2006). "Risk factors for Buruli ulcer, Benin." *Emerging Infectious Diseases,* **12**, 1325-1331.

Duker, A.A., Carranza, E.J.M., and Hale, M. (2004). "Spatial dependency of Buruli ulcer prevalence on arsenic-enriched domains in Amansie West District, Ghana:

implications for arsenic mediation in *Mycobacterium ulcerans* infection." *International Journal of Health Geographics*, **3**, 19.

Dzigbodi-Adjimah, K. (1993). "Geology and geochemical patterns of the Birimian gold deposits, Ghana, West Africa." *Journal of Geochemical Exploration*, **47**, 305-320.

Eddyani, M., Ofori-Adjei, D., Teugels, G., De Weirdt, D., Boakye, D., Meyers, W.M., and Portaels, F. (2004). "Potential role for fish in transmission of *Mycobacterium ulcerans* disease (Buruli ulcer): an environmental study." *Applied and Environmental Microbiology*, **70**, 5679-5681.

Faber, N.M., Buydens, L.M.C., and Kateman, G. (1993). "Standard errors in the eigenvalues of a cross-product matrix: theory and applications." *Journal of Chemometrics*, **7**, 495-526.

Fyfe, J.A.M., Lavender, C.J., Handasyde, K.A., Legione, A.R., O'Brien, C.R., Stinear, T.P., Pidot, S.J., Seemann, T., Benbow, M.E., Wallace, J.R., McCowan, C., Johnson, P.D.R. (2010). "A Major role for mammals in the ecology of *Mycobacterium ulcerans*." *PLoS Neglected Tropical Diseases*, **4**, e791.

Ghana Meteorological Service (2010). Daily precipitation data 1960-2010. Obtained from GMet 28 June 2010.

Hausermann, H. (n.d.). "Health governance and practice surrounding Buruli ulcer." Manuscript in preparation.

Hayman, J. (1991). "Postulated epidemiology of *Mycobacterium ulcerans* infection." *International Journal of Epidemiology*, **20**, 1093-1098.

Iivanainen, E.K., Martikainen, P.J., Väänänen, P.K., and Katila, M.-L. (1993). "Environmental factors affecting the occurrence of Mycobacteria in brook waters." *Applied and Environmental Microbiology*, **59**, 398-404.

Johnson, P.T.J., Townsend, A.R., Cleveland, C.C., Glibert, P.M., Howarth, R.W., McKenzie, V.J., Rejmankova, E., and Ward, M.H. (2010). "Linking environmental nutrient enrichment and disease emergence in humans and wildlife." *Ecological Applications*, **20**, 16-29.

Legendre, P. and Legendre, L. (1998). *Numerical Ecology*. Elsevier: New York. 853 p.

Marsollier, L., Sévérin, T., Aubry, J., Merritt, R.W., Saint André, J.-P., Legras, P., Manceau, A.L., Chauty, A., Carbonnelle, B., and Cole, S.T. (2004). "Aquatic snails, passive hosts of *Mycobacterium ulcerans*." *Applied and Environmental Microbiology*, **70**, 6296-6298.

McCune, B. and Grace, J.B. (2002). *Analysis of Ecological Communities*. MjM Software Design: Gleneden Beach, Oregon. 300 p.

Merritt, R.W., Benbow, M.E., Small, P.L.C. (2005). "Unraveling an emerging disease associated with disturbed aquatic environments: the case of Buruli ulcer." *Frontiers in Ecology and the Environment*, **3**, 323-331.

Merritt, R.W., Walker, E.D., Small, P.L.C., Wallace, J.R., Johnson, P.D.R., Benbow, M.E., and Boakye, D.A. (2010). "Ecology and transmission of Buruli ulcer disease: a systematic review." *PLoS Neglected Tropical Diseases*, **4**, e911.

Multilateral Investment Guarantee Agency (2000). "Geological Map: Ghana." *Africa Mining 2000: Investment and Business Opportunities Symposium*, 4-6 December 2000, Ouagadougou, Burkina Faso.

<http://www.fdi.net/documents/WorldBank/conferences/mining2000/Africadata/Ghana/Maps/Geol.pdf>

National Buruli Ulcer Control Programme-Ghana (2008). *Summary Feedback Report, Buruli Ulcer Control Programme, January-March 2008*.

NIST/SEMATECH e-Handbook of Statistical Methods, http://www.itl.nist.gov/div898/handbook/, 30 March 2012.

Portaels, F., Elsen, P., Guimaraes-Peres, A., Fonteyne, P.-A., and Meyers, W.M. (1999). "Insects in the transmission of *Mycobacterium ulcerans* infection." *The Lancet*, **353**, 986.

R Development Core Team: *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2011.

Raghunathan, P.L., Whitney, E.A.S., Asahoa, K., Stienstra, Y., Taylor, T.J. Jr., Amofah, G.K., Ofori-Adjei, D., Dobos, K., Guarner, J., Martin, S., Pathak, S., Klutse, E., Etuaful, S., van der Graaf, W.T.A., van der Werf, T.S., King, C.H., Tappero, J.W., and Ashford, D.A. (2005). "Risk factors for Buruli ulcer disease (*Mycobacterium ulcerans* infection): Results from a case-control study in Ghana." *Clinical Infectious Diseases*, **40**, 1445-1453.

Silliman, S.E., Boukari, M., Crane, P., Azonsi, F., and Neal, C.R. (2007). "Observations on elemental concentrations of groundwater in central Benin." *Journal of Hydrology*, **335**, 374-388.

Sopoh, G.E., Barogui, Y.T., Johnson, R.C., Dossou, A.D., Makoutodé, M., Anagonou, S.Y., Kestens, L., and Portaels, F. (2010). "Family relationship, water contact and occurrence of Buruli ulcer in Benin." *PLoS Neglected Tropical Diseases*, **4**, e746.

Tschakert, P., Parker, E., Kromel, L., Hagarty, J., and Hausermann, H. (n.d.). "Buruli ulcer surveys: Risk factors for Buruli ulcer in gold mining areas of Ghana." Manuscript in preparation.

United Nations (2000). *Millennium Declaration*, New York, NY, 8 September 2000.

Vandelannoote, K., Durnez, L., Amissah, D., Gryseels, S., Dodoo, A., Yeboah, S., Addo, P., Eddyani, M., Leirs, H., Ablordey, A., and Portaels, F. (2010). "Application of real-time PCR in Ghana, a Buruli ulcer-endemic country, confirms the presence of *Mycobacterium ulcerans* in the environment." *FEMS Microbiology Letters*, **304**, 191-194.

Veitch, M.G., Johnson, P.D., Flood, P.E., Leslie, D.E., Street, A.C., and Hayman, J.A. (1997). "A large localized outbreak of *Mycobacterium ulcerans* infection on a temperate southern Australian island." *Epidemiology and Infection*, **119**, 313-318.

Wagner, T., Benbow, M.E., Burns, M., Johnson, R.C., Merritt, R.W., Qi, J., and Small, P.L.C. (2008). "A Landscape-based model for predicting *Mycobacterium ulcerans* infection (Buruli ulcer disease) presence in Benin, West Africa." *EcoHealth*, **5**, 69-79.

Walker, C.H., Hopkin, S.P., Sibly, R.M., and Peakall, D.B. (2006). *Principles of Ecotoxicology*. Taylor & Francis: New York, NY. 316 p.

Williamson, H.R., Benbow, M.E., Nguyen, K.D., Beachboard, D.C., Kimbirauskas, R.K., McIntosh, M.D., Quaye, C., Ampadu, E.O., Boakye, D., Merritt, R.W., and Small, P.L.C. (2008). "Distribution of *Mycobacterium ulcerans* in Buruli ulcer endemic and non-endemic aquatic sites in Ghana." *PLoS Neglected Tropical Diseases*, **2**, e205.

World Health Organization (2007). "Buruli ulcer disease (*Mycobacterium ulcerans* infection). Fact Sheet No. 199. Geneva, Switzerland.

World Health Organization (2010). *First WHO report on neglected tropical diseases: working to overcome the global impact of neglected tropical diseases*. Geneva, Switzerland.

World Health Organization (2011). "Buruli ulcer: Number of new cases reported, 2010." Geneva, Switzerland.

World Health Organization (2012). *Buruli Ulcer*. Geneva, Switzerland. <www.who.int/buruli>

Appendix A

GPS Coordinates of Sampling Locations

Sample ID	Community	Type of Source	Date	Time	Latitude	Longitude	Elevation
P-BW-1	Pokukrom	Built Water	7/14/2011	12:00	6° 00' 57.0" N	1° 48' 40.7" W	118 m
P-BW-2	Pokukrom	Built Water	7/14/2011	12:10	6° 00' 53.6" N	1° 48' 46.7" W	117 m
P-BW-3	Pokukrom	Built Water	7/14/2011	12:20	6° 00' 53.6" N	1° 48' 49.6" W	116 m
P-BW-4	Pokukrom	Built Water	7/14/2011	12:30	6° 00' 54.9" N	1° 48' 50.8" W	114 m
P-BW-5	Pokukrom	Built Water	7/14/2011	12:35	6° 00' 50.1" N	1° 48' 48.9" W	110 m
P-RS-1	Pokukrom	River/Stream	7/13/2011	13:45	6° 00' 41.5" N	1° 48' 52.0" W	116 m
P-RS-2	Pokukrom	River/Stream	7/13/2011	14:00	6° 00' 46.0" N	1° 48' 43.9" W	110 m
P-GP-1	Pokukrom	Galamsey Pit	7/13/2011	10:50	6° 01' 08.8" N	1° 48' 43.9" W	98 m
P-GP-2	Pokukrom	Galamsey Pit	7/13/2011	11:05	6° 01' 09.3" N	1° 48' 44.1" W	109 m
P-GP-3	Pokukrom	Galamsey Pit	7/13/2011	11:55	6° 00' 57.1" N	1° 48' 53.7" W	106 m
P-GP-4	Pokukrom	Galamsey Pit	7/13/2011	12:05	6° 00' 57.1" N	1° 48' 51.7" W	117 m
P-BU-1	Pokukrom	BU Hot Spot	7/14/2011	11:45	6° 01' 02.6" N	1° 48' 42.5" W	116 m
P-SW-1	Pokukrom	Swamp	7/13/2011	13:35	6° 00' 42.4" N	1° 48' 55.1" W	113 m
P-SW-2	Pokukrom	Swamp	7/14/2011	13:10	6° 00' 46.1" N	1° 48' 49.2" W	120 m
B-BW-1	Betenase	Built Water	7/7/2011	14:20	6° 09' 12.8" N	2° 02' 26.0" W	154 m
B-BW-2	Betenase	Built Water	7/7/2011	14:55	6° 09' 14.7" N	2° 02' 13.9" W	155 m
B-RS-1	Betenase	River/Stream	7/7/2011	15:15	6° 09' 09.6" N	2° 02' 09.7" W	155 m
B-RS-2	Betenase	River/Stream	7/7/2011	16:05	6° 09' 26.2" N	2° 02' 10.6" W	138 m
B-GP-1	Betenase	Galamsey Pit	7/7/2011	12:35	6° 09' 51.2" N	2° 01' 33.4" W	136 m
B-GP-2	Betenase	Galamsey Pit	7/7/2011	12:35	6° 09' 51.2" N	2° 01' 33.4" W	136 m
B-GP-3	Betenase	Galamsey Pit	7/7/2011	12:35	6° 09' 51.2" N	2° 01' 33.4" W	136 m
B-GP-4	Betenase	Galamsey Pit	7/7/2011	12:35	6° 09' 51.2" N	2° 01' 33.4" W	136 m
B-BU-1	Betenase	BU Hot Spot	7/7/2011	14:35	6° 09' 13.1" N	2° 02' 22.2" W	160 m

B-BU-2	Betenase	BU Hot Spot	7/7/2011	14:35	6° 09' 13.0" N	2° 02' 22.1" W	160 m
B-SW-1	Betenase	Swamp	7/7/2011	15:20	6° 09' 10.2" N	2° 02' 10.8" W	156 m
B-SW-2	Betenase	Swamp	7/7/2011	15:30	6° 09' 10.7" N	2° 02' 10.6" W	156 m
K-BW-1	Kedadwen	Built Water	7/6/2011	7:40	5° 01' 35.6" N	2° 05' 06.1" W	65 m
K-BW-2	Kedadwen	Built Water	7/6/2011	8:00	5° 01' 42.5" N	2° 05' 02.7" W	75 m
K-BW-3	Kedadwen	Built Water	7/6/2011	8:10	5° 01' 44.3" N	2° 05' 01.6" W	88 m
K-BW-4	Kedadwen	Built Water	7/6/2011	8:15	5° 01' 43.8" N	2° 05' 06.9" W	62 m
K-BW-5	Kedadwen	Built Water	7/6/2011	8:30	5° 01' 42.0" N	2° 05' 11.0" W	73 m
K-BW-6	Kedadwen	Built Water	7/6/2011	8:45	5° 01' 48.0" N	2° 05' 10.5" W	76 m
K-RS-1	Kedadwen	River/Stream	7/5/2011	11:40	5° 01' 48.2" N	2° 04' 42.9" W	99 m
K-RS-2	Kedadwen	River/Stream	7/5/2011	13:00	5° 01' 17.7" N	2° 05' 09.9" W	66 m
K-RS-3	Kedadwen	River/Stream	7/5/2011	13:10	5° 01' 57.6" N	2° 05' 12.2" W	76 m
K-GP-1	Kedadwen	Galamsey Pit	7/5/2011	13:35	5° 02' 37.9" N	2° 05' 45.1" W	75 m
K-GP-2	Kedadwen	Galamsey Pit	7/5/2011	13:50	5° 02' 37.3" N	2° 05' 44.8" W	102 m
K-GP-3	Kedadwen	Galamsey Pit	7/5/2011	14:00	5° 02' 37.2" N	2° 05' 45.8" W	75 m
K-GP-4	Kedadwen	Galamsey Pit	7/5/2011	14:15	5° 02' 36.6" N	2° 05' 47.7" W	94 m
K-SW-1	Kedadwen	Swamp	7/6/2011	7:30	5° 01' 36.4" N	2° 05' 06.6" W	72 m
K-SW-2	Kedadwen	Swamp	7/6/2011	9:00	5° 01' 48.6" N	2° 05' 06.6" W	64 m
A-BW-1	Ayanfuri	Built Water	7/10/2011	13:00	5° 57' 53.2" N	1° 53' 38.0" W	191 m
A-BW-2	Ayanfuri	Built Water	7/10/2011	13:10	5° 58' 02.9" N	1° 53' 43.0" W	187 m
A-BW-3	Ayanfuri	Built Water	7/10/2011	13:25	5° 57' 49.3" N	1° 53' 40.8" W	190 m
A-BW-4	Ayanfuri	Built Water	7/10/2011	13:45	5° 58' 06.6" N	1° 53' 46.7" W	165 m
A-BW-5	Ayanfuri	Built Water	7/10/2011	14:00	5° 57' 46.8" N	1° 53' 47.3" W	171 m
A-BW-6	Ayanfuri	Built Water	7/10/2011	14:10	5° 57' 31.4" N	1° 53' 51.6" W	187 m
A-RS-1	Ayanfuri	River/Stream	7/10/2011	14:25	5° 57' 43.4" N	1° 53' 49.7" W	166 m
A-RS-2	Ayanfuri	River/Stream	7/10/2011	14:45	5° 58' 21.6" N	1° 52' 44.0" W	182 m
A-RS-3	Ayanfuri	River/Stream	7/10/2011	15:20	5° 57' 44.9" N	1° 53' 27.3" W	167 m
A-RS-4	Ayanfuri	River/Stream	7/12/2011	15:20	5° 57' 57.3" N	1° 53' 49.5" W	128 m
A-GP-1	Ayanfuri	Galamsey Pit	7/11/2011	10:10	5° 58' 02.1" N	1° 53' 30.6" W	158 m
A-GP-2	Ayanfuri	Galamsey Pit	7/11/2011	10:25	5° 58' 03.7" N	1° 53' 33.6" W	158 m

A-GP-3	Ayanfuri	Galamsey Pit	7/11/2011	11:35	5° 56' 19.7" N	1° 53' 34.2" W	177 m
A-GP-4	Ayanfuri	Galamsey Pit	7/11/2011	11:45	5° 56' 19.5" N	1° 53' 34.1" W	165 m
A-SW-1	Ayanfuri	Swamp	7/11/2011	12:05	5° 56' 56.4" N	1° 53' 55.4" W	183 m
A-SW-2	Ayanfuri	Swamp	7/12/2011	15:30	5° 57' 57.8" N	1° 53' 48.7" W	150 m
N-BW-1	Nangruma	Built Water	6/28/2011	10:30	10° 22' 37.9" N	1° 29' 24.0" W	136 m
N-RS-1	Nangruma	River/Stream	6/28/2011	11:50	10° 21' 47.1" N	1° 29' 37.0" W	153 m
N-RS-2	Nangruma	River/Stream	6/28/2011	15:30	10° 21' 59.2" N	1° 29' 27.2" W	158 m
N-RS-3	Nangruma	River/Stream	6/29/2011	13:05	10° 21' 47.3" N	1° 29' 40.3" W	158 m
N-RS-4	Nangruma	River/Stream	6/29/2011	13:40	10° 22' 00.8" N	1° 29' 26.5" W	164 m
N-GP-1	Nangruma	Galamsey Pit	6/29/2011	10:30	10° 21' 52.9" N	1° 29' 33.0" W	160 m
N-GP-2	Nangruma	Galamsey Pit	6/28/2011	11:25	10° 21' 48.8" N	1° 29' 38.7" W	167 m
N-GP-3	Nangruma	Galamsey Pit	6/29/2011	11:30	10° 22' 01.6" N	1° 29' 28.1" W	169 m
N-GP-4	Nangruma	Galamsey Pit	6/28/2011	15:50	10° 21' 59.3" N	1° 29' 25.4" W	158 m
N-SW-2	Nangruma	Swamp	6/28/2011	14:05	10° 21' 47.3" N	1° 29' 40.3" W	154 m

Appendix B

June-July 2011 Data

All data are reported as mg/L.

Sample IDs are organized by the convention: first letter of community name-water body type-number. Water bodies are coded as follows: BW stands for built water sources, including wells and boreholes. RS stands for rivers and streams, GP for galamsey pit, BU for BU hot spot, and SW for swamp. DUP-# indicates the duplicate sample for that study site (number corresponds to study area number). Duplicate samples are listed in the table immediately below the sample with which they pair. TB-# corresponds to the trip blank associated with each study site. Again, the number represents study area number.

Sample ID	Alk	Cl	TH	NO3	NO2	PO4	NH4	F	Mn	SO4	S	Cd	Cu	Fe	Pb	Zn	As	Se
P-BW-1	40	2.84	56.8	0.12	0.09	6.4	0	0	0.004	8	0.01	0.001	0.01	0.3	0.005	0.025	0.002	0.0015
P-BW-2	60	6.39	127.8	0.08	0.06	20	0.34	1	0.007	10	0.01	0.013	0.01	2.4	0.005	0.025	0.019	0.0015
P-BW-3	70	4.26	85.2	0.16	0.11	15	0.14	1.3	0.018	14	0.01	0.001	0.01	0.2	0.005	0.025	0.004	0.0015
P-BW-4	160	2.84	56.8	0.16	0.12	19	0.2	1.4	0.018	16	0.01	0.001	0.01	0.05	0.005	0.025	0.002	0.0015
P-BW-5	110	4.26	85.2	0.02	0.01	5.2	0.18	N	0.009	18	0.01	0.001	0.01	0.3	0.005	0.08	0.006	0.0015
DUP-1	20	30.53	610.6	0.06	0.03	0.7	N	N	0.005	32	0.01	0.001	0.01	0.05	0.005	0.025	0.002	0.0015
P-RS-1	300	2.84	56.8	0.14	0.11	23	0.47	1.1	0.006	16	0.01	0.001	0.01	8	0.005	0.025	0.026	0.0015
P-RS-2	60	11.36	227.2	0.29	0.22	12	0.31	1.2	0.024	14	0.01	0.001	0.01	21.5	0.005	0.025	0.06	0.0015
P-GP-1	240	2.13	42.6	0.08	0.05	12	0.21	N	0.007	16	0.01	0.015	0.37	180	0.07	0.25	0.46	0.01
P-GP-2	130	3.55	71	0.05	0.04	7.5	0.06	N	0.002	12	0.01	0.04	0.73	510	0.63	1.12	0.74	0.004
P-GP-3	200	2.84	56.8	0.09	0.07	15	0.07	N	0.006	17	0.01	0.001	0.01	6.2	0.005	0.025	0.016	0.0015
P-GP-4	350	2.84	56.8	0.05	0.04	12	0.07	N	0.001	14	0.01	0.001	0.01	6.8	0.005	0.025	0.006	0.0015
P-BU-1	250	9.94	198.8	0.29	0.22	14	0.35	1.3	0.023	18	0.01	0.12	0.86	1640	0.81	2.91	3	0.016

P-SW-1	150	2.84	56.8	0.12	0.09	18	0.14	1.3	0.008	13	0.01	0.001	0.01	5	0.005	0.025	0.012	0.0015
P-SW-2	60	6.39	127.8	0.03	0.02	1.4	N	N	0.004	21	0.01	0.001	0.01	12.5	0.005	0.025	0.044	0.0015
B-BW-1	10	2.13	42.6	0.08	0.06	12	<0.01	<0.05	0.009	17	0.01	0.001	0.1	0.05	0.02	0.025	0.002	0.0015
B-BW-2	20	2.13	42.6	N	N	6.4	N	0.01	0.004	9	0.01	0.001	0.01	0.05	0.01	0.025	0.001	0.0015
B-RS-1	50	2.13	42.6	0.08	0.05	15	<0.01	0.5	0.006	10	0.01	0.001	0.01	5	0.005	0.025	0.006	0.0015
B-RS-2	60	2.13	42.6	0.19	0.17	2.9	0.03	0.2	0.002	13	0.01	0.001	0.01	2.6	0.005	0.025	0.004	0.0015
DUP-2	50	2.13	42.6	0.06	0.03	1.8	N	0.05	0.002	10	0.01	0.001	0.01	2.9	0.005	0.025	0.016	0.0015
B-GP-1	220	4.26	85.2	0.11	0.08	16	0.49	1.2	0.008	21	0.01	0.024	0.44	344	0.24	0.8	0.26	0.01
B-GP-2	90	4.26	85.2	0.02	0.01	2.9	N	N	N	16	0.01	0.023	0.41	332	0.23	0.79	0.16	0.006
B-GP-3	250	8.52	170.4	0.06	0.05	9.9	0.04	N	0.01	17	0.01	0.052	0.94	728	0.54	1.38	0.34	0.01
B-GP-4	400	4.26	85.2	0.12	0.09	16	0.26	N	0.005	18	0.01	0.026	0.45	364	0.32	0.67	0.26	0.016
B-BU-1	190	4.97	99.4	0.08	0.05	8.7	<0.01	0.2	0.013	15	0.01	0.002	0.05	38.3	0.04	0.42	0.032	0.004
B-BU-2	280	58.22	1164.4	0.08	0.05	5.2	0.01	0.05	0.007	10	0.01	0.005	0.08	73.9	0.07	0.71	0.046	0.0015
B-SW-1	35	3.55	71	0.12	0.09	9.9	0.01	0.1	0.006	16	0.01	0.001	0.01	4.6	0.005	0.025	0.004	0.0015
B-SW-2	50	3.55	71	0.62	0.48	8.7	0.03	0.4	0.004	7	0.02	0.001	0.01	10.3	0.005	0.025	0.001	0.0015
K-BW-1	110	1.42	28.4	0.41	0.32	0.7	N	0.25	N	3	0.01	0.001	0.01	0.05	0.005	0.025	0.002	0.0015
K-BW-2	20	2.13	42.6	0.11	0.08	2.9	0.01	0.1	0.006	8	0.01	0.001	0.01	0.2	0.005	0.025	0.01	0.0015
K-BW-3	50	6.39	127.8	0.62	0.47	1.8	N	N	0.009	9	0.01	0.001	0.01	0.6	0.005	0.025	0.001	0.0015
K-BW-4	10	9.23	184.6	0.97	0.75	1.8	0.01	0.2	0.016	20	0.01	0.001	0.01	0.05	0.005	0.025	0.004	0.0015
K-BW-5	60	5.68	113.6	0.57	0.44	2.9	N	0.05	0.002	8	0.01	0.001	0.01	0.05	0.005	0.07	0.006	0.0015
DUP-3	60	2.84	56.8	0.57	0.44	7.5	0.03	1.4	0.005	14	0.01	0.001	0.01	0.05	0.005	0.05	0.001	0.0015
K-BW-6	40	4.97	99.4	0.05	0.04	14	0.18	1.1	0.023	10	0.01	0.001	0.001	0.025	0.005	0.025	0.026	0.0015
K-RS-1	60	1.42	28.4	0.09	0.07	0.7	0.01	0.05	0.005	3	0.01	0.001	0.01	1.8	0.005	0.025	0.006	0.0015
K-RS-2	40	2.84	56.8	0.16	0.12	2.9	N	N	0.001	13	0.01	0.001	0.01	2.4	0.005	0.025	0.002	0.0015
K-RS-3	40	1.42	28.4	0.09	0.06	15	<0.01	<0.05	0.005	18	0.01	0.001	0.01	6.3	0.005	0.025	0.008	0.0015
K-GP-1	150	1.42	28.4	0.66	0.51	6.4	0.01	1.45	0.006	56	0.01	0.001	0.01	6.9	0.005	0.025	0.006	0.0015
K-GP-2	250	2.84	56.8	0.03	0.01	0.7	N	N	0.001	46	0.01	0.077	1.64	1110	0.32	1.43	0.2	0.006
K-GP-3	150	2.84	56.8	0.09	0.042	11	0.25	N	0.008	19	0.01	0.1	1.7	1280	0.39	2.39	0.2	0.012
K-GP-4	20	2.13	42.6	0.15	0.1	6.4	0.01	0.25	0.002	8	0.01	0.024	0.54	356	0.94	0.33	0.054	0.004
K-SW-1	160	1.42	28.4	0.09	0.06	6.4	0.01	1.15	0.006	14	0.01	0.001	0.01	17	0.005	0.025	0.01	0.0015

K-SW-2	90	3.55	71	0.61	0.47	9.9	0.05	1.45	0.007	18	0.01	0.001	0.01	10.3	0.005	0.025	0.004	0.0015
A-BW-1	110	4.97	99.4	0.13	0.09	6.4	N	N	0.006	9	0.01	0.001	0.01	0.2	0.005	0.18	0.006	0.0015
A-BW-2	110	3.55	71	0.28	0.22	4.1	N	0.05	0.012	9	0.01	0.001	0.01	0.1	0.005	0.15	0.004	0.0015
A-BW-3	200	4.26	85.2	0.08	0.06	21	0.23	0.9	0.018	12	0.01	0.001	0.01	0.2	0.005	0.15	0.002	0.0015
A-BW-4	110	2.13	42.6	0.84	0.64	4.1	0.01	1.25	0.005	12	0.01	0.001	0.04	0.6	0.005	0.22	0.006	0.0015
A-BW-5	130	2.84	56.8	0.33	0.26	1.8	N	0.1	0.007	9	0.01	0.001	0.01	1.6	0.005	0.16	0.002	0.0015
A-BW-6	110	2.84	56.8	0.75	0.58	2.9	N	0.25	0.017	17	0.01	0.001	0.03	0.2	0.005	0.24	0.002	0.0015
A-RS-1	110	2.84	56.8	0.53	0.41	8.7	0.04	1.45	0.004	13	0.01	0.001	0.01	3.9	0.005	0.025	0.012	0.0015
A-RS-2	30	2.13	42.6	0.21	0.16	7.5	N	0.3	0.004	9	0.01	0.001	0.01	4.4	0.005	0.025	0.008	0.0015
A-RS-3	40	2.84	56.8	0.26	0.2	11	N	N	0.005	13	0.01	0.001	0.01	3.7	0.005	0.025	0.008	0.0015
A-RS-4	70	2.84	56.8	0.13	0.09	2.9	N	1.25	0.006	9	0.01	0.001	0.01	2.5	0.005	0.025	0.022	0.0015
DUP-4	80	2.84	56.8	0.16	0.12	1.8	N	N	0.003	13	0.01	0.001	0.01	2.6	0.005	0.025	0.02	0.0015
A-GP-1	380	1.42	28.4	0.04	0.07	3.9	N	N	0.016	19	0.01	0.18	0.5	2220	1.19	0.52	114	0.024
A-GP-2	60	4.97	99.4	0.07	0.05	12	0.05	N	0.012	13	0.01	0.2	2.01	2410	2.48	1.57	30	0.02
A-GP-3	20	2.13	42.6	0.22	0.17	8.7	N	N	0.006	16	0.01	0.001	0.01	14.2	0.01	0.025	0.36	0.0015
A-GP-4	150	1.42	28.4	0.12	0.09	16	0.08	1.4	0.007	12	0.01	0.001	0.01	8.8	0.005	0.025	0.18	0.0015
A-SW-1	100	1.42	28.4	0.44	0.39	9.9	N	1.5	0.007	16	0.01	0.001	0.01	2.3	0.005	0.025	0.03	0.0015
A-SW-2	90	2.13	42.6	0.66	0.51	4.1	N	1.3	0.007	13	0.01	0.001	0.01	4.5	0.005	0.025	0.018	0.006
N-BW-1	300	11.36	227.2	0.79	0.61	6.4	N	N	0.003	10	0.01	0.001	0.01	0.7	0.005	0.025	0.002	0.0015
N-RS-1	220	2.13	42.6	0.12	0.09	19	0.14	1.35	0.009	17	0.01	0.015	0.22	214	0.13	0.65	3.4	0.0015
N-RS-2	40	2.84	56.8	0.7	0.54	6.4	0.01	1.3	0.004	10	0.01	0.001	0.01	12	0.005	0.025	1.2	0.0015
N-RS-3	310	105.79	2115.8	0.62	0.47	9.9	N	N	0.006	14	0.01	0.001	0.01	0.9	0.005	0.025	0.016	0.0015
N-RS-4	130	2.84	56.8	0.11	0.08	26	0.12	0.03	0.004	14	0.01	0.008	0.1	130	0.03	0.14	24	0.0015
DUP-5	80	2.13	42.6	0.21	0.17	1.8	N	N	0.003	7	0.01	0.01	0.11	150	0.03	0.15	16.2	0.0015
N-GP-1	120	2.13	42.6	0.13	0.09	4.1	N	N	0.005	14	0.01	0.04	0.79	570	0.92	1.86	340	0.004
N-GP-2	20	1.42	28.4	0.66	0.51	0.7	N	N	0.001	5	0.01	0.043	0.68	640	0.37	2	38	0.0015
N-GP-3	320	1.42	28.4	0.44	0.34	2.9	N	N	0	9	0.01	0.034	0.54	482	0.11	1.09	194	0.0015
N-GP-4	80	1.42	28.4	0.22	0.16	6.4	N	N	0.004	10	0.01	0.006	0.1	78.3	0.09	5.82	19.6	0.0015
N-SW-2	220	2.13	42.6	0.52	0.4	4.1	N	N	0.004	12	0.01	0.001	0.01	4	0.05	0.025	0.42	0.0015
TB-1	20	1.42	28.4	0.04	0.03	0.7	N	0.05	<0.001	3	0.01	0.001	0.01	0.05	0.005	0.025	0.001	0.0015

TB-2	10	1.42	28.4	N	N	7.5	0.01	0.05	0.003	8	0.01	0.001	0.01	0.05	0.005	0.025	0.001	0.0015
TB-3	20	7.1	142	0.16	0.12	1.8	N	N	0.003	13	0.01	0.001	0.01	0.05	0.005	0.025	0.004	0.0015
TB-4	20	2.84	56.8	0.22	0.17	7.5	0.01	1.2	0.008	14	0.01	0.001	0.01	0.05	0.005	0.025	0.002	0.0015
TB-5	20	1.42	28.4	0.16	0.12	4.1	N	N	0.002	9	0.01	0.001	0.01	0.05	0.005	0.025	0.006	0.0015

Appendix C

Scores and Loadings for Principal Components Analysis

Analysis is presented in main body. Raw data are shown in tabular format in Appendix B.

Scores

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
P-BW-1	-1.11258	-0.72308	1.386047	-0.45881	0.717111	0.448797	0.388864	-0.02809	0.142832	0.193881
P-BW-2	-0.52591	2.785911	0.05622	-0.04838	0.421138	1.067757	0.205415	0.24341	-0.00904	0.571108
P-BW-3	-0.82918	2.238277	-0.00165	0.846544	1.520767	-0.09031	-0.83476	-0.30539	-1.19806	-0.26262
P-BW-4	-0.51971	3.081123	-0.55572	0.746326	0.598578	0.137976	-0.47046	-0.53065	-0.61306	-0.01385
P-BW-5	-0.51215	0.513482	1.013891	-0.39189	1.409985	-0.42894	0.856859	-1.01653	-0.49323	-0.89132
P-RS-1	0.04113	4.181378	-0.59598	-0.698	-0.91148	-0.41189	-0.19229	0.075999	1.127967	-0.43786
P-RS-2	-0.75516	2.446942	-1.64807	1.083637	0.23337	1.508227	-0.3364	-0.37903	-0.9878	0.367805
P-GP-1	1.255014	1.232384	0.267081	-0.97394	-0.71785	-0.27771	0.602047	-0.15264	1.15146	-0.34214
P-GP-2	1.828226	-0.72935	1.125605	-0.87079	-0.3346	0.111425	0.34606	0.891535	0.012102	0.372817
P-GP-3	-0.3338	0.957625	0.571402	-1.25995	-1.01952	-0.15816	0.352593	-0.32496	0.437471	0.198745
P-GP-4	-0.1169	0.772953	0.603549	-2.02889	-1.33385	-0.55862	0.572354	-0.49403	1.677258	0.009042
P-BU-1	5.623951	1.916367	-2.88497	1.759982	-0.48506	1.243546	-0.02179	1.120487	-0.27761	-0.5954
P-SW-1	-0.64161	2.252332	0.044617	-0.1324	-0.64365	0.300518	-0.2872	0.094907	0.130855	0.877503
P-SW-2	-0.77731	-0.86604	1.378682	-0.95937	0.143508	-0.89029	0.56356	-0.57613	-0.75726	0.022923
B-BW-1	-0.81131	0.204709	1.579049	-0.14418	1.842358	-0.91521	-0.13679	-0.2889	-1.26906	-0.77614

B-BW-2	-0.97961	-0.43934	2.117845	-0.73782	1.570483	-0.32172	-0.02178	-0.00832	-0.38394	-0.36669
B-RS-1	-0.91458	0.732756	1.360202	-0.50089	0.742102	-0.1083	-0.45223	0.270329	-0.36544	0.1798
B-RS-2	-1.16807	-0.9111	1.232556	-0.30662	0.207113	-0.59027	0.009319	0.146629	0.10261	-0.1453
B-GP-1	1.883786	3.323037	-0.62465	0.111839	-0.23662	-0.81041	0.041831	0.380092	0.321168	-0.88076
B-GP-2	0.777895	-1.15342	1.579059	-0.8773	0.395297	-0.81561	0.531796	0.268565	-0.2317	-0.27667
B-GP-3	3.142115	-0.14623	-0.03372	-0.84509	0.13081	-0.53526	0.347774	0.362354	-0.04948	-0.47799
B-GP-4	2.647287	1.70149	-0.46831	-1.02819	0.11348	-1.40743	0.559687	-0.18727	2.073	-1.61377
B-BU-1	-0.11586	0.294301	0.04267	-0.35166	-0.04388	0.873775	1.102361	-1.06967	0.12573	0.197792
B-BU-2	-0.32073	-1.14526	-3.26634	-3.82796	0.636962	1.619512	1.450512	-0.83846	-0.08587	0.886483
B-SW-1	-0.88533	-0.0458	1.308136	-0.67713	0.247416	-0.62099	-0.2979	0.128603	-0.66714	-0.10892
B-SW-2	-1.83983	-1.00947	-0.31833	0.931209	0.078226	-0.14642	-0.89581	1.065931	0.723329	-0.33833
K-BW-1	-1.59811	-1.6712	0.532753	0.175574	-0.10868	0.356526	-0.0113	0.532634	1.469438	0.108918
K-BW-2	-1.16805	-0.78317	1.773187	-0.27345	1.665109	-0.31472	-0.21162	0.017345	-0.31554	-0.55465
K-BW-3	-1.92196	-1.90982	-0.55715	1.459045	1.610585	-0.01314	-0.00561	-0.02325	0.311414	-0.99742
K-BW-4	-2.50211	-2.49071	-2.99417	4.04556	1.651573	0.954957	1.632347	-1.23937	-0.21568	-0.59855
K-BW-5	-1.89324	-2.05736	-0.34969	0.976139	0.503178	0.40458	0.50044	0.2348	0.947187	-0.18794
K-BW-6	-0.68421	2.439193	-0.03338	0.881778	1.769784	1.349638	-0.19802	-0.73375	-1.25462	0.16363
K-RS-1	-0.95594	-0.9138	1.643164	-1.17485	-0.81785	0.871335	-0.32216	0.568995	0.641091	0.734125
K-RS-2	-1.09431	-1.09703	1.500859	-0.79951	-0.3022	-0.45488	-0.02885	0.323253	0.032521	0.157762
K-RS-3	-0.75385	0.402363	1.422273	-0.70547	0.086681	-0.56779	-0.01843	0.025686	-0.6767	0.025977
K-GP-1	-1.16696	-0.1049	-1.79139	2.019553	-3.0573	-4.11338	0.042109	-0.87738	-1.81086	-0.07287
K-GP-2	3.958897	-1.68517	-0.08118	-0.01183	-2.34306	-1.88391	3.38716	-1.05657	-1.2727	0.991154
K-GP-3	4.953144	-0.00934	-0.43712	0.656867	-0.67565	1.299503	2.466695	0.772561	-0.27602	0.225039
K-GP-4	0.849049	-1.28342	1.303652	0.296473	1.230227	0.298354	0.071729	0.733004	-0.03414	0.936217
K-SW-1	-0.87701	0.373945	0.29917	0.251578	-0.51201	0.664306	1.022847	-0.83944	0.394852	1.356861
K-SW-2	-1.69418	0.128069	-1.17626	1.732152	-0.64926	-0.79154	-0.81704	0.501164	-0.01325	0.257082
A-BW-1	-0.88406	-0.54925	0.739467	-0.63065	0.059904	0.879055	0.751415	-0.3287	0.41397	0.393987
A-BW-2	-1.09431	-0.76076	0.101089	0.331213	0.424845	0.977	0.547438	-0.52251	0.314612	-0.02326
A-BW-3	-0.20519	3.180264	-0.59911	0.051726	0.254087	1.254357	0.245699	-0.71446	-0.01804	0.197372
A-BW-4	-2.12572	-1.31926	-1.71685	2.400664	-0.96206	-0.06163	-0.45046	0.710855	0.990092	0.206244

A-BW-5	-1.2354	-1.2381	0.15441	0.310659	0.121493	0.627454	0.644224	-0.34513	0.781583	-0.02828
A-BW-6	-1.62732	-1.36234	-1.75938	2.211313	-0.29615	0.361668	0.105631	-0.37542	0.109383	-0.60954
A-RS-1	-1.64486	-0.03655	-0.90366	1.29137	-1.18526	-0.02053	-0.44007	0.570097	0.647911	0.929053
A-RS-2	-1.20303	-0.45685	1.273922	-0.38649	0.167571	-0.2189	-0.61099	0.61861	-0.01991	0.078068
A-RS-3	-1.1101	-0.32867	0.938039	-0.40899	-0.07204	-0.36382	-0.46464	0.493018	-0.15968	-0.10174
A-RS-4	-1.17066	-0.00308	0.911857	-0.01494	-0.15973	0.056533	-0.51555	0.241287	-0.06646	0.925269
A-GP-1	7.339016	-1.252	-0.44942	0.913763	2.103553	-1.51818	-1.49811	-2.08276	1.69425	-0.81462
A-GP-2	9.279212	-1.77714	0.157703	1.385454	2.166947	-0.19621	-1.22081	2.061766	-0.6581	1.662031
A-GP-3	-1.03006	-0.40096	1.205402	-0.4156	-0.03061	-0.75082	-0.57154	0.387775	-0.58498	-0.24629
A-GP-4	-0.72487	1.93557	0.515976	-0.29481	-0.49413	-0.41692	-0.97834	0.411454	0.041493	0.600632
A-SW-1	-1.43707	0.383081	-0.43797	0.926694	-1.31647	-0.86236	-1.36507	0.80587	-0.03101	0.605876
A-SW-2	-1.52218	-0.80944	-1.18404	1.837764	-1.01779	-0.29342	-0.77762	0.676258	0.81979	0.429552
N-RS-3	-1.44795	-2.00379	-7.35267	-5.93275	1.238571	-0.91665	-1.44582	0.699721	-0.7914	-0.04814
N-GP-1	3.513172	-2.63677	1.217601	-0.67898	-3.21396	2.486833	-3.8837	-3.32501	-0.57788	-0.10812
N-GP-4	0.840462	-1.36754	0.86403	-0.7873	-3.12207	2.692726	0.4344	2.228824	-1.46971	-2.94354

	PC11	**PC12**	**PC13**	**PC14**	**PC15**	**PC16**	**PC17**	**PC18**	**PC19**
P-BW-1	0.136934	-0.02685	0.057657	-0.29628	-0.07536	-0.04934	0.017147	0.012137	3.80E-17
P-BW-2	1.54043	0.14104	-0.122	-0.91093	0.061311	0.107163	-0.13722	0.003024	5.70E-17
P-BW-3	-0.28917	0.494388	0.245253	0.460129	-0.26089	0.050779	-0.01341	-0.04903	-5.96E-16
P-BW-4	-0.21675	0.111792	0.526382	0.406839	-0.07602	0.128354	-0.01843	-0.01235	-3.73E-16
P-BW-5	0.393691	0.50035	-0.70275	0.150069	0.435038	0.251191	-0.0235	-0.01417	-8.36E-16
P-RS-1	1.422352	-0.06309	-0.20663	0.100139	0.485669	0.702186	0.020437	0.000862	-1.91E-16
P-RS-2	-0.65431	-0.61293	-1.41687	0.032367	0.226984	-0.05087	0.016478	0.008108	-1.04E-15
P-GP-1	-0.13278	-0.81764	-0.62949	-0.04586	-0.19395	-0.64738	-0.04143	-0.02869	-1.88E-16
P-GP-2	0.348984	-0.06106	0.016647	0.434612	0.354363	0.208067	-0.01475	0.028635	4.62E-16
P-GP-3	-0.25926	-1.22694	0.47491	-0.20076	0.185571	0.132747	0.016844	0.005164	4.21E-16
P-GP-4	-0.29646	-0.45722	0.646869	0.363982	0.352959	0.513306	-0.01511	-0.02422	3.69E-16
P-BU-1	-1.11343	0.29645	-0.52934	-0.40298	0.040057	0.225564	0.164269	0.003304	-3.54E-16

P-SW-1	0.078084	0.144688	0.642591	-0.19505	-0.09818	-0.02084	0.018474	0.00613	-1.31E-17
P-SW-2	-0.32889	0.289705	-0.6743	-0.2444	0.291304	-0.07564	-0.00328	-0.00389	-6.00E-16
B-BW-1	0.366254	-0.20861	0.529552	0.078804	-0.29618	-0.01724	-0.00289	0.013485	8.16E-17
B-BW-2	0.222971	0.454974	0.012269	-0.03828	-0.15657	0.057978	0.004936	0.002393	-1.82E-16
B-RS-1	0.145391	-0.09492	0.832018	-0.20645	-0.37715	-0.05754	0.038618	-0.02343	2.68E-16
B-RS-2	0.123284	0.350023	-0.50262	-0.07397	0.088442	0.122951	-0.02186	0.083607	-1.19E-16
B-GP-1	1.261591	0.845165	-1.16191	0.00284	0.194318	-0.33381	0.022576	0.016309	-1.10E-15
B-GP-2	0.130311	0.781113	-0.21127	-0.01948	-0.09717	-0.24717	0.030244	0.01026	-3.28E-16
B-GP-3	-0.60227	-0.25786	0.345255	0.904657	-0.32872	-0.17114	0.003292	0.039889	1.80E-16
B-GP-4	0.481125	-0.11544	0.305106	0.348193	-0.05087	-0.90021	-0.00828	0.007136	-4.15E-16
B-BU-1	-1.04804	-0.36669	0.508345	0.04681	0.132136	-0.40965	0.002654	-0.01959	-3.16E-16
B-BU-2	-0.25403	0.798889	0.337394	-0.19677	0.30891	-0.15816	0.001163	0.015668	-1.11E-15
B-SW-1	-0.13688	-0.55254	-0.02423	-0.26758	-0.09065	-0.04195	0.019183	-0.00035	8.72E-17
B-SW-2	0.56912	-0.51338	0.085675	-0.02234	-0.28199	0.230609	0.041841	-0.00113	7.50E-16
K-BW-1	0.096979	0.365001	-0.34728	0.167006	-0.00434	0.27246	-0.00644	-0.00359	2.47E-16
K-BW-2	-0.03824	0.613602	-0.39199	0.241765	-0.13233	0.146472	-0.01055	-0.02132	-3.50E-16
K-BW-3	0.424326	-0.00975	-0.12252	0.360166	-0.08865	0.253766	0.00081	-0.04106	2.18E-16
K-BW-4	1.045921	-0.41884	0.429491	-0.27333	0.180978	-0.19764	0.008251	0.043305	3.22E-16
K-BW-5	0.867264	0.018756	0.086803	-0.23217	-0.001	0.105029	0.020851	0.012345	4.64E-16
K-BW-6	-0.64856	0.110657	-0.17788	0.253316	-0.13478	-0.11865	-0.02049	0.030458	-8.66E-16
K-RS-1	-1.2419	-0.4859	-1.28699	-0.02139	0.054328	0.040967	-0.00821	-0.0107	-1.89E-16
K-RS-2	-0.1282	-0.09323	-0.66704	-0.30407	0.088385	0.010943	0.009549	-0.01275	-4.44E-17
K-RS-3	0.220458	-0.93994	0.629606	-0.5729	-0.13547	-0.13244	0.04998	-0.00861	4.47E-16
K-GP-1	-0.19572	-0.09258	0.026989	-0.31976	0.655096	-0.17998	-0.04048	-0.03535	9.22E-17
K-GP-2	0.326898	0.221456	-0.10434	0.387505	-0.46426	0.445525	0.036498	0.015605	6.71E-16
K-GP-3	1.023112	-0.26762	-0.43184	-0.16275	-1.17605	0.099473	-0.04849	-0.02718	9.87E-16
K-GP-4	1.01225	0.375016	0.380175	0.345696	0.861942	-0.40084	0.113224	0.009053	1.42E-16
K-SW-1	-0.41637	1.131819	0.559948	-0.13599	0.135476	-0.19314	0.004638	-0.01516	-4.18E-16
K-SW-2	0.261966	0.412258	0.343354	0.032825	-0.15642	0.036981	0.011256	-0.01371	2.60E-16
A-BW-1	-0.29588	-0.23208	0.127197	-0.20246	0.08679	-0.04645	0.007618	-0.02113	3.10E-17

A-BW-2	-0.72971	-0.47001	-0.13302	0.219407	0.091458	0.054434	-0.02422	0.014356	-2.41E-17
A-BW-3	-0.2025	-0.63478	0.607934	0.149437	0.072579	0.115517	-0.00266	0.012839	-1.79E-16
A-BW-4	0.289633	0.451925	0.223119	0.210635	-0.11574	0.129414	-0.00503	-0.0527	6.16E-16
A-BW-5	-0.34974	-0.00709	-0.16842	0.192213	0.165139	0.144254	-0.02333	0.013714	1.09E-17
A-BW-6	-0.84023	-1.21618	-0.30938	0.471166	0.152473	0.132041	-0.03485	-0.0214	4.22E-16
A-RS-1	0.164305	0.548299	0.327941	-0.14851	-0.11344	-0.03604	0.007421	0.000146	2.95E-16
A-RS-2	-0.11122	-0.17319	-0.18636	-0.12977	-0.20229	0.040023	0.019315	-0.00702	1.99E-16
A-RS-3	0.025625	-0.97421	0.105417	-0.31239	-0.1624	0.013893	0.032094	0.004571	4.83E-16
A-RS-4	-0.81296	1.286561	-0.41902	0.19159	-0.12698	-0.0755	-0.02254	-0.04913	-6.21E-16
A-GP-1	-1.65876	0.712191	0.3632	-0.8708	-0.11061	0.421202	-0.02484	0.011985	2.59E-16
A-GP-2	0.412365	-0.8365	0.304422	0.167859	0.683399	-0.08099	-0.08557	-0.02591	1.92E-15
A-GP-3	-0.22635	-0.86324	-0.27703	-0.23527	-0.08504	0.010477	0.032592	0.003122	2.74E-16
A-GP-4	-0.29323	0.485893	0.629837	0.183844	-0.26216	0.099225	0.005483	-0.01811	-3.08E-17
A-SW-1	-0.6033	0.185282	0.162353	0.173732	-0.24587	0.024055	-0.05276	0.16146	2.40E-16
A-SW-2	-0.51634	0.318503	-0.27438	-0.08171	-0.30491	-0.63424	-0.01776	-0.0098	1.83E-17
N-RS-3	0.330727	-0.11642	-0.14371	0.028511	-0.2271	0.073615	-0.009	-0.01143	-8.57E-16
N-GP-1	1.67272	-0.0397	-0.19988	0.253077	-0.16452	-0.14335	0.017725	-0.00158	3.26E-17
N-GP-4	-0.7536	0.80064	0.948749	-0.23479	0.412955	0.019564	-0.05806	-0.00459	6.65E-17

Loadings

	PC1	**PC2**	**PC3**	**PC4**	**PC5**	**PC6**	**PC7**	**PC8**	**PC9**	**PC10**
% Variance Explained	28.10%	13.20%	12.70%	10.80%	7.47%	5.84%	5.20%	3.92%	3.33%	2.67%
pH	0.055803	0.025405	-0.05888	-0.19808	-0.68338	0.210099	-0.0742	0.156474	0.084625	0.331528
Conductivity	-0.03897	-0.08543	-0.20934	0.256252	0.001611	0.497528	0.560615	-0.29031	0.125905	0.295176
Alkalinity	0.190817	0.167863	-0.28668	-0.24169	-0.18214	-0.17238	0.14449	-0.28278	0.543918	-0.17095
Chloride	-0.02674	-0.10658	-0.46783	-0.42476	0.162314	0.029756	-0.0508	0.033332	-0.1771	0.07911
Total Hardness	-0.02674	-0.10658	-0.46783	-0.42476	0.162314	0.029756	-0.0508	0.033332	-0.1771	0.07911

Nitrate	-0.18273	-0.22624	-0.37651	0.35138	-0.09773	-0.07364	-0.13875	0.163924	0.134497	-0.14375
Nitrite	-0.18113	-0.22474	-0.37265	0.354926	-0.09413	-0.08894	-0.15694	0.158368	0.153955	-0.14766
Phosphate	0.06037	0.542482	-0.06508	-0.0752	0.027069	0.019255	-0.14579	0.148154	-0.05457	0.009418
Ammonium	0.114251	0.507688	-0.1283	0.040485	-0.01005	0.114213	0.083554	0.052783	0.121991	-0.21547
Fluoride	-0.08369	0.368297	-0.18783	0.298533	-0.18769	-0.05774	-0.26386	0.118243	-0.098	0.387369
Manganese	0.080619	0.266171	-0.24353	0.272352	0.28027	0.280581	-0.09035	-0.26296	-0.31052	-0.12274
Sulphate	0.102807	0.037186	-0.14296	0.125728	-0.32436	-0.59307	0.280291	-0.34042	-0.4771	-0.05283
Cadmium	0.406786	-0.0865	-0.05578	0.110553	0.115833	-0.05267	-0.03588	0.070757	0.034975	0.119372
Copper	0.377166	-0.10552	-0.01049	0.050134	-0.06186	-0.01283	0.19412	0.137804	-0.15998	0.23414
Iron	0.410746	-0.08899	-0.0589	0.106252	0.09029	-0.05461	-0.01647	0.057082	0.023446	0.102479
Lead	0.375071	-0.13505	0.005957	0.080316	0.129219	0.01783	-0.22816	0.116716	-0.02192	0.233351
Zinc	0.243036	-0.09056	-0.0223	-0.0165	-0.32248	0.365511	0.125187	0.324158	-0.31378	-0.57891
Arsenic	0.16148	-0.16305	0.05759	-0.01734	-0.22043	0.239033	-0.56417	-0.62047	-0.03033	-0.101
Selenium	0.385091	0.002063	-0.07677	0.083407	0.119227	-0.139	-0.02088	0.049042	0.31287	-0.15799

	PC11	**PC12**	**PC13**	**PC14**	**PC15**	**PC16**	**PC17**	**PC18**	**PC19**
% Variance Explained	2.38%	1.64%	1.32%	0.53%	0.51%	0.36%	0.01%	0.01%	0.00%
pH	-0.2958	-0.36801	-0.22504	-0.12758	0.041497	-0.06125	0.003752	-0.00068	6.59E-17
Conductivity	0.179268	0.057028	0.194978	-0.20541	0.062851	-0.11292	0.006924	0.016221	1.06E-17
Alkalinity	-0.18807	0.119811	0.255502	0.376123	0.09456	0.199518	-0.02053	-0.01901	3.35E-17
Chloride	0.055677	0.080211	-0.05315	-0.06908	-0.02098	-0.04974	0.000878	0.006822	-7.07E-01
Total Hardness	0.055677	0.080211	-0.05315	-0.06908	-0.02098	-0.04974	0.000878	0.006822	7.07E-01
Nitrate	0.13119	-0.16735	0.012868	0.046439	-0.02848	0.023329	0.152529	-0.69151	2.81E-16
Nitrite	0.10176	-0.16409	0.013315	0.02344	-0.01384	0.076272	-0.14637	0.688778	8.74E-17
Phosphate	0.290459	-0.4003	0.589718	-0.16703	-0.15034	-0.0366	0.021718	0.01119	3.48E-16
Ammonium	0.429404	0.032559	-0.61859	-0.06523	0.199342	0.121461	0.004774	0.002298	-2.33E-16
Fluoride	-0.10563	0.647838	0.0731	0.063403	-0.07969	-0.06678	-0.00977	-0.00267	-2.02E-16
Manganese	-0.5341	-0.29256	-0.12127	0.238898	-0.01279	-0.00748	-0.01443	-0.00562	-1.33E-16
Sulphate	0.059205	-0.08253	0.004505	-0.1649	0.158238	-0.08604	-0.00404	0.002543	-5.59E-17

Cadmium	-0.0888	0.012144	0.006256	-0.34151	-0.07822	0.429336	-0.65626	-0.1671	1.32E-15
Copper	0.303009	-0.10775	-0.10605	0.609741	-0.45794	-0.0533	-0.0603	0.016379	2.35E-16
Iron	-0.09523	0.031739	-0.0109	-0.24961	-0.1212	0.398178	0.720479	0.134413	-8.21E-16
Lead	0.12867	-0.06663	0.132285	0.209133	0.763505	-0.17748	0.016091	0.006322	-8.77E-17
Zinc	-0.05779	0.285816	0.225258	0.00041	0.06298	-0.00505	-0.01546	0.00476	-5.39E-17
Arsenic	0.304383	0.069129	-0.00144	-0.0487	-0.15355	-0.01366	-0.00552	-0.00022	-5.07E-17
Selenium	-0.14572	0.04583	-0.09964	-0.26561	-0.22048	-0.72591	-0.00864	0.013464	-2.99E-16

Appendix D

Sampling Protocols and Field Data Sheet

Water Sampling Protocols

1. Surface water samples can be collected by directly filling the container from the surface water body being sampled. Avoid stirring up sediment when collecting samples.
2. During borehole water sample collection, if collecting water from a pump, make sure that the pump does not come in contact with the sample containers.
3. Label each container with sample ID, sample type (metals, nutrients, major ions), and preservation technique (if any).
4. Document all field observations and any deviations from standard sampling procedures on the field data sheet.

Notes:

1. Sample bottles must be new and certified-clean for collection of samples for trace metals or organic compound analyses. Cleaned bottles may be used for major ions.
2. All samples requiring preservation must be preserved as soon as practically possible, ideally immediately at the time of sample collection.
3. All samples preserved using pH adjustment must be checked, using pH strips, to ensure that they were adequately preserved. This is done by pouring a small volume of sample over the strip. Do not place the strip in the sample. Samples requiring reduced temperature storage should be placed on ice immediately.

Specific Sampling Requirements: (http://www.uga.edu/sisbl/epatab1.html)

Major Anions/Sulfate: collect 500 mL, keep cool and analyze within 28 days.

Ammonia/Nitrate/Nitrite/Phosphorus: collect 500 mL, acidify with H2SO4 to pH 2, keep cool and analyze within 28 days.

Major Cations/Metals: collect 100 mL, acidify with HNO3 to pH 2 and analyze within 6 months. **Do not use metal or rubber components for trace-element sampling.**

Quality Assurance/Quality Control

To assure quality data, we must collect a series of additional samples.

Trip Blanks

A trip blank is a bottle of blank water that is filled **before** sampling begins and travels with the samples during collection and shipment to the lab. It quantifies the bias introduced by contamination during sampling and transportation. Fill **two** bottles with blank water in the lab and label as a Trip Blank. Keep these samples with the other collected in the field, carrying them with you until the end of the sampling period.

Duplicate Samples

Duplicate samples identify and/or quantify the variability in the sampling and analysis system. Environmental samples collected in duplicate are considered identical or nearly identical in composition and are analyzed for the same chemical properties. Field personnel should be careful to keep detailed notes on exactly how the duplicate samples were collected and processed, to help distinguish the sources of variability that affected the samples.

Concurrent Duplicates

Concurrent duplicates are two samples of environmental water that are collected simultaneously or at approximately the same time. Concurrent duplicates should be taken for surface and ground water samples **every 10th sample.** 1.) Use identical sampling procedures and supplies, collecting the sample for each analysis one after the other, and 2.) Preserve the duplicate-sample set in the same order as the order in which the samples were collected.

Split Duplicates

Split duplicates are samples obtained by dividing one sample that is designated for a specific laboratory analysis into two subsamples (duplicates). Split replicates can determine variability from field plus laboratory processes (the field-replicate split sample) or from laboratory procedures (the lab-replicate split sample). **Split every 10th sample before processing in the lab.**

From http://www.epa.gov/region4/sesd/fbqstp/Surfacewater-Sampling.pdf and http://water.usgs.gov/owq/FieldManual/chapter4/pdf/Chap4_v2.pdf

Bottles were purchased from U.S. Plastics: http://www.usplastic.com/catalog/item.aspx?itemid=23905&catid=897

Water Sample Data Collection Sheet

Community: ______________________ Date: ___________ Time: _______________
Sampler's Name: ___________________________

Sampling Location: Borehole Well Stream Pond Galamsey Pit River Swamp Pipe
Spring BU Hot Spot

Latitude: __________ Longitude: _________ Elevation: ____________

Weather: sunny overcast raining

Air temperature(°C): _______________ Photograph taken? yes no
Image number: ________________

Field Measurements

Water temperature (°C): _____________ pH: ____________
DO (mg/L):_____________ Conductivity (uS/cm): ____________
Turbidity (cm): ______________ Hardness (test strip): ____________

Duplicate Collected? Yes No (Fill out separate form for duplicate sample)

Field Blank with Sample? Yes No

Major Ions

Use one **500 mL** bottle.

Metals

Use one **100 mL** bottle.

Add 15 drops Nitric Acid.

Nitrogen/Phosphorus

Use one **500 mL** bottle.

Add 25 drops Sulfuric Acid.

Comments:

Appendix E

Rainfall Data Analysis

A working hypothesis for the incidence of Buruli ulcer revolves around the prevalence of flooded areas in and near communities. Flooded areas are, by definition, transient and associated with rainfall patterns. In an effort to find a correlation between flooding and incidence of Buruli ulcer, monthly rainfall patterns were compared between gauging stations throughout Ghana. These data were provided by the Ghana Meteorological Service and spans Buruli-endemic and non-Buruli-endemic communities.

The analysis presented here was conducted by fast Fourier transform (FFT). FFT accepts time-series data and extracts dominant frequency signals. Consider $y=\sin(2x)$: we know that this sinusoidal equation has a period of π. If data from $y=\sin(2x)$ were analyzed by FFT, the output would be a dominant frequency of $\pm 1/\pi$, as shown in Figure E-1.

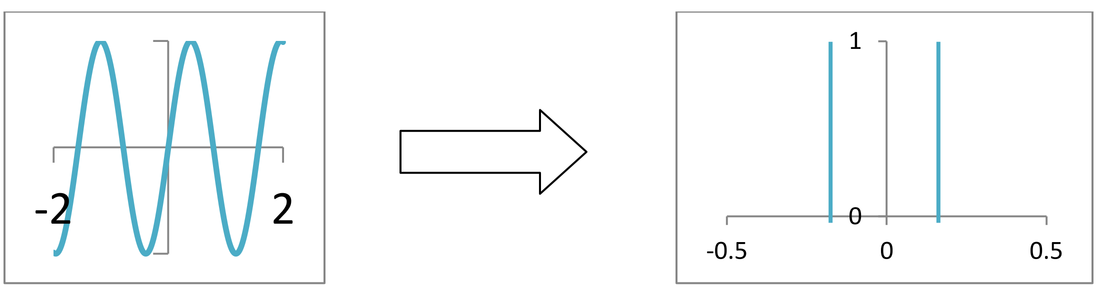


Figure E-1. Conversion of periodic data to dominant frequencies by FFT.

The vertical axis on FFT plots represents the relative strength of the signal, or the proportion of data that aligns at that frequency. When input data is non-ideal, the frequencies reflect complexities by way of multiple frequencies of varying strength. In the context of rainfall, we can see climatic patterns in rainfall at each gauging station and detect differences between these patterns. If there are differences in these patterns, we can then compare this to known incidences of Buruli ulcer to determine a possible relationship between climatic rainfall and incidence of the disease.

Figure E-2 shows the location of seven gauging stations operated by the Ghana Meteorological Service (Axim, Dunkwa, Enchi, Kumasi, Navrongo, Sefwi Bekwai, and Tarkwa). Each of these sites has complete records for 30-50 years of monthly rainfall totals.

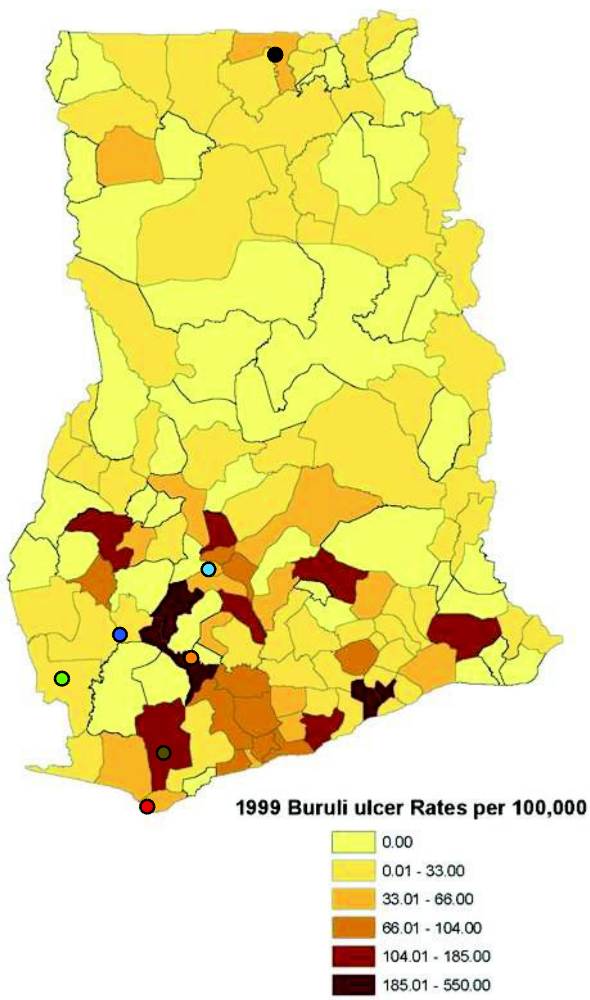
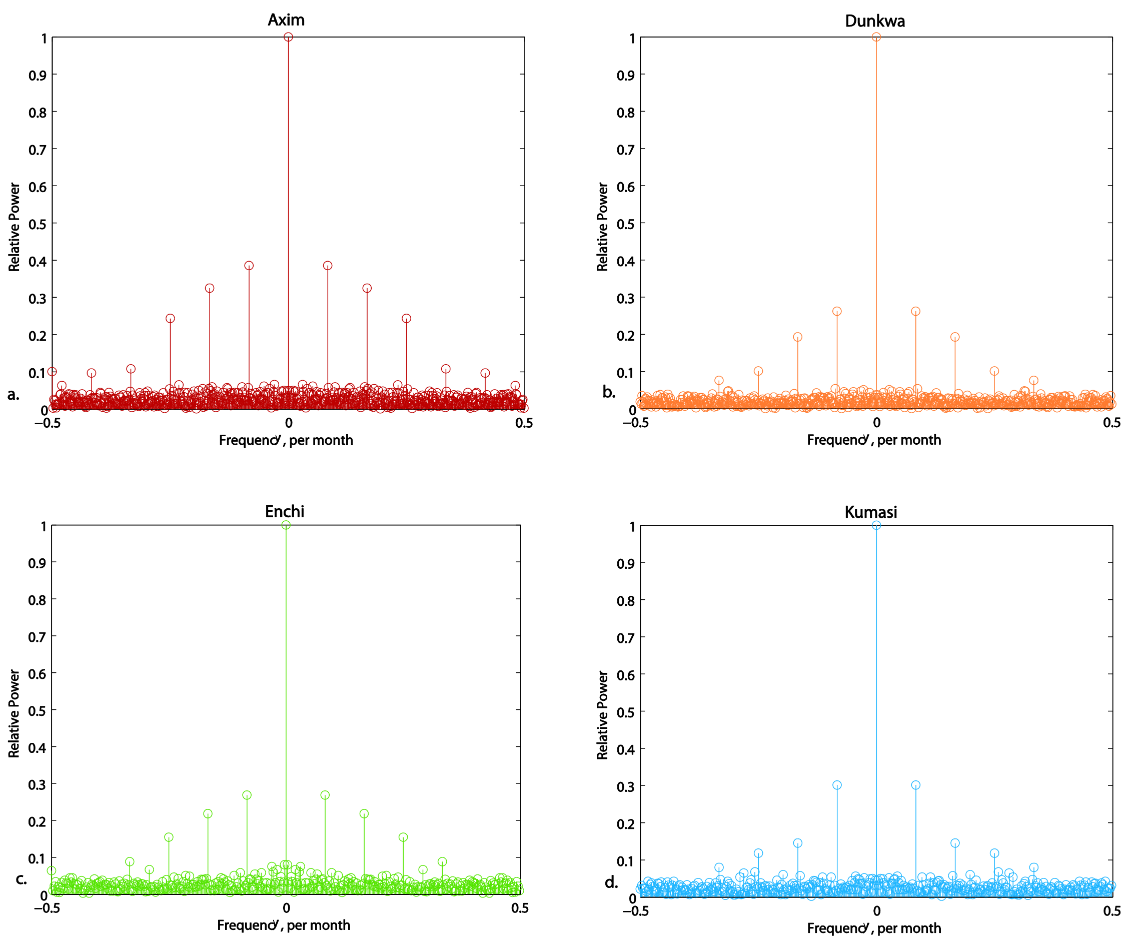


Figure E-2. Buruli ulcer incidences by district, 1999, from David Ferring (University of North Texas). Colored dots indicate gauging stations.

As seen in Figure E-3, there is very little variation between gauging stations. The one acception to this is Figure E-3g, Navrongo, which is located in arid northern Ghana. As the climate is substantially different here, it is logical that Navrongo should have a different climatic pattern in FFT as well. Considering only the southern sites, climate is very similar throughout endemic and non-endemic areas. This suggests that climatic rainfall patterns do not affect incidence of Buruli ulcer. However, individual rainfall events may affect incidence.



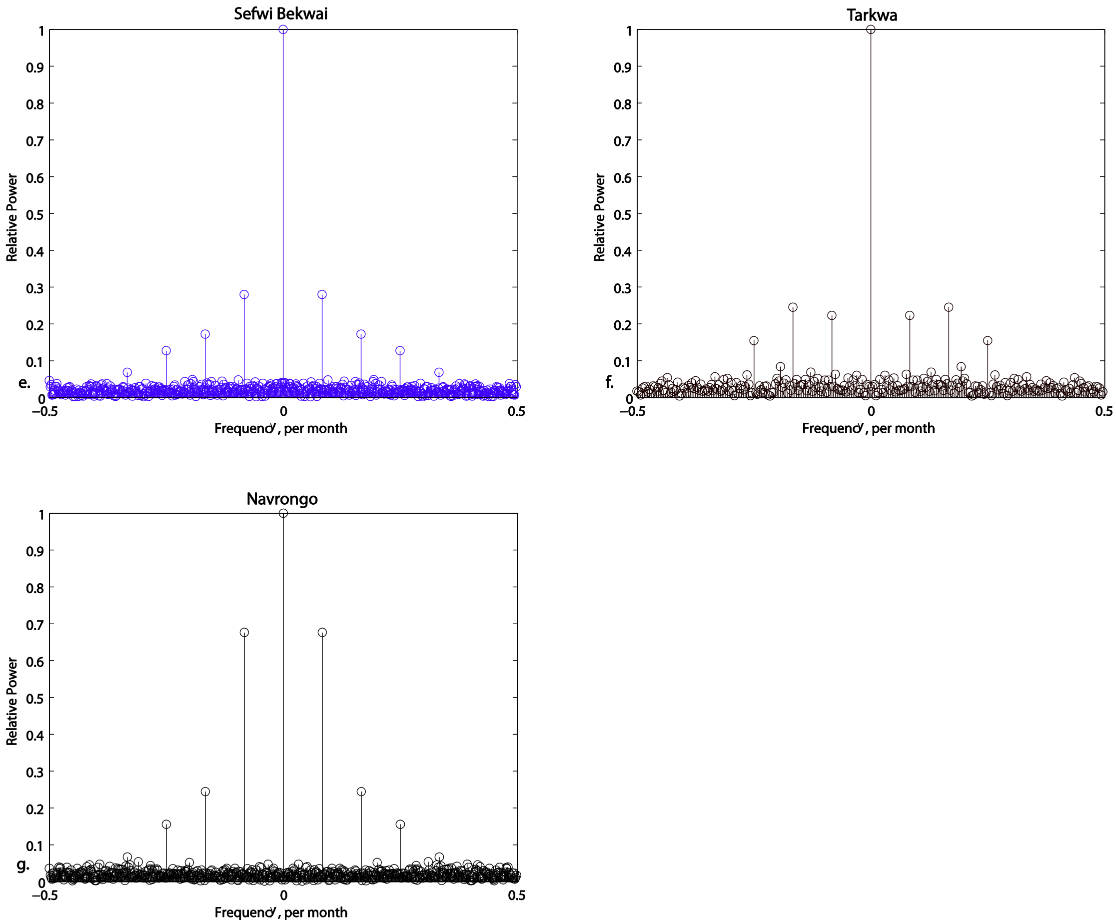


Figure E-3. FFTs of monthly rainfall at seven gauging stations: a) Axim, b) Dunkwa, c) Enchi, d) Kumasi, e) Sefwi Bekwai, f) Tarkwa, g) Navrongo. Colors correspond to map locations in Figure E-2.

Appendix F

January 2011 Data Analysis

In January 2011, samples were collected from all eleven locations listed in Table 1. The following analyses are presented here, rather than in the main body of this text, because sampling locations and field and laboratory methods differ between these samples and those of June-July 2011. Forty-six water samples were collected in two acid-washed 1.5 liter drinking water bottles. These water samples were not acidified and were stored in coolers during filed work and transportation to KNUST for analysis. Samples were analyzed in a manner similar to that described in the Field Methods section of this document. Larger quantities of water were collected in each sample, and all trace metal analysis was performed at KNUST using ICP-AAS. Resulting data (shown in tabular format in Appendix G) were analyzed by PCA, as described in the Statistical Methods section.

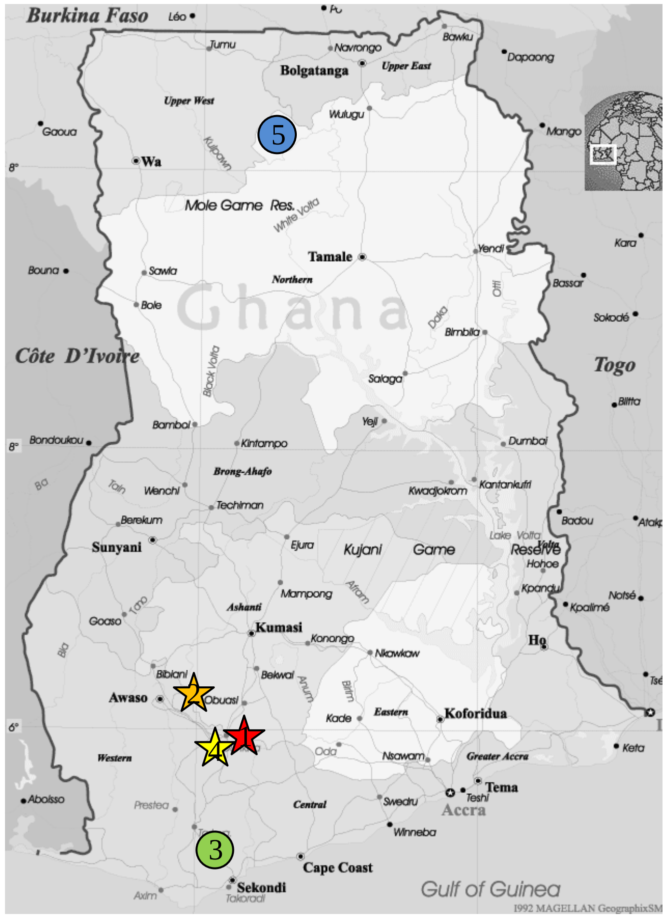


Figure F-1. Map of Ghana showing sampling locations from January 2011 sampling. Stars indicate Buruli-endemic areas, while circles indicate non-endemic communities.

A biplot of these data shows the contribution of chemical variables to the principal components (Figure F-2). The analyses here consider only two principal components, PC1 and PC2. These components explain only 42.1% of the total variance in the data, but they are the most significant components. PC1 is controlled largely by calcium, magnesium, and total hardness, while PC2 is dominated by iron, arsenic, alkalinity, phosphate, and pH.

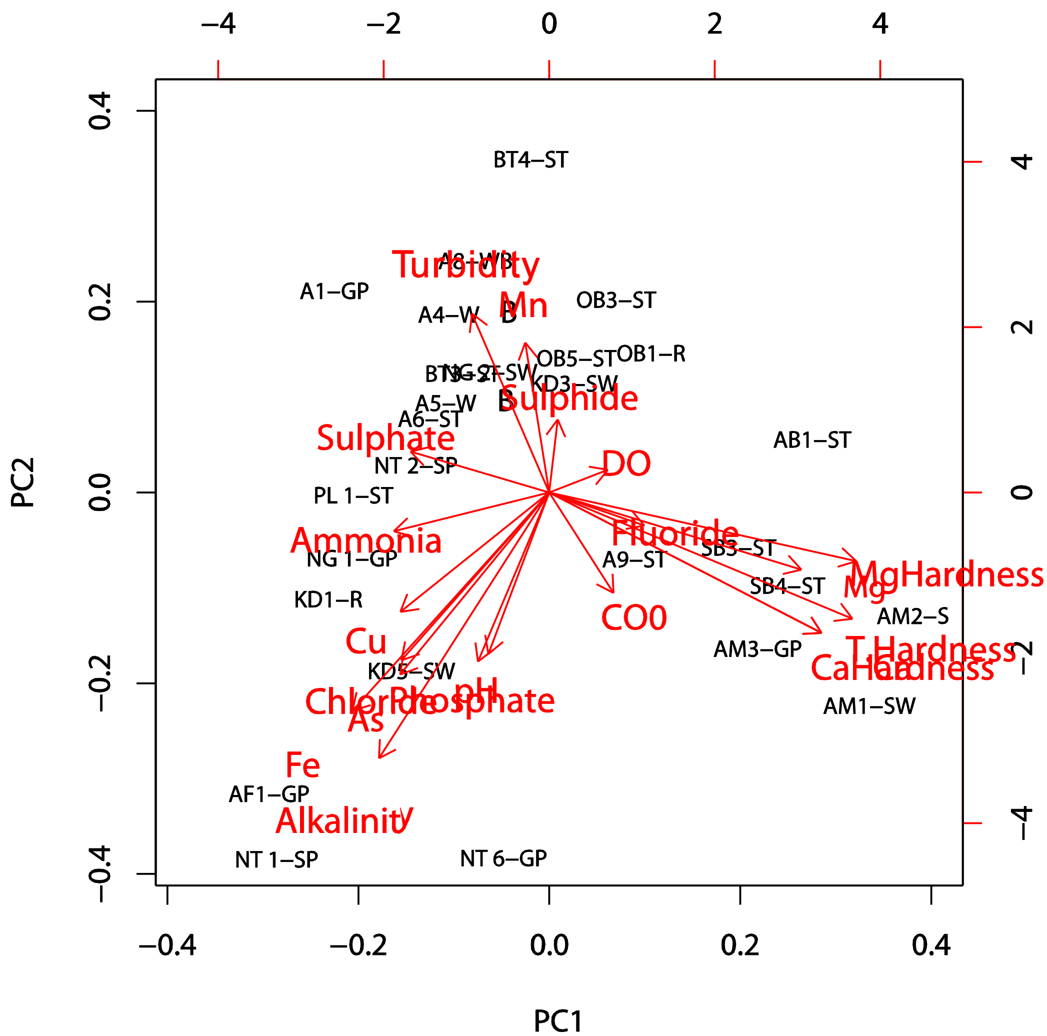


Figure F-2. Biplot of PCA data. PC1 is controlled by calcium, magnesium, and total hardness; PC2 is dominated by iron, arsenic, alkalinity, phosphate, and pH.

PCA coded by community, as presented in Figure F-3, shows some differentiation between communities. However, there is no such differentiation between Buruli-endemic and non-Buruli-endemic communities.

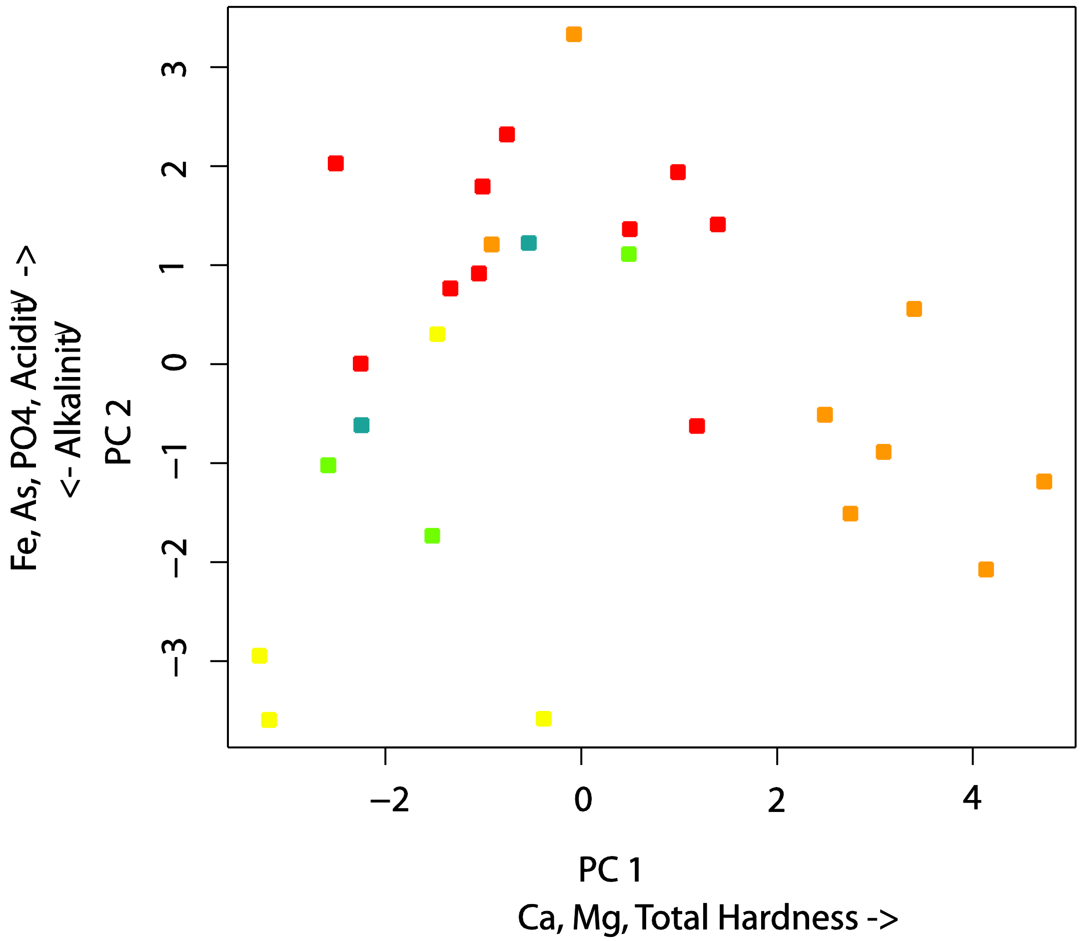


Figure F-3. PCA coded by study area. Colors correspond to Figure E-1. PC1 is controlled largely by calcium, magnesium, and total hardness, while PC2 is dominated by iron, arsenic, alkalinity, phosphate, and pH.

When coded by type of water body, the PCA results show a division between two groups of water bodies: fast-moving and/or well-connected water bodies plot significantly differently from slow-moving, poorly-connected, and/or stagnant water bodies (Figure F-4). This division is driven primarily by PC2.

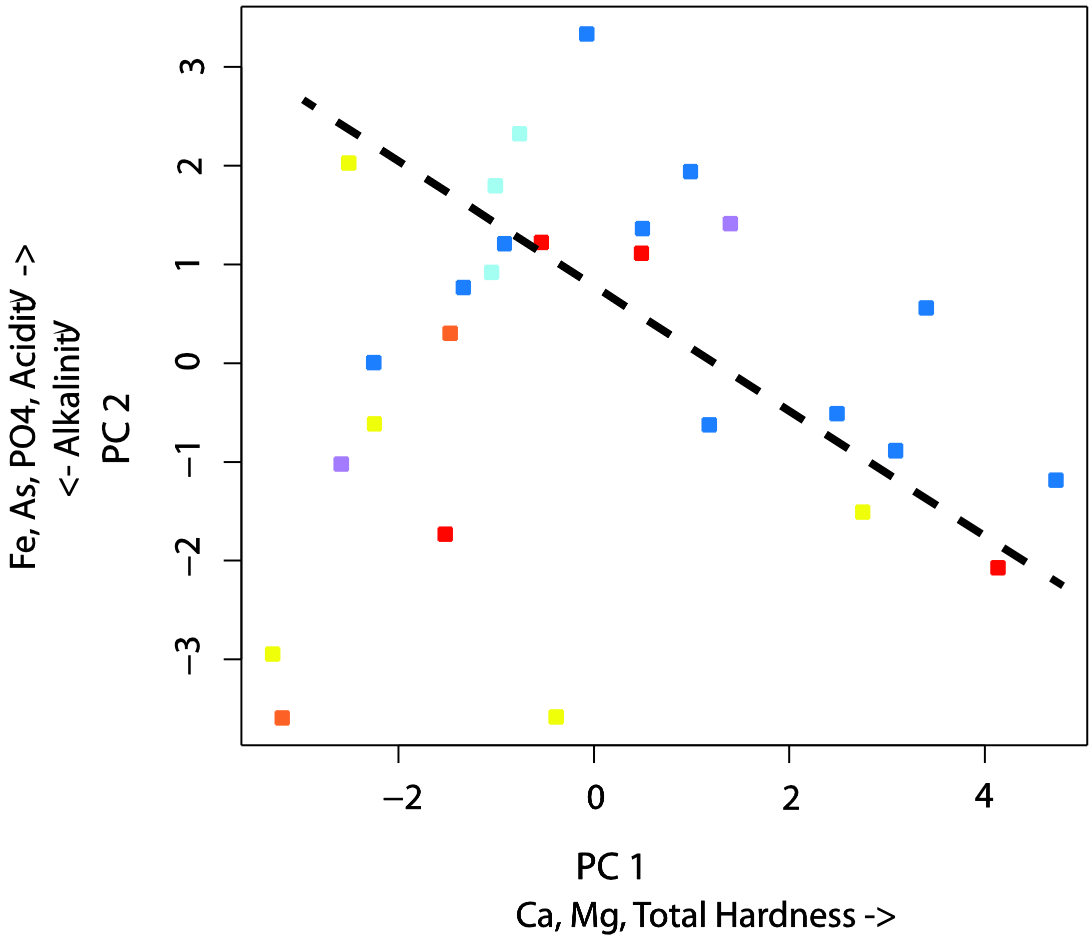


Figure F-4. PCA coded by type of water body. Dark blue represents streams, light blue ponds/lakes, purple rivers, red swamps, orange springs, and yellow galamsey pits. PC1 is controlled largely by calcium, magnesium, and total hardness, while PC2 is dominated by iron, arsenic, alkalinity, phosphate, and pH.

Interestingly, the factors that control PC2 agree with the expected environmental niche for *Mycobacterium ulcerans* as outlined by Iivanainen et al. (1993). These results suggest that, in these communities, *M. ulcerans* may proliferate in galamsey pits, spring outlets, and swamps. While these data show a distinct difference between water bodies, no such difference is found in the data presented in the main body of this thesis. This may be an artifact of changes in field and laboratory methods or changes in specific sampling locations. The difference could also be an effect of time; water bodies could be chemically different between dry and rainy seasons.

Appendix G

January 2011 Data

pH is presented in standard units, and conductivity is reported in uS/cm. All other data are reported in mg/L.

Sample coding convention is community code, sample number-water body type. Codes for communities are as follows: A represents Pokukrom, OB Obiaradaneden, PL Powerline, BT Betenase, SB Subin, AM Ameyaw, AB Ampabena, KD Kedadwen, AF Ayanfuri, and NT Nkotumso. Water body types are coded as GP for galamsey pit, BH for borehole, W for well, WB for water body (pond), ST for stream, R for river, SW for swamp, and SP for spring.

Code	pH	DO	Cond	Turb	SO4	S	PO4	F	NH4	Alk	Cl	T.H.	CaH	MgH	Ca	Mg	Mn	As	Fe	Cu
A1-GP	6.80	3.16	124.90	50.60	120.00	0.10	5.60	0.00	0.26	40.00	2.66	70.00	45.36	24.64	0.540	0.160	0.024	0.004	1.783	0.024
A2-BH	5.60	1.80	137.20	0.20	3.00	0.06	4.10	0.40	0.00	32.50	2.84	85.00	51.74	33.26	0.616	0.234	0.021	0.001	0.539	0.036
A3-W	4.56	3.44	324.00	0.90	3.00	0.06	0.00	0.35	0.00	10.00	2.84	55.00	30.66	24.34	0.365	0.185	0.021	0.000	0.040	0.018
A4-WB	5.92	2.40	103.60	46.18	25.00	0.06	2.10	0.00	0.01	32.50	2.49	56.50	33.60	22.90	0.400	0.165	0.006	0.011	2.679	0.021
A5-WB	5.87	2.36	101.30	41.61	19.00	0.06	15.00	0.00	0.00	40.00	1.78	59.50	25.20	34.30	0.300	0.295	0.004	0.008	2.909	0.092
A6-ST	6.47	2.78	135.00	56.00	20.00	0.06	19.00	0.10	0.04	50.00	1.60	60.00	37.80	22.20	0.450	0.150	0.004	0.010	2.248	0.021
A7-BH	5.93	2.80	261.60	16.32	20.00	0.07	12.00	0.60	0.00	52.50	2.84	150.00	96.60	53.40	1.150	0.350	0.011	0.001	0.993	0.027
A8-WB	6.37	2.47	40.30	100.00	87.00	0.12	18.00	0.00	0.00	27.50	1.78	112.00	63.00	49.00	0.750	0.370	0.018	0.005	3.039	0.028
A9-ST	6.52	2.93	319.00	64.00	22.00	0.08	20.00	0.00	0.05	47.50	2.49	184.00	100.80	83.20	1.200	0.640	0.003	0.011	2.384	0.018
OB1-R	5.16	4.02	356.00	20.81	7.00	0.07	0.00	0.00	0.00	15.00	1.78	132.50	79.80	52.70	0.950	0.375	0.010	0.000	1.325	0.012
OB2-W	5.26	1.82	134.20	3.40	0.00	0.07	1.80	0.30	0.00	30.00	1.24	105.00	71.40	33.60	0.850	0.200	0.021	0.000	0.187	0.028
OB3-ST	5.07	2.93	110.70	8.95	3.00	0.07	2.90	1.15	0.00	20.00	2.84	110.00	67.20	42.80	0.800	0.300	0.018	0.000	0.000	0.006
OB4-W	5.85	2.42	115.70	18.80	12.00	0.07	6.40	0.00	0.00	20.00	2.84	75.00	50.40	24.60	0.600	0.150	0.005	0.001	0.002	0.023

OB5-ST	5.67	1.66	52.50	17.43	0.00	0.05	0.70	0.00	0.00	25.00	1.78	105.00	54.60	50.40	0.650	0.400	0.003	0.003	2.017	0.020
PL 1-ST	5.41	0.44	68.60	11.31	17.00	0.09	14.00	0.50	0.01	76.00	12.80	24.56	7.56	17.00	0.090	0.170	0.008	0.031	5.061	0.014
BT2-BH	5.71	1.51	61.60	0.00	0.00	0.60	0.70	0.20	0.00	20.00	2.13	65.00	46.20	18.80	0.550	0.100	0.030	0.000	0.044	0.052
BT3-ST	6.13	0.96	28.60	26.26	13.00	0.07	18.00	0.20	0.00	30.00	2.31	61.50	37.80	23.70	0.450	0.165	0.006	0.002	2.423	0.023
BT4-ST	4.90	1.25	31.60	83.00	0.00	0.06	0.00	0.60	0.00	10.00	1.95	80.00	54.60	25.40	0.650	0.150	0.011	0.000	0.000	0.004
SB1-BH	5.86	2.05	198.30	24.29	14.00	0.13	21.00	0.10	0.01	90.00	20.20	26.56	7.56	19.00	0.090	0.190	0.022	0.219	0.611	0.027
SB2-BH	6.15	1.42	61.90	36.10	10.00	0.07	19.00	0.05	0.00	62.50	4.26	164.50	117.60	46.90	1.400	0.245	0.004	0.009	2.008	0.011
SB3-ST	6.54	1.42	61.90	39.32	16.00	0.06	12.00	0.65	0.00	35.00	3.20	256.00	162.96	93.04	1.940	0.620	0.005	0.005	2.146	0.023
SB4-ST	6.54	1.14	54.60	19.48	16.00	0.06	16.00	0.60	0.00	20.00	1.60	320.00	243.60	76.40	2.900	0.300	0.005	0.004	1.407	0.007
AM1-SW	6.93	3.49	135.20	0.00	14.00	0.03	15.00	0.65	0.00	40.00	1.60	380.00	292.32	87.68	3.480	0.320	0.005	0.000	0.014	0.014
AM2-ST	6.46	3.57	179.40	7.50	7.00	0.08	2.90	0.10	0.00	35.00	0.71	342.00	203.28	138.72	2.420	1.000	0.007	0.008	2.271	0.010
AM3-GP	6.22	3.14	179.40	0.00	3.00	0.08	19.00	1.40	0.00	55.00	2.49	228.00	127.68	100.32	1.520	0.760	0.006	0.017	0.293	0.024
AM4-BH	6.60	6.58	88.00	16.10	13.00	0.05	4.10	0.00	0.00	30.00	1.78	290.00	181.44	108.56	2.160	0.740	0.006	0.000	0.000	0.037
AM5-BH	5.82	1.99	49.30	2.72	0.00	0.08	2.90	0.65	0.00	15.00	2.31	284.00	146.16	137.84	1.740	1.100	0.024	0.000	0.045	0.017
AB1-ST	5.16	1.38	67.40	21.01	3.00	0.05	3.90	1.10	0.00	30.00	1.42	266.00	162.96	103.04	1.940	0.720	0.021	0.002	2.118	0.016
AB2-BH	5.04	1.25	62.70	1.81	10.00	0.10	15.00	0.70	0.00	20.00	1.42	314.00	208.32	105.68	2.480	0.660	0.002	0.000	0.036	0.018
AB3 BH	5.91	2.2	136.1	33.45	13.00	0.11	21.00	0.40	0.21	80.00	10.10	46.96	24.00	22.96	0.280	0.240	0.007	0.102	0.021	0.014
KD1-R	7.30	4.54	91.10	15.24	24.00	0.05	12.00	0.00	0.03	90.00	1.80	43.92	31.92	12.00	0.380	0.120	0.008	0.011	22.633	0.092
KD3-SW	6.18	2.12	80.80	21.72	50.00	0.08	11.00	0.60	0.01	47.00	1.42	173.40	155.40	18.00	1.850	0.180	0.030	0.005	1.052	0.009
KD4-BH	5.63	0.66	300.00	14.81	16.00	0.06	7.50	0.20	0.01	65.00	6.03	283.52	233.52	50.00	2.780	0.500	0.016	0.023	17.026	0.196
KD5-SW	6.12	0.32	97.70	19.59	14.00	0.03	7.50	0.55	0.06	47.00	2.13	136.32	124.32	12.00	1.480	0.120	0.009	0.072	20.553	0.040
KD6-BH	5.47	1.12	186.10	30.34	24.00	0.01	14.00	0.60	0.05	47.00	4.97	85.40	71.40	14.00	0.850	0.140	0.016	0.034	0.460	0.028
AF1-GP	6.26	0.12	98.40	12.31	52.00	0.07	6.40	1.10	0.08	125.00	5.30	90.76	74.76	16.00	0.890	0.160	0.003	0.031	24.362	0.296
AF 2-W	5.74	0.74	108.50	33.10	24.00	0.10	16.00	0.60	0.01	90.00	14.90	171.16	146.16	25.00	1.740	0.250	0.015	0.006	2.121	0.010
AF 3-W	5.38	3.90	179.00	21.50	19.00	0.10	18.00	0.10	0.06	85.00	27.70	30.36	3.36	27.00	0.040	0.270	0.012	0.034	5.216	0.017
AF 4-W	5.58	0.37	80.60	4.10	20.00	0.14	44.00	0.05	0.06	168.00	15.10	27.12	15.12	12.00	0.180	0.120	0.030	0.013	2.711	0.009
AF 5-W	5.14	2.70	62.90	7.71	17.00	0.09	27.00	0.60	0.01	52.00	12.78	57.36	45.36	12.00	0.540	0.120	0.017	0.017	0.000	0.003
NT 1-SP	6.95	0.23	178.20	11.27	26.00	0.01	26.00	0.05	0.08	117.00	25.20	82.20	62.20	20.00	0.740	0.200	0.009	0.026	18.509	0.017
NT 2-SP	5.61	0.27	65.30	13.39	18.00	0.10	16.00	0.65	0.01	52.00	13.10	51.24	15.12	36.00	0.180	0.360	0.017	0.037	0.280	0.024
NT 3-BH	5.71	0.84	210.00	24.67	20.00	0.12	27.00	0.50	0.04	95.00	20.20	144.28	119.28	25.00	1.420	0.250	0.020	0.032	0.807	0.023

NT 4-BH	5.67	1.44	104.70	36.23	15.00	0.12	16.00	0.10	0.01	73.00	22.60	56.80	37.80	19.00	0.450	0.190	0.008	0.016	0.032	0.027
NT 5-W	5.85	1.39	53.70	22.13	16.00	0.13	18.00	0.05	0.04	30.00	12.80	317.12	309.12	8.00	3.680	0.080	0.028	0.020	16.996	0.044
NT 6-GP	6.84	3.89	202.70	32.24	26.00	0.13	19.00	0.05	0.05	135.00	12.80	203.52	149.52	54.00	1.780	0.540	0.000	0.034	17.525	0.020

Appendix H

WATSAN Data Analysis

The data presented here come from the Community Water and Sanitation Agency (WATSAN) in Ghana. These data are presented here because, while relevant, they were collected by a third party and do not correspond to sampling locations presented in the main body of this thesis. Many communities' primary source of drinking water is a borehole drilled by the Ghanaian government. Each borehole drilled for drinking water is sampled extensively before it is mad available for public use. These data are collected by region at the regional WATSAN office. The WATSAN data presented here represent two districts: Upper Denkyira, the location of some of my field sites, and Twifo Heman Lower Denkyira, an adjacent district. Upper Denkyira is an endemic district for Buruli Ulcer, while Twifo Heman Lower Denkyira has no reported cases of the disease. The water chemistry data collected encompass turbidity, pH, conductivity, total dissolved solids, major ions (sodium, potassium, etc.), nutrients (nitrate, phosphate, etc.), hardness (calcium, magnesium, etc.), and alkalinity (carbonate, bicarbonate, etc.), making 22 consistently measured variables.

Data in these analyses are overlain with aquifer rock type and geographic location. Aquifer rocks in this dataset are classified discretely as Phyllite, Granite, or Schist/Quartzite (Shown in Figure H-1). Geographic location is considered more subjectively. The sampling sites in this dataset are located in a roughly northwest-southeast orientation. These sites are situated in groups, as shown in Figure H-2.

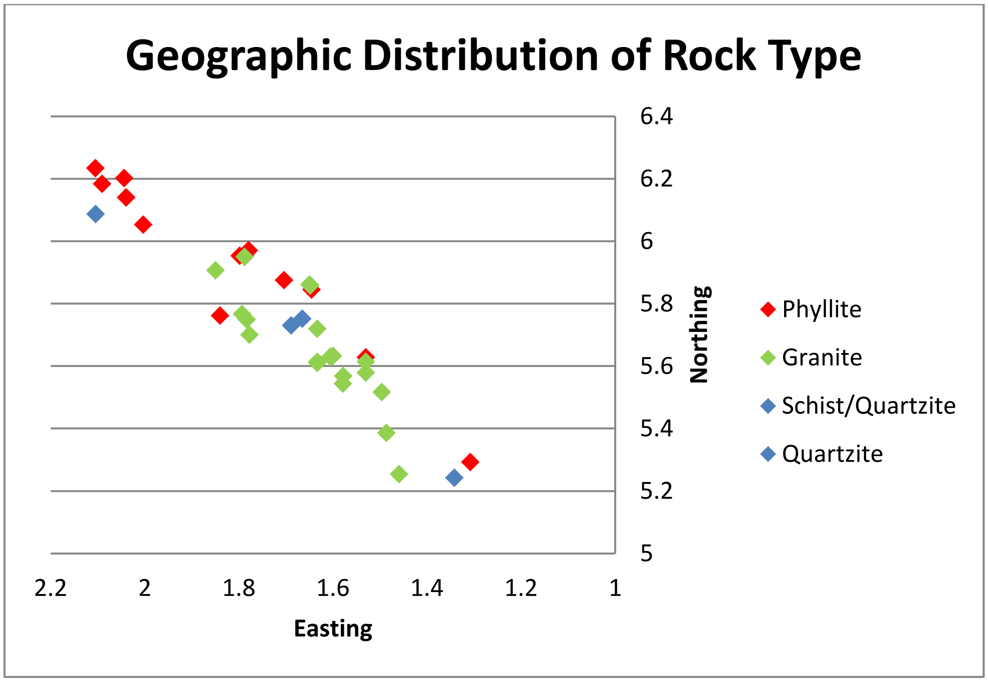


Figure H-1. Geographic distribution of aquifer rock type.

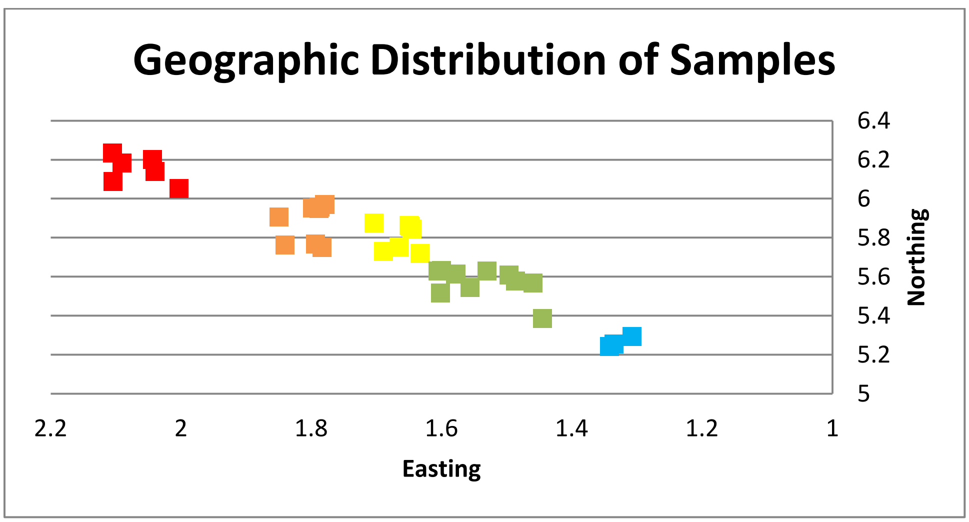


Figure H-2. Geographic distribution of samples.

Cluster Analysis

The dendrogram relating to rock type (Figure H-3) shows some relation between rock type and similarity of samples. Kyerepo, Ahwiawia, and Amobaka group together;

these samples are all composed of phyllite. Likewise, Burger's Estate Site B, Dunkwa Sec. Tech., District Assembly Guest House, and Dunkwa Sec. Sch. group together and are all phyllite. This trend is seen among some other small groupings within the dendrogram. Aquifer rock type, however, is likely not the primary explanation for the similarity among samples. Many small groupings are not organized by rock type, and no large groups appear to be associated with aquifer rock type.

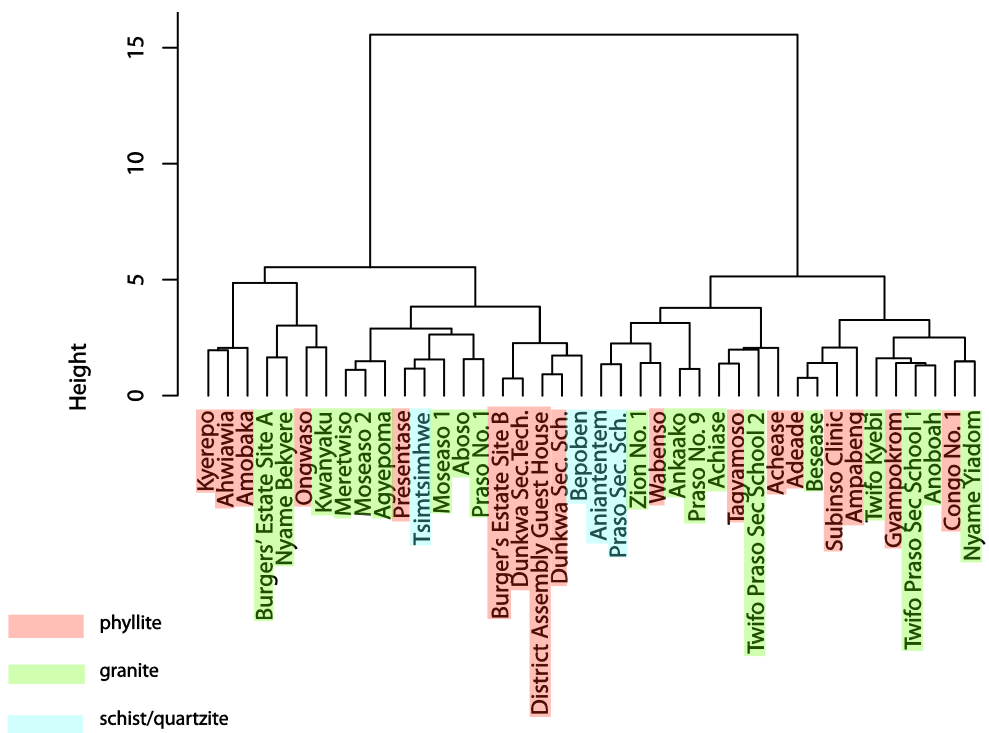


Figure H-3. Dendrogram representing similarity among WATSAN sampling sites with respect to aquifer rock type.

Figure H-4, showing the relation between geographic location and similarity among samples. Some of the same groups appear as did with respect to rock type: Kyerepo, Ahwiawia, and Amobaka group together, as do Adeade, Besease, Subinso Clinic, and Ampabeng. However, many small groups are not associated with geographic location.

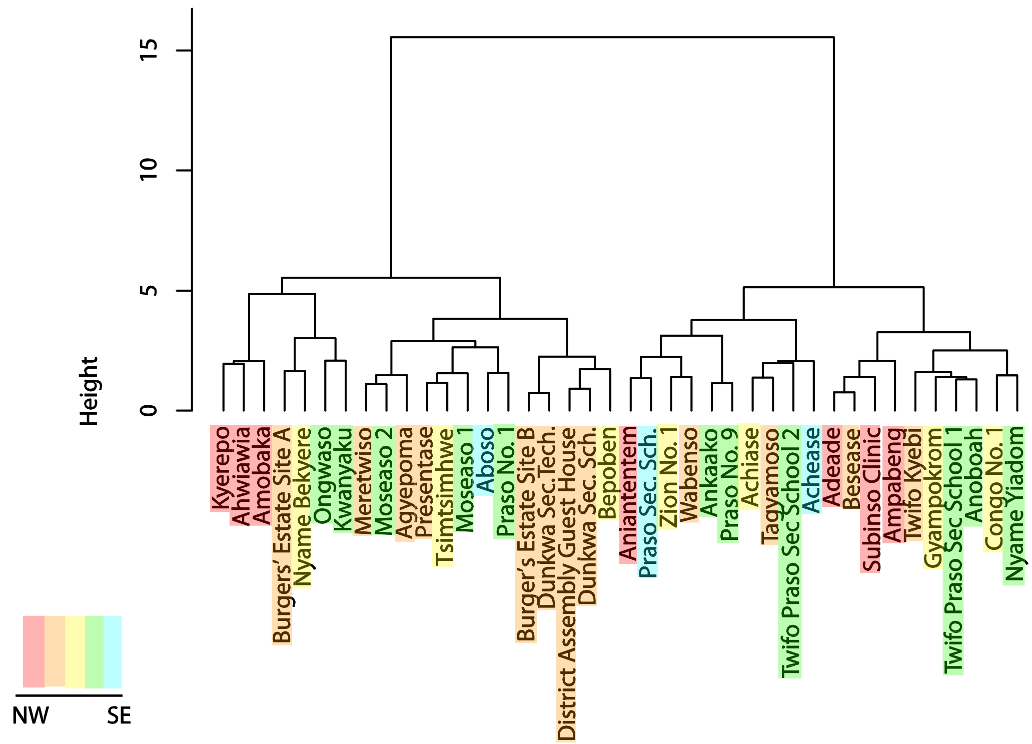


Figure H-4. Dendrogram representing similarity among WATSAN sampling sites with respect to geographic location.

Figure H-5 shows that groundwater samples from communities with reported cases of Buruli Ulcer are, in some way, similar to each other. There may be something about the water quality in Adeade, Besease, Subinso Clinic, and Ampabeng that makes Buruli Ulcer more common in these communities than others. These communities have higher turbidity and more intense color than the other communities considered in this study; they also have lower concentrations of total dissolved solids and each of the major ions, as well as lower values for conductivity, hardness, alkalinity, and pH. Acidic water may create a favorable environment for growth of *Mycobacteria*, and high turbidity and color values suggest a high suspended solids load that may relate to bacterial concentration. A potentially unfair assumption was made with Buruli Ulcer incidence data: Burger's Estate Sites A and B, Dunkwa Sec. Sch., Dunkwa Sec. Tech., and District

Assembly Guest House are all located within Dunkwa; as such, the number of reported of cases of Buruli Ulcer in Dunkwa (26 cases) was divided evenly among these locations.

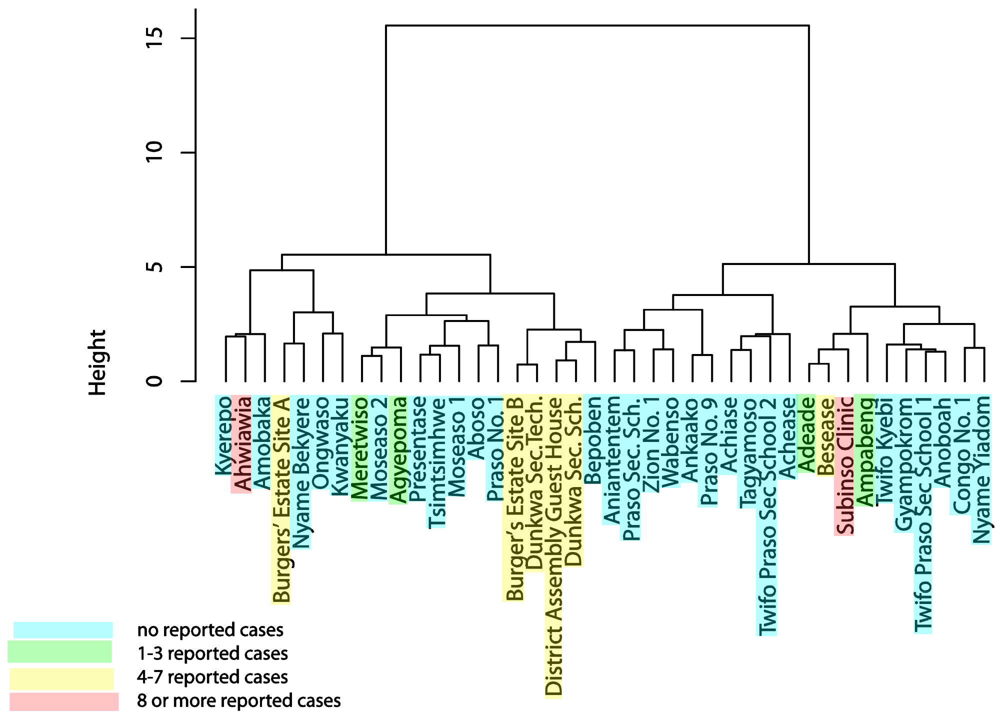


Figure H-5. Dendrogram representing similarity among WATSAN sampling sites with respect to incidence of Buruli Ulcer.

Clustering of variables is far more easily interpreted. Figure H-6 shows three distinct clusters. The cluster of Turb (turbidity), Col (color), SO_4, PO_4, Mn, NO_2, and NO_3 represents a collection of the major parameters responsible for and measuring cloudiness. Turbidity is a measure of the cloudiness of the water, which is typically due to suspended solids, a parameter not recorded consistently in this dataset. However, the grouping of sulfate, phosphate, nitrite, and nitrate in this cluster is logical, as the nutrients support the growth of microorganisms in some water systems. This microbial growth can contribute to suspended solids concentration. Color is often related to iron concentration, as iron complexes can affect the color of water; iron is, however, not consistently measured in this dataset. The presence of manganese in this cluster suggests that iron would be present in this cluster if it had been included in the analysis because iron and manganese concentrations in water are often closely related.

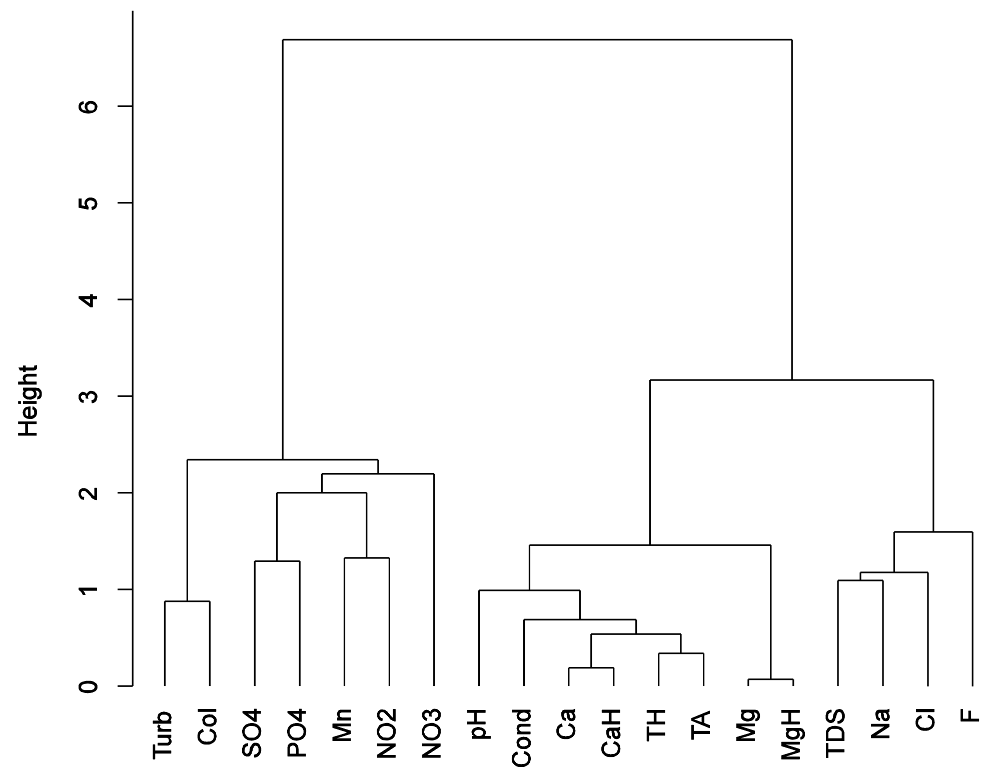


Figure H-6. Dendrogram representing similarity among responses of geochemical variables measured by Community Water and Sanitation Agency.

Another distinct cluster shown in Figure H-6 is comprised of TDS (total dissolved solids), F, Cl, and Na. Most of these elements should logically group together. Total dissolved solids is a measure of total ions in solution. This measure typically corresponds to F, Cl, and Na, as well as Ca and Mg. Ca and Mg do not appear in this cluster because they play a more important role in another grouping.

Mg, MgH (magnesium hardness), pH, Conductivity, Ca, CaH (calcium hardness), TH (total hardness), and TA (total alkalinity) make up the third distinct cluster in Figure H-6. These elements should be expected to group together. Magnesium and magnesium hardness are two ways of expressing magnesium concentration; thus they should be very closely related to each other. The same explanation applies for the close relationship between calcium and calcium hardness. Total hardness is an expression of the concentration of multivalent cations in solution and is, in essence, equal to the sum of calcium hardness and magnesium hardness. Thus, all of these elements should cluster together. Alkalinity, closely related to bicarbonate concentration, is a measure of the ability of water to neutralize acid. pH and total alkalinity should cluster together in natural water samples. Alkalinity and hardness are typically related to each other, as most multivalent cations are derived from carbonaceous rocks. When cations such as calcium and magnesium are released into solution, carbonate and bicarbonate are, as well.

Principal Components Analysis

Calcium, calcium hardness, magnesium, magnesium hardness, total hardness, and total alkalinity arrows seen in Figure H-7 point opposite the cluster of Phyllite samples concentrated on the right side of the plot in Figure H-8. Consistent with this, data shows that Phyllite samples had relatively low concentrations of these components. In Figure H-9, a similar trend is seen, with samples from the extreme northwest area of the data set clustering on the right of the plot. These data all have pH values less than the mean for the data set, agreeing with their location opposite the pH arrow shown in Figure H-7. Comparison between Figures H-7 and H-10 shows little correlation between any water quality parameter and incidence of Buruli Ulcer. It should be noted that the city of

Dunkwa had 26 reported cases of Buruli Ulcer between 2007 and 2010. As Burger's Estate A, Burger's Estate B, Dunkwa Secondary School, Dunkwa Secondary Tech., and District Assembly Guest House are all spatially co-located with the city of Dunkwa; cases of disease were assumed equally distributed among these sites. This assumption may be unfair, but it does not appear to affect the findings of cluster analysis and PCA in this dataset.

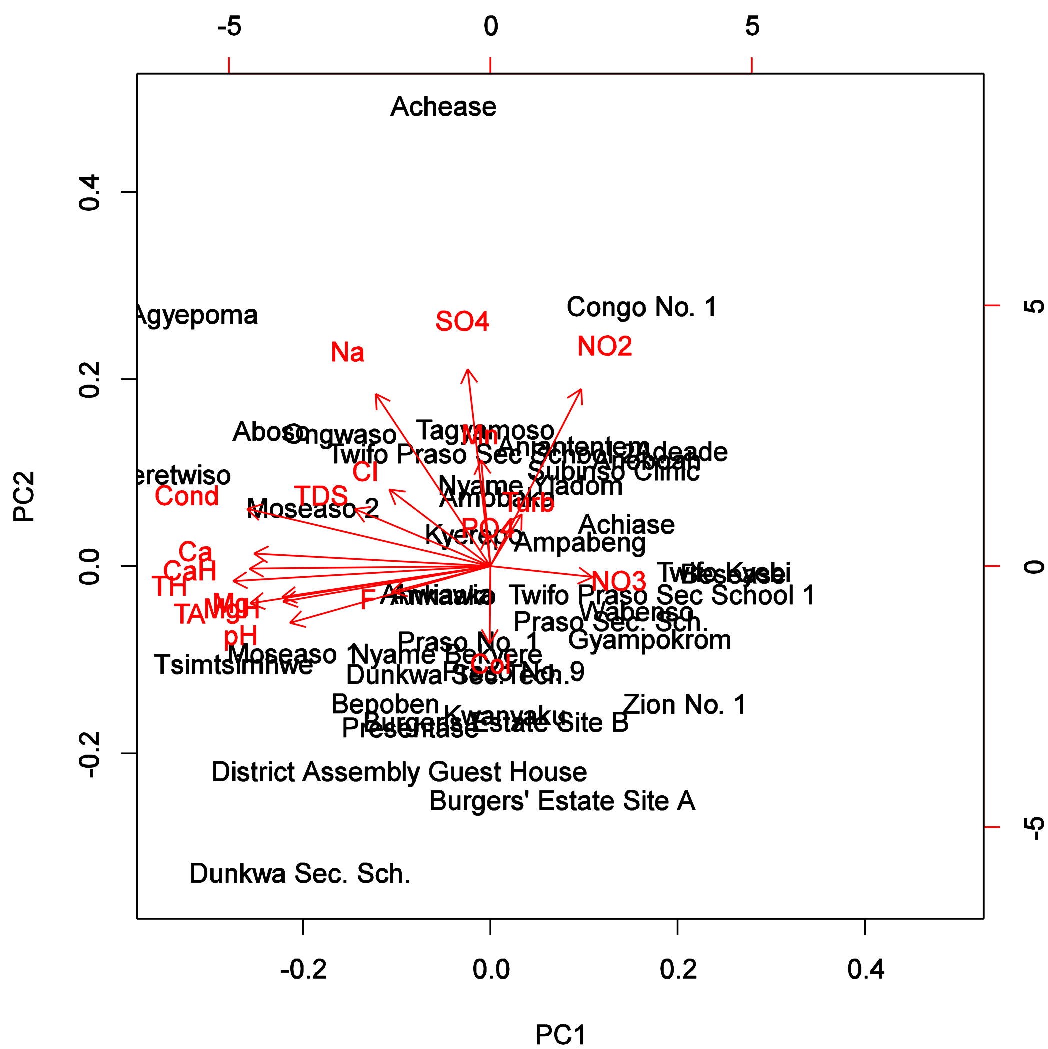


Figure H-7. Biplot of sample scores and variable loading.

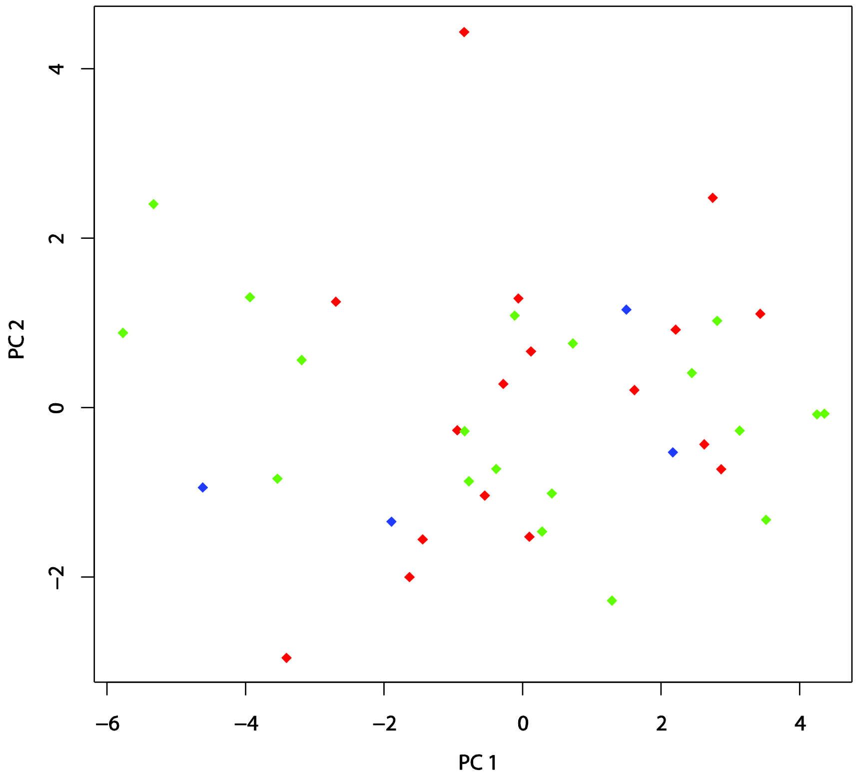


Figure H-8. PCA coded by aquifer rock type. Principal component 1 is controlled by hardness and alkalinity, while principal component 2 is controlled by turbidity, color, and iron.

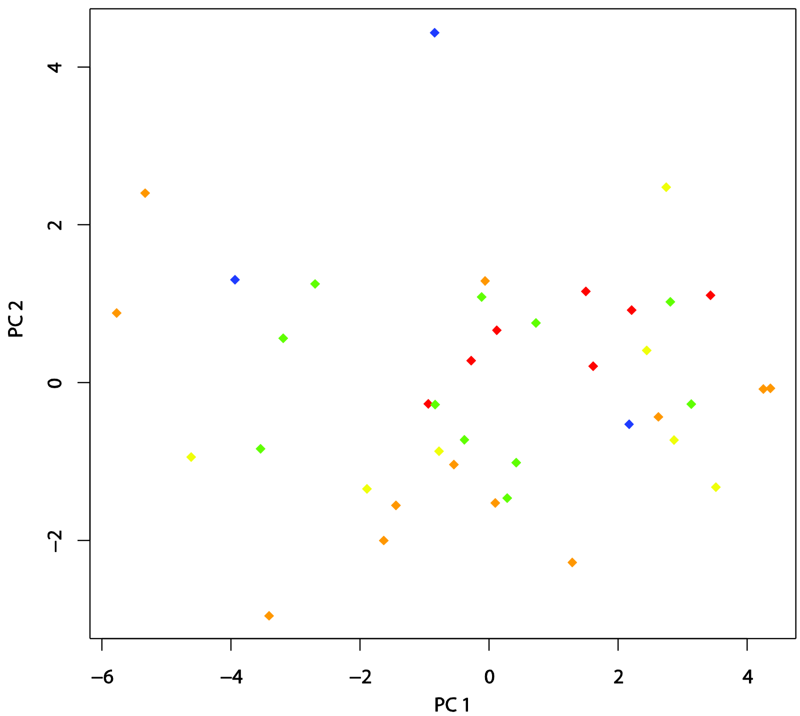


Figure H-9. PCA coded by geographic location. Principal component 1 is controlled by hardness and alkalinity, while principal component 2 is controlled by turbidity, color, and iron.

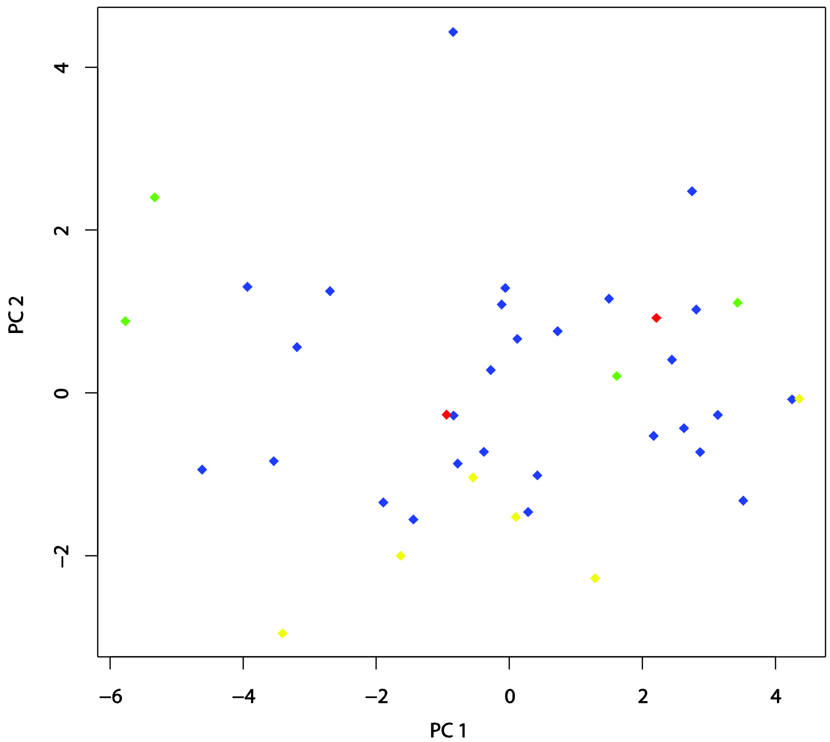


Figure H-10. PCA coded by incidence of Buruli Ulcer. Principal component 1 is controlled by hardness and alkalinity, while principal component 2 is controlled by turbidity, color, and iron.

Factor analysis suggests that total hardness, calcium hardness, calcium, total alkalinity, and conductivity are the largest drivers of difference among water samples in this dataset. Turbidity, color, phosphate, and manganese play the most minor roles in describing the total variance among samples.

While cluster analysis yields little information about the relation of water quality to either aquifer rock type or geographic location, there is evidence that these factors may contribute to water quality in some way. Water conditions in communities with reported cases of Buruli Ulcer were such that growth of *Mycobacteria* may be favorable over that of other bacteria. Further analysis will yield a clearer explanation of the extent of this relationship. Clustering of variables shows that these water samples fit the profile that would be expected from groundwater. This alignment between real data and expectations suggests that water samples were properly collected and analyzed and that the data is reliable.

The results of this analysis suggest that calcium, magnesium, hardness, alkalinity, and pH values are consistently low for phyllite samples. With the exception of this observation, there is little relation between any parameter of groundwater quality and any of the external variables considered here. Variance among water samples is explained in large part by hardness and alkalinity as well as calcium and magnesium concentration.