Report Regional Planning, Spring 2016

Lauren Frazier, Yu Jin Hur, Disha Mendhekar

May 19, 2016

Abstract

The scope of the semester was to analyze the feasibility and need for an AguaClara plant in India. The team decided to focus on India as a country of potential implementation as India already has AguaClara LLC workers on the ground building connections. The team approached the analysis through large-scale research, gathering data on India's water sources, demographics, geography, and economy. There are ten ideal characteristics that would suggest a potential site. The team conducted research on India to determine locations with these characteristics. The team also created an optimization model using Matlab to visually display the deliverable in the form of a color-coded map.

Introduction

The Regional Planning team was created this semester to research new potential sites for AguaClara plants. The team's research will help AguaClara successfully expand its technology and take a big step towards its mission to provide everyone access to safe drinking water. AguaClara has spent most of its history in Honduras, but with recent forays into India, Colombia, and Puerto Rico, a Planning Team is necessary to gather data on new potential sites and analyze specific areas for feasibility and needs assessments. The process of choosing a location for a potential plant is complicated, involving many factors for an ideal site. After reading papers written by Karim Beers, Kenichi Victor, and Nishikawa Chavez, eleven ideal characteristics for an AguaClara plant have been identified. The plant must be placed where (1) the population is between 5,000 and 25,000, (2) there is a certain percentage of average per capita daily water use (yet to decide specific boundaries), (3) surface water is available consistently all year, (4) there is consistent aquifer recharge (rainfall or mountain runoff), (5) the surface water must be between 10 NTU and 1000 NTU, (6), there is a distribution system and surface water treatment plants available, (7) the consumers are willing to pay the tariff for plant operations, (8) the government set up is accepting of the facility, (9) there is little political conflict or sufficient stability to allow for community engagement, (10) local partnership is available, and (11) the groundwater has an undefined amount of Arsenic and Fluoride contamination. The team focused on the selection of a potential site in India. India already has AguaClara LLC workers on the ground meeting with local partners and building connections which is integral to the success of a plant. The team's job is to approach through the research angle, gathering data on India's water sources, demographics, geography, and economy. The team built an optimization model through a Python and Matlab combination in order to create a system of potential sites, giving weights to each of the considerations representing relative importance with regard to the other considerations.

Literature Review

The literature review consisted of three papers that together gave the basis of the team's research platform. The team's goal was to define a new ideal site for an AguaClara plant, and to develop a method to select future sites. The team extracted the potential steps and ideas from three papers. The first paper was a thesis by Kenichi Victor Nishikawa Chavez of the Cornell Institute of Public Affairs entitled "Introduction to AguaClara". In the paper, Chavez defined twelve steps to establishing a new pilot plant, which are: new market identification, needs assessment, local partner selection, knowledge cluster identification, local partner training, pilot community selection, investment procurement, community participatory planning, infrastructure construction, community capacity building, water treatment operations, and initial monitoring. Chavez states that the three main steps to be completed in the short term are to select a local partner, target a locality, and reach an agreement with a funding agency [11]. The step the team focused first on the needs assessment: why this place and nowhere else? It looked at three geopolitical levels: regional, state, and municipal; first looking at total natural water availability, surface water runoff, and replenishment; then at the percentage of population with drinking water access; then poverty levels by region correlated with access to drinking water. There are many different models for each step in the pilot process, but these twelve steps highlight the basic route for establishment. For example, in the local partner selection step, there are multiple subsequent steps for analyzing the stakeholders and involving the community versus basing the project on the government (or a hybrid model if the government and people are aligned). The three steps that must be completed first are to select a local partner, select a corresponding area, and then find funding. From reading this paper the team guided its own research process towards fulfilling his goals as well. We were involved in parts of the method, such as training the community or direct engagement with the local partner, but we will have to focus on locating the partner and also the potential funding. This paper gives a good framework for an overall process. His emphasis on the local partner gives a good perspective on how important this step will be.

The second paper in the literature report was a collaboration from Gonzalez, M., Beers, K., Weber-Shirk, M., and Warner, M. in 2014, titled "Analyzing the Potential of Community Water Systems: the Case of AguaClara".The purpose was to develop a framework for community water management system through analyzing the case of AguaClara in Honduras. It identified three broad challenges in terms of water provision, which are: technology, management, and governance. It emphasized the interactions of the various forms of capital and the domains of technology, management and governance to develop a community capitals framework for provision of water, and then analyzed AguaClara on the basis of the framework developed. The conclusions drawn are threefold: one, that there are three keys for sustainable community water systems: "technology that is resilient, affordable and reliable; management that is within the reach of the human and financial resources of the community; governance system that is trusted and supported by the community and by higher levels of government."; two, that AguaClara strengthens the interactions between the domains of technology, management and governance by focusing on capacity building and engaging in cooperative governance where information and technical expertise is shared to create innovative solutions and building effective communication and resource between the local communities; and that three, the key area of concern is to bring the undeserved population into the network through a method of cross-subsidization of tariffs by building connection with higher levels of government.

"Analyzing the Potential of Community Water Systems: the Case of AguaClara", however, does not take into account the feature of conflicts and protests, which will be crucial taking the context of India. How to resolve conflicts related to the varying interests of communities is one area that needs to be explored further. The other area could be the growing corruption and negligence towards the under-served populations, on the part of the political capital. It also brings up the question of how to help the poorest of the poor and the socially backwards cultures and draw them into this network, because they are often excluded from access to safe drinking water; the concept of water being an economic good versus a human right is also a complicating theme. This article helped the team in developing a conceptual framework for expansion of AguaClara technology in the Indian context. AguaClara has strong technological links in India but needs a framework to fortify the domains of governance and management as these systems can be very complex given the site location. To provide sustainable community water systems in India, strong domains of technology, governance and management which are continuously interacting with each other are very important. These interactions can be in terms of adaptive learning which emphasizes problem-solving and value creation for which AguaClara in Honduras provides the best example.

The third paper in the literature review is the thesis written by Karim Beers, "Governance Models for Community Water Systems: the Case of AguaClara" [2]. The method presented was a multiple case analysis to assess the strengths and weaknesses of AguaClara's governance model. The five main categories that AguaClara should consider before building a plant is the natural (natural resources available in the community), physical (built aspects of the water system such as treatment plant and distribution system), financial capital, social capital(community ownership and trust, good management, and learning), human capital, and political capital(presence of ongoing external support, and an enabling policy, legal, and regulatory environment).

It would be helpful if the Village Water Board (VWB) or community owned the water source because it would mean the people have direct say in their water distribution, and it is easier to work with a local board than a state government who may be physically removed from the village.

When assessing physical capital, planners should assess the distribution system as there could be problems such as leaky water pipes which could raise demand for water to levels that the plant is unable to provide, leading to sub-optimal water. If the pipe bringing water from river to plant is in poor condition that is also a problem.

When assessing financial capital the team should focus mainly on engineering studies, construction, and operation and maintenance. Along with everybody paying the required amount, administrators must properly manage funds, pay employees on time, and set aside money for maintenance.

For capital costs, AguaClara has depended on third party funding (mostly from NGOs). There will be a cost associated with connecting people who are not yet connected to the water plant by methods such as subsidizing tariffs and putting in pipes, and that should be considered.

In regard to social capital we need community ownership and trust, good management, and learning. AguaClara's technology is sensitive to the ability of its operators since it is designed to be controlled by a plant operator so the plant needs capable and committed operators. There will need to be ongoing external support for small community organizations to maintain services over a long term.

"The AguaClara model involves strong partnerships between a university, an NGO, and community organizations" [2].Cornell University in this case provides technical assistance, the NGO provides administrative assistance, and the community provides social capital. The team should research policy, legal and regulatory framework of a country to account for these considerations, and it would be helpful to have a direct source on the ground, or to visit the sites directly. However, the author doesn't seem to have visited many plants as his analysis came from information he gathered through papers, web, and interviews.

Summary There were many takeaways from the literature review used throughout the semester. The most important of these takeaways was understanding each author's emphasis on the social and cultural aspect of the process. The team previously had considered the most important characteristic to be the physical culture: the geography, surface water availability, and the physical ability to implement a successful plant. While the physical factors were, in the team's mind, still of chief concern, it was now understood that cultural and social factors are equally important. An AguaClara plant will never succeed without support of the local people, as well as a demand and willigness from them to accept a foreign company building infrastructure in their homes. The AguaClara model focuses on a strong partnership between many different organizations and understands the importance of establishing connections on the ground. Without a local partner and without a strong needs assessment, the feasibility of a plant is low.

From Beers, the team learned that funding must be considered a chief concern, and thus will consider the poverty ratios in an area to be a very high factor, meaning: will the people be able to pay the necessary tariff for the plant? Beers also gave the team a good perspective on the pros and cons of AguaClara's governance model. Gonzalez et al. gave insight into three broad challenges of water provision: technology, management, and governance. The team also was able to conclude that the AguaClara success in Honduras was built upon the foundation of connections with locals and governments. Without trust and relationships on the ground, it will be difficult to understand the needs of the people and what they want. Since the foundation for the success in Honduras was due to prior connections, the team realized that the case of choosing sites in India was going to be different than the case in Honduras as we had no prior connections in India. Finally from Chavez's paper "An Introduction to AguaClara", the team identified some of the steps to establishing a pilot plant, and decided which of Chavez's steps are the most relevant to this semester's research. The team also realized that it had to look at places that require this technology. In other words, places where a demand for this technology existed as it would be no use to install a water treatment plant in a place that had poor quality surface water for nobody to use it.

Methods and Discussion

India consists of 29 states, and the states are further split into 532 districts in total. Within each district, there are up to 1,000 villages. The total number of villages to analyze in India is roughly 700,000. The team first looked at data on the state level. The methods section will present the data based on the ideal site characteristics listed in the introduction section, which are: population size, surface water availability, annual rainfall, water turbidity, availability of distribution system, consumers' willingness to pay, government set-up, ethnic homogeneity index, and other considerations. On a side note, all the Union Territories have been eliminated in the analysis. Union territories have a different government set-up as these are ruled directly by the central government. These federal territories are majorly urbanized and work like any of the census town in India. Since this analysis analysis is primarily focused on the rural areas, these union territories have not been considered [1].

Union Territories in India -

- Andaman and Nicobar Islands
- Chandigarh
- Dadra and Nagar Haveli
- Daman and Diu
- Delhi (National Capital Territory of Delhi)
- Lakshadweep
- Pondicherry

After analyzing the ideal site characteristics a weight was placed on each of these characteristics. The higher the weight, the more important the characteristic is for the ideal site. The weights was a number between 0 and 1 and represented the relative importance of the characteristic. The weight of each characteristic was decided based on both the literature review and on previous AguaClara plants that have been successful in the past. Many different weights were analyzed for the optimization program and tested for validity; the team decided on the current set of weights after iterating through many trials and judging which set of weights produced logical results. A further discussion of the weighting system can be found under "Matlab Optimization Program" on page 20. For the state level analysis, the data was normalized by its variability to provide equitable scaling. For example, larger numbers in rainfall might overshadow smaller numbers such as poverty ratio if the data was not normalized; the equations for the optimization can be found underneath the section titled "Matlab optimization program". The result of the optimization program was the sum of each state's characteristics multiplied by the weights. The higher the final number, the more likely the state was the optimal state for an AguaClara plant. The optimization code was written as inputs into a Matlab code. The code for producing a color-coded map that shows where each state lies on the scale from most to least ideal using the optimized scores of each state (from optimization code from Matlab) was written in Python. Matlab and Python were chosen because members the Regional Planning team had the most expertise in those languages. Using this algorithm the most desirable states were chosen and then the team worked down towards district scale using a similar method. This program was ideal from a planning perspective because it required simple weighting inputs from a user and can be easily adjusted, giving visually appealing and cognitively simple results. Below is a figure of the Administrative State Boundaries of India (Figure 1).

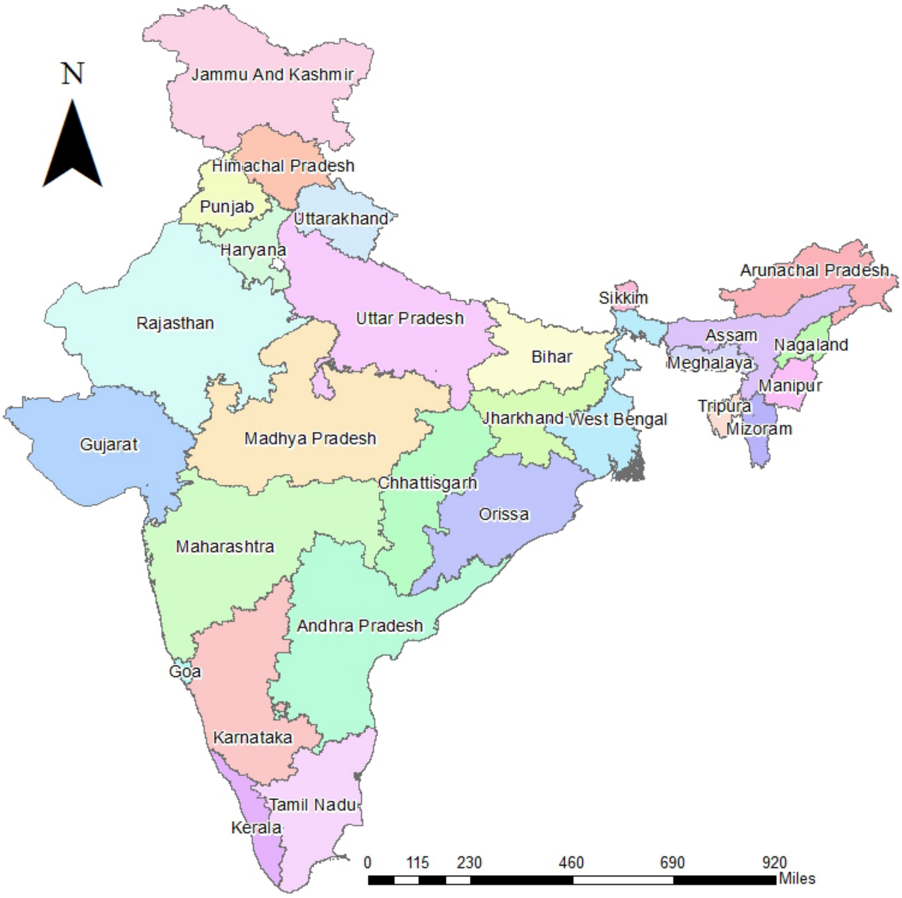


Figure 1: Administrative Boundaries of India - States (Census of India, 2011)

0.1 Population Size

The first constraint is the population range ideal for an AguaClara plant capacity. Based on discussion with Monroe, a population range between 5,000 and 25,000 people was chosen for the final site. This population limit ensures that the revenue collected from the community is sufficient to finance the operation and maintenance costs of the plant. Also this is the range of population that previous successful plants in Honduras have been able to support. This range will be ideal for the next AguaClara plant because it is within the scope of experience and knowledge. On the state level, this constraint did not apply and the optimization program did not account for it because the constraint only applies to the village sizes. The population constraint will only apply once the program runs on the village level.

0.2 Surface Water Availability

According to the AguaClara website [3], one of the technical requirements for establishing a water treatment plant in rural communities is that the water source should be surface water or ground water influenced by surface water.

India has one hundred and thirteen river basins. Fourteen of these are large, forty-four are medium and fifty-five are minor river basins. The most likely AguaClara sites will be located nearest a river basin, and the most ideal would be a larger basin that is unlikely to be dry throughout the year. The major river basins of India in descending order of area are: the Ganga, Indus, Godavari, Krishna, Brahmaputra, Luni, Mahanadi, Narmada, Kaveri, Tapi, Pennar, Brahmani, Mahi, Sabarmati, Barak, and Subarnarekha. The major river basins form about 84 per cent of the total drainage area of the country. The natural water bodies existing in India are shown in Figure 2 and the availability of surface water is shown in Figure 3.

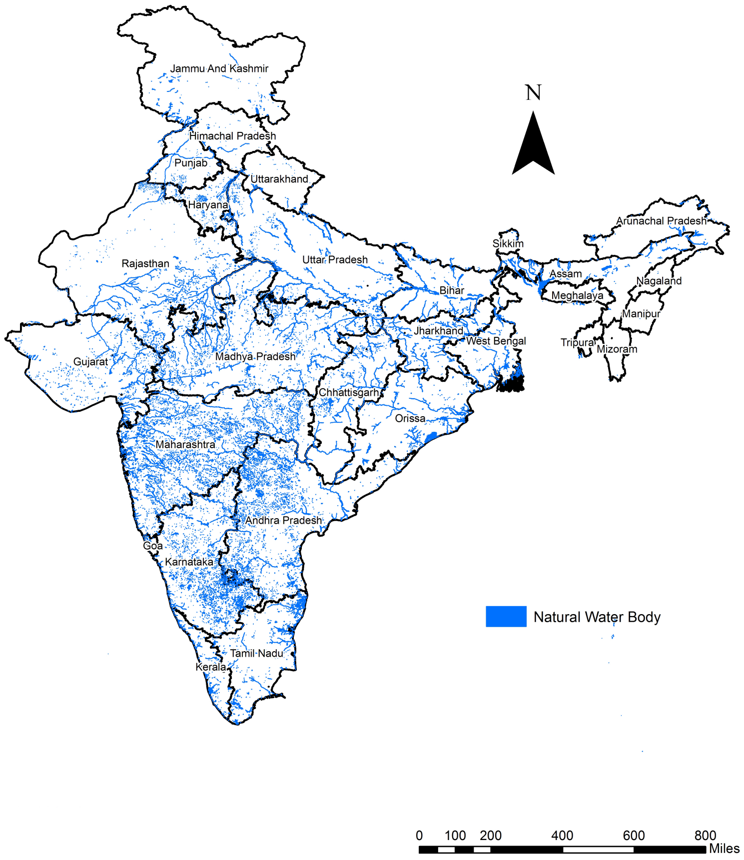


Figure 2: State-wise distribution of Natural Water Bodies [4]

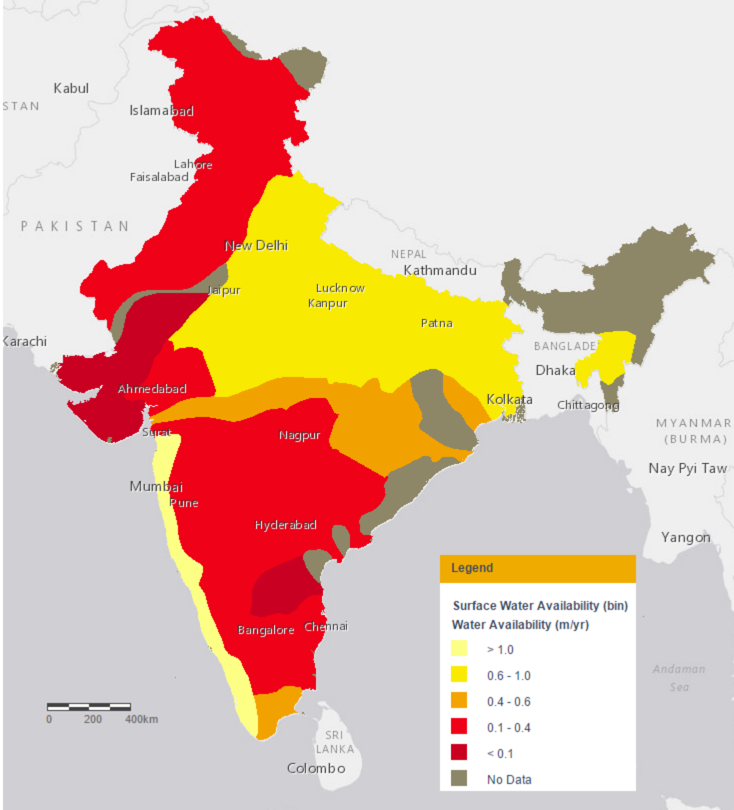


Figure 3: Surface Water Availability [5]

Since the location of dams and reservoirs will have high correlation with availability and demand of surface water, the locations of various dams and reservoirs are also mapped out in Figure 4. Currently, according to the Constitution of India, it is the state's responsibility to develop and manage its water resources. All the states are required to develop their state's water policy within the framework of national water policy and accordingly set up a master plan for water resources development. India's National Water Policy adopted in 1987 emphasizes the need for water use planning based on river basins. Water allocation priority has been given to drinking water followed by irrigation, hydro-power, navigation and industrial or other uses. Hence, it will be a tedious task to locate the sites of all the water treatment plants in India. Also, it is unknown whether these water treatment plants use surface water or ground water. This data can be further pursued once the team has narrowed down India into potential towns.

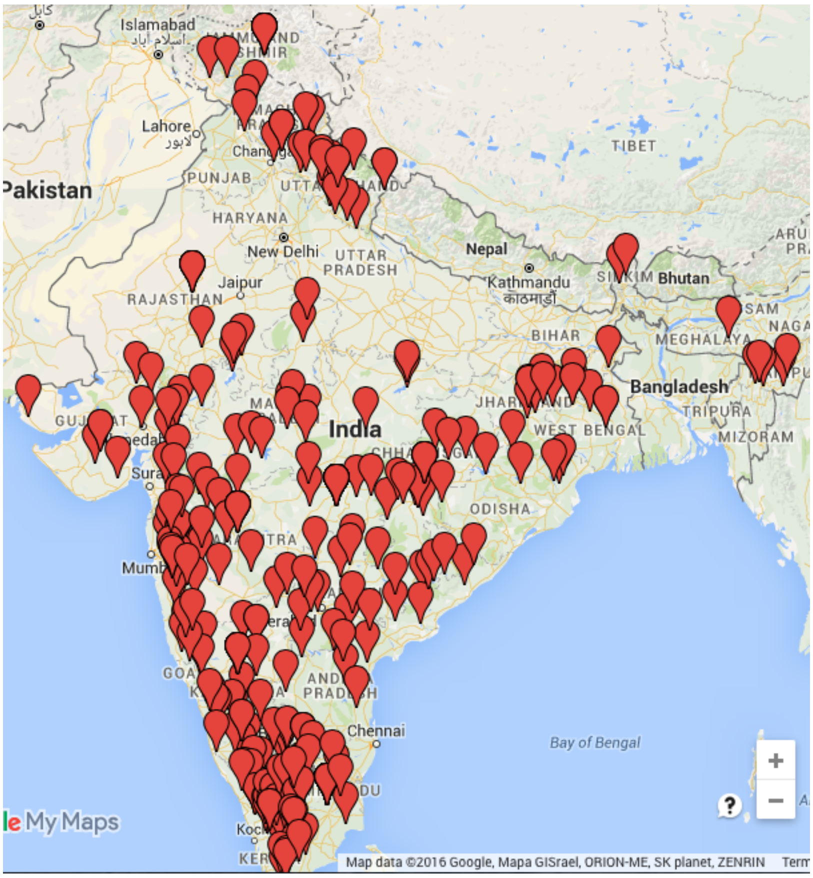


Figure 4: Dams and Reservoirs in India, (Google Maps)

Also, data from Ministry of Water Resources, India on state-wise percentage of rural population dependent on surface water for domestic purposes was gathered [5]. The surface water sources include rivers, canals, springs, tanks, ponds and lakes. This data gave an idea of which states in India have the highest demand for surface water. This data was mapped out on ArcGIS and the team analyzed which states have higher percentage of population dependent on surface water for drinking purposes as shown in Figure 5 and 6.

When Figure 3 and Figure 6 were compared, the team analyzed which areas had high surface water availability along with a higher than average amount of Rural Population dependent on Surface Water for Drinking Purposes. The eastern region of Rajasthan, most of Uttar Pradesh, parts of Jharkhand, parts of Orissa, Goa, a thin strip going down all of the western side of Karnataka and Maharashtra, and a good portion of Kerala all had surface water availability of more than 0.4 m/year and more than 2% of the rural population were dependent on surface water (Figure 6). While the states in the north-east part of India are highly dependent on surface water for drinking purposes, they do not have high surface water availability. The high dependence on surface water despite low availability could be because of the topography of these regions. These regions are located in the Himalayas, high above the sea level, and extract water from the streams that flow downhill from the glaciers. Due to this reason, the surface water here is not so polluted as compared to the the western coast where there exist instances of high pollution in the rivers owing to industrialization [8]

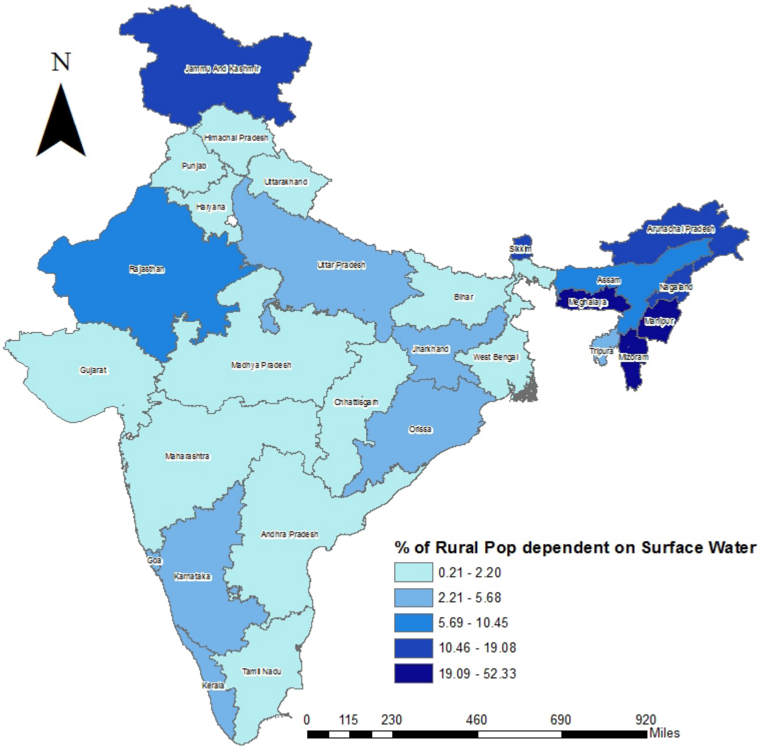


Figure 5: Percentage of Rural Population dependent on Surface Water For Drinking Purposes [5]

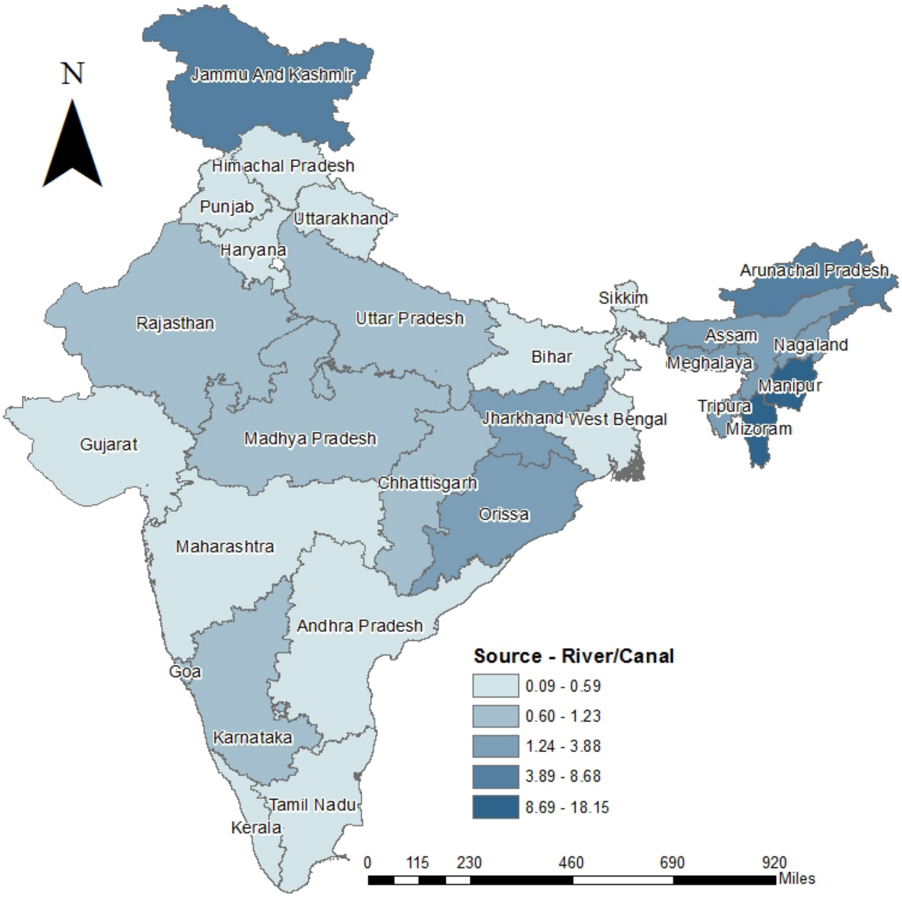


Figure 6: Percentage of Rural Population dependent on Rivers or Canals For Drinking Water [8]

0.3 Annual Rainfall

The annual rainfall is optimally more than 1000 millimeters. Although this is not a limiting constraint since it is possible that a drier region has a continuous source of surface water from a lake or from mountain snow runoff or other sources, the Matlab optimization code included annual rainfall as a very small factor (weight of 0.01) because even though rainfall is not a limiting nor very important factor, it is still relevant to surface water availability and would still be a helpful addition to the site selection. A map of annual rainfall data per state (Figure 7), sourced from the Ministry of Water Resources, India, is generated below. The regions of North-east India and states of Kerala, Goa, Orissa and West Bengal receive the highest amount of rainfall annually.

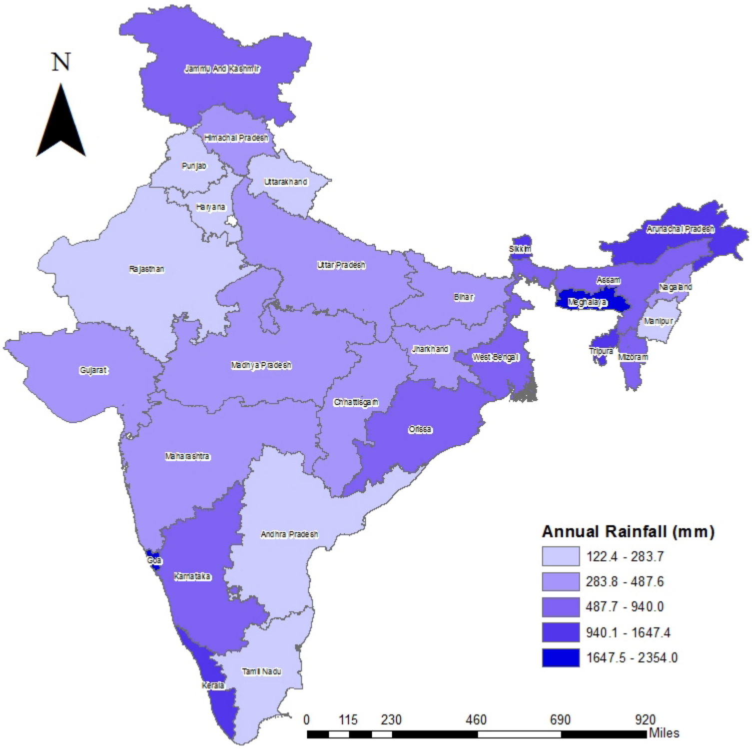


Figure 7: Annual Rainfall in millimeters, [5]

0.4 Water turbidity

The constraint for water turbidity is one of the most important factors but also has proven to be one of the most difficult to evaluate. The team was unable to locate data on the water turbidity of various surface water bodies in India. The constraint between 10 NTU and 1000 NTU for an optimal AguaClara plant will prove to be useful when the water turbidity data is located. However, upon speaking with members of AguaClara LLC, the team has decided to forgo further research and instead recommend on-the-ground contacts to uncover this data. AguaClara LLC depends on communication with local water agencies and field experiments using turbidity tests in order to determine this metric; they were also not able to find the information online.

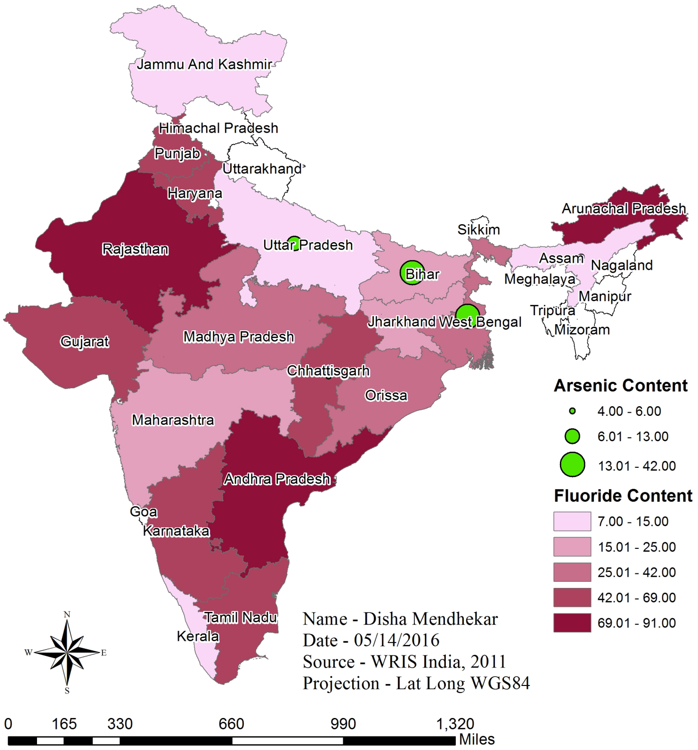
0.5 Fluoride and Arsenic Contaminated Groundwater

Arsenic or fluoride contaminated groundwater could be another factor correlated with a need for surface water treatment.

The contamination level for Fluoride according to the Central Groundwater Board, India, is more than 1.5 mg per liter and for Arsenic it is more than 0.05 mg per liter. According to Table 1 and Figure 8, the states of Rajasthan, Andhra Pradesh and Arunachal Pradesh have the highest proportion of districts with Fluoride contaminated groundwater. For Arsenic, Bihar and West Bengal show the highest contamination.

Table 1: State-wise Districts with Fluoride and Arsenic Contaminated Groundwater [9]

States	**Fluoride contaminated districts**		**Arsenic contaminated districts**		**Total No. of districts in state**
	Number	**Proportion (%)**	**Number**	**Proportion (%)**	
Andhra Pradesh	19	83	-	-	23
Arunachal Pradesh	13	81	-	-	16
Assam	4	15	1	4	27
Bihar	9	24	15	39	38
Chhattisgarh	12	67	1	6	18
Delhi	5	56	-	-	9
Gujarat	18	69	-	-	26
Haryana	14	67	-	-	21
Jammu & Kashmir	2	9	-	-	22
Jharkhand	6	25	-	-	24
Karnataka	20	67	-	-	30
Kerala	1	7	-	-	14
Madhya Pradesh	19	38	-	-	50
Maharastra	8	23	-	-	35
Odisha	11	37	-	-	30
Punjab	11	55	-	-	20
Rajasthan	30	91	-	-	33
Tamil Nadu	17	53	-	-	32
Uttar Pradesh	10	14	9	13	71
West Bengal	8	42	8	42	19
Total	**237**	**42**	**34**	**6**	**558**



Name - Disha Mendhekar
Date - 05/14/2016
Source - WRIS India, 2011
Projection - Lat Long WGS84

Figure 8: State-wise Districts with Fluoride and Arsenic Contaminated Groundwater, [9]

0.6 Socio-economic Characteristics

Because households will be paying for the operation of the plant through a tariff, the optimal community will 1) be financially able to pay for this tariff, and 2) have enough demand for water treatment that they will want to pay the tariff in exchange for cleaner water. To begin quantifying willingness to pay, the team gathered data from the Census of India with poverty rates throughout India. The higher the poverty level, the more unlikely that there will be funds to pay the necessary tariff, considering that the tariff will be, at current estimate, over 10 percent of the average rural family's income. Another parameter found is data on the percentage of people with access to clean drinking water. This is an important parameter because this will correlate with a higher demand for water treatment and accessibility.

Looking at the rural poverty ratios of each state (Figure 9), the data shows that the states of Jharkhand, Chattisgarh, Arunachal Pradesh and Manipur have the highest poverty ratio averaging to 41% compared to a national average of 26%.

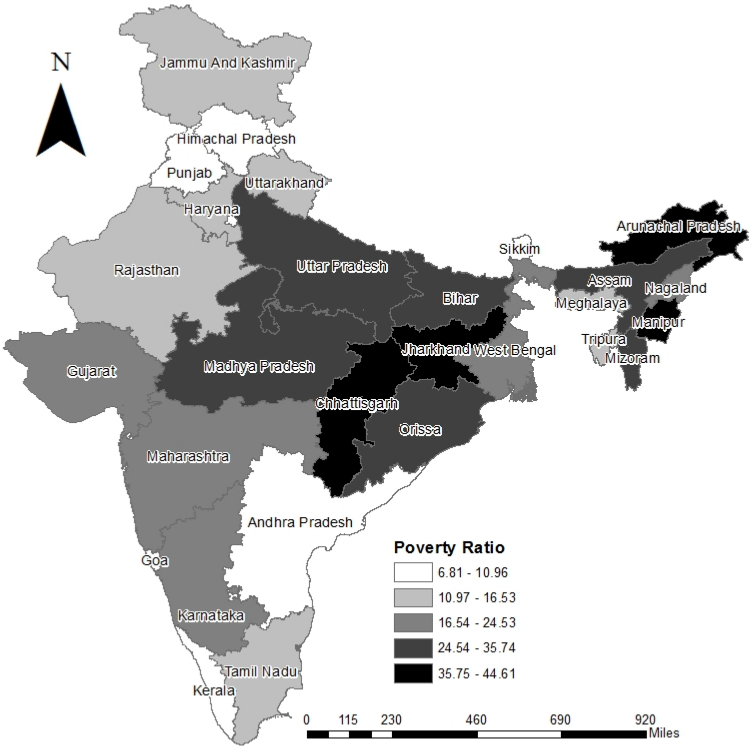


Figure 9: Rural Poverty Ratio [7]

Also in Table 3 and Figure 11, the Human Development Index (HDI) produced by the Institute of Applied Manpower Research and Planning Commission in Indian Human Development Report, 2011, has been added, as well as the column for the coverage of drinking water in the rural areas of each state. This data was drawn from Ministry of Rural Development India for the year 2011. The national average HDI was 0.52 in the year 2011 while the national average coverage of drinking water in rural India is 90%. Table 3 showed that the states of Bihar, Jharkhand, Chattisgarh, Orissa and Uttar Pradesh have the lowest HDI while the states of Kerala, Manipur, Meghalaya and Mizoram have the lowest percentages of rural population with access to drinking water.

Current estimates show that households will need to pay a tariff of $4 a month. This is expensive for rural households in India, and past the upper limit for a rough figure on a rural household's willingness to pay for better quality water. This upper limit is based on a study conducted in 2000, which analyzed 15 contingent value studies from various developing nations and found the range of willingness to pay between 0.29 and 10.7 percent of monthly income [12]. The average annual income of rural Indian households was found to be 22,400 rupees, which equals a monthly income of 28.03 US dollars[12]. Based on this value, requiring $4 of tariff per month would take away 14 percent of a rural household's income. Therefore, the current tariff estimate is outside the range of willingness to pay that was found through multiple studies in developing nations. The rural median household income is shown in Figure 10.

According to another World Bank study conducted in the states of Maharashtra and Kerala for the year 2012, villagers paid Rs. 50 for the first 12,000 liters of water per households and Rs. 10 was charged per additional 1000 liters [10]. This water tariff is less than a dollar and accounts for 10% of their monthly income. Therefore it seems that $4 will be a high charge for villagers to pay for their water because it is over 10%.

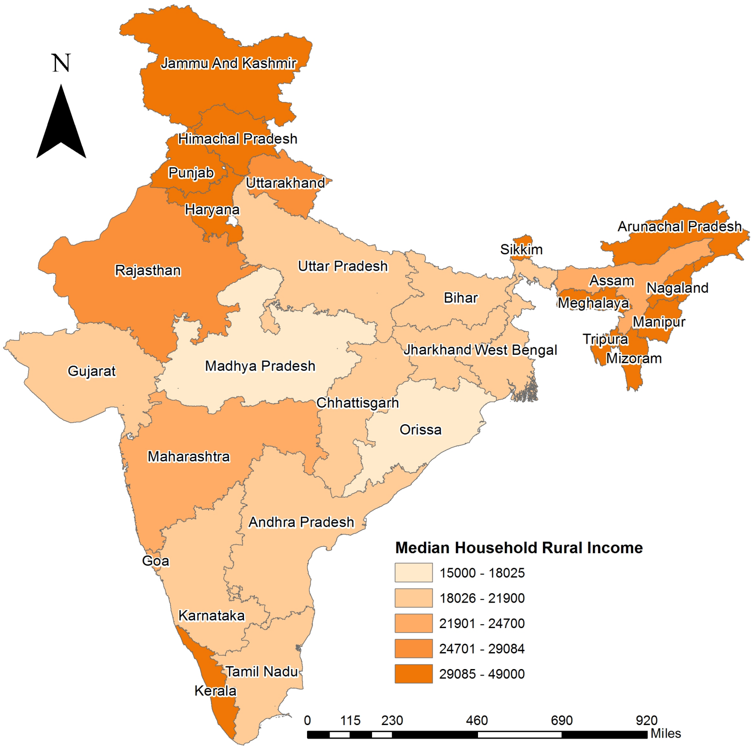


Figure 10: Rural Median Household Income [6]

Table 3: State-wise HDI and Drinking Water [7]

S.No	**State Name**	**Human Development Index**	**% of rural pop with access to drinking water**
1	Andhra Pradesh	0.47	76.9
2	Arunachal Pradesh	0.57	73.7
3	Assam	0.44	56.8
4	Bihar	0.37	86.1
5	Chhattisgarh	0.36	66.2
6	Goa	0.62	58.3
7	Gujarat	0.53	76.9
8	Haryana	0.55	81.1
9	Himachal Pradesh	0.65	87.5
10	Jammu And Kashmir	0.53	54.9
11	Jharkhand	0.38	35.5
12	Karnataka	0.52	80.5
13	Kerala	0.79	16.9
14	Madhya Pradesh	0.38	61.5
15	Maharashtra	0.57	68.4
16	Manipur	0.57	29.3
17	Meghalaya	0.57	29.5
18	Mizoram	0.57	23.8
19	Nagaland	0.57	47.5
20	Orissa	0.36	62.9
21	Punjab	0.61	96.9
22	Rajasthan	0.43	60.4
23	Sikkim	0.57	67
24	Tamil Nadu	0.57	85.3
25	Tripura	0.57	45
26	Uttar Pradesh	0.38	85.5
27	Uttarakhand	0.49	83
28	West Bengal	0.49	87

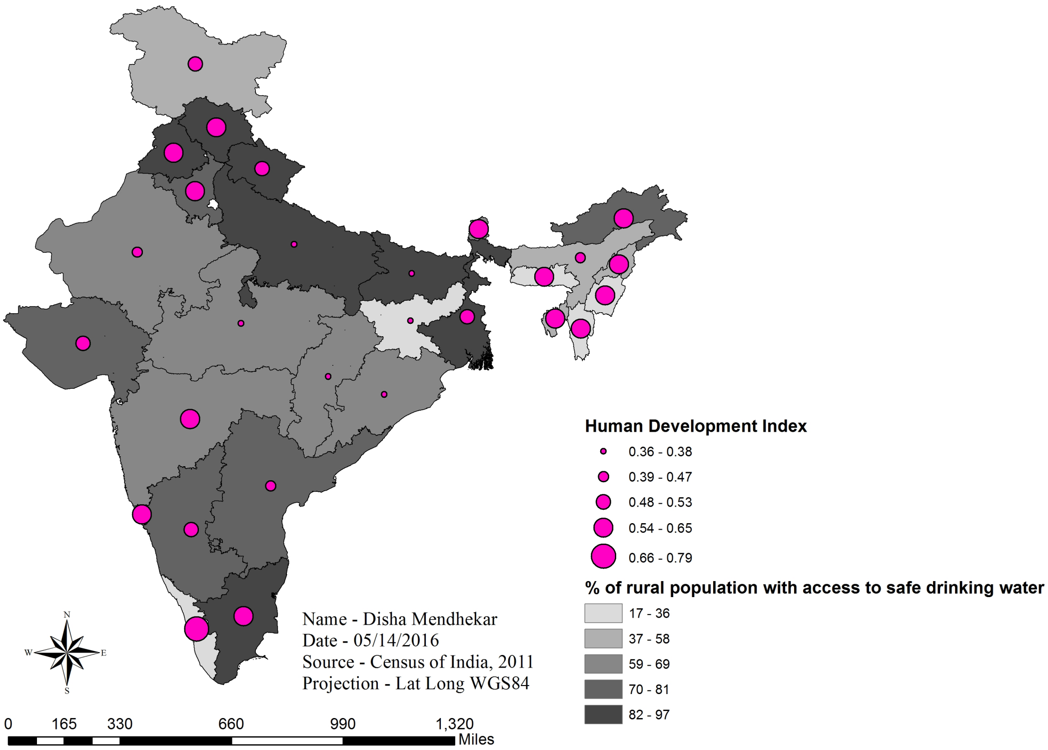


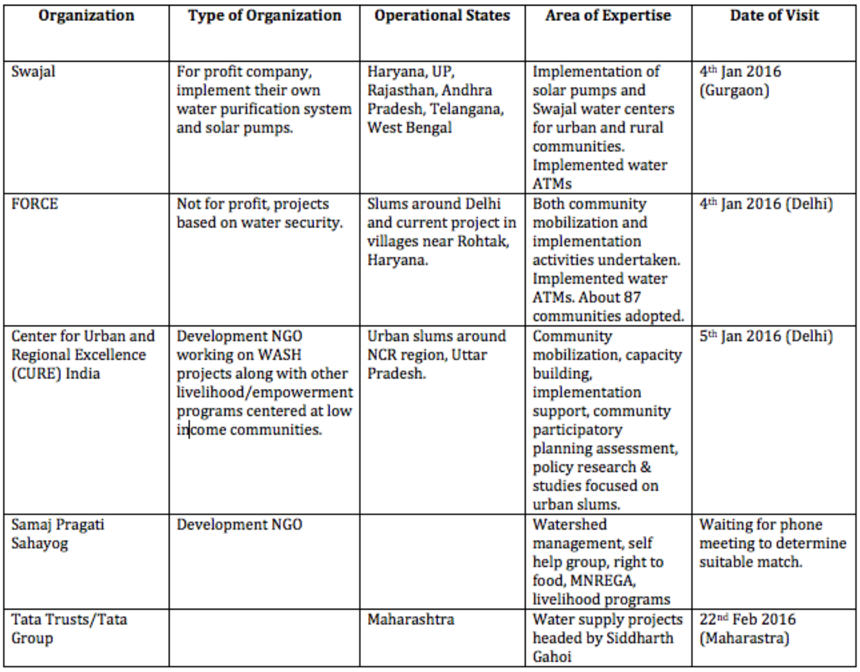
Figure 11: State-wise HDI and Drinking Water, [7]

0.7 Other Considerations

The team acknowledged that there are considerations for an AguaClara plant outside the eleven listed above. Elements not included in the original AguaClara objectives list have also been considered, one of which is political conflict. Many states in India are high in political conflict and therefore might not be receptive to foreigners or stable enough to communicate with an outside government. It will fall upon the next team to analyze the districts optimized in this report for the existence of political conflict and the viability of that district.

Another consideration also included is the existence of a local partner. AguaClara LLC currently has people on the ground speaking with local partners and gauging their ability, reach, and infrastructure. John Finn, senior director at AguaClara LLC, provided a report detailing their progress on the ground in India. Within his report is a list of potential local partners for AguaClara's future work based on the groups they have gotten into contact with to this date. The current list has eight potential partners and spans a large region of India. The full list of potential partners, courtesy of Mr. Finn, is shown in Table 5 and is shown graphically in Figure 12. The AguaClara LLC contacts on the ground will be vital to the success of a new AguaClara plant, which is why the existence of a partner was included in the optimization program. It is not, however, the most important objective, because it was recognized that AguaClara LLC has not been able to reach many parts of India and therefore the lack of partners in that area may not be due to viability but other constraints such as time and movement.

Table 5: Potential Partners in India (Finn, 2016)



Organization	Type of Organization	Operational States	Area of Expertise	Date of Visit
Swajal	For profit company, implement their own water purification system and solar pumps.	Haryana, UP, Rajasthan, Andhra Pradesh, Telangana, West Bengal	Implementation of solar pumps and Swajal water centers for urban and rural communities. Implemented water ATMs	4th Jan 2016 (Gurgaon)
FORCE	Not for profit, projects based on water security.	Slums around Delhi and current project in villages near Rohtak, Haryana.	Both community mobilization and implementation activities undertaken. Implemented water ATMs. About 87 communities adopted.	4th Jan 2016 (Delhi)
Center for Urban and Regional Excellence (CURE) India	Development NGO working on WASH projects along with other livelihood/empowerment programs centered at low income communities.	Urban slums around NCR region, Uttar Pradesh.	Community mobilization, capacity building, implementation support, community participatory planning assessment, policy research & studies focused on urban slums.	5th Jan 2016 (Delhi)
Samaj Pragati Sahayog	Development NGO		Watershed management, self help group, right to food, MNREGA, livelihood programs	Waiting for phone meeting to determine suitable match.
Tata Trusts/Tata Group		Maharashtra	Water supply projects headed by Siddharth Gahoi	22nd Feb 2016 (Maharastra)

Figure 12: Locations of Partners' Work (Finn, 2016)

0.8 Matlab optimization program

The team has written a program in Matlab that mirrors the steps outlined in the beginning of the Methods section. The code processes data for seven objectives: 1) the percent of people dependent upon surface water in each state; 2) the poverty ratio in each state; 3) the annual rainfall in each state; 4) the HDI index per state; 5) the rural population of each state normalized by the total population in the state; 6) the presence or absence of an established local partner, as supplied by AguaClara LLC; and 7) the total rural population in a state by the total rural population in India. There are seven objectives in the program because they are the ones with consistent, quantifiable data; other objectives, such as the governmental set-up and lack of political conflict, are harder to quantify and will not be used in the optimization code. Instead, the team used the program to identify potential states, and then cross-check the "most ideal" states against quantifiable factors such as low political turmoil, etc.

The weights assigned to these objectives were relative to each other and are dimensionless. They were decided based upon the discussion in the literature review about which considerations are most important to choosing an AguaClara site, and upon past successful AguaClara plants that required these characteristics. The weights were as follows:

Weight1=1.3; The dependence upon surface water.

Weight2=0.75; The poverty ratio.

Weight3=.01; The annual rainfall.

Weight4=.25; The HDI index.

Weight5=.3; Rural population percentage per state.

Weight6=.15; The presence of a local partner.

Weight7=.3; The rural population in the state by the total rural population in India.

The characteristic that proved most important was the dependence upon surface water. This is the core of an AguaClara plant's success: the demand for surface water treatment. This characteristic must be further correlated with water turbidity data, but this semester's team was not able to obtain that data and it must be analyzed in future semesters. The next highest characteristic was the poverty ratio; the reason for this is that it is almost equally important that the chosen site be able to pay the maintenance and operations tariff. The tariff, as discussed earlier in the report, is relatively high relative to the average income of a rural Indian family, and a site with many families living beneath the poverty line will likely not be willing or able to pay the tariff. A further discussion of this can be found under the heading "Socio-Economic Characteristics". The remaining weights were decided based on relativity and are less important than the first two. The annual rainfall parameter is almost negligible and therefore had a weight of 0.01. It was included only because it may be important at a lower level and the weight may be changed then. A higher HDI index is desirable because it correlates with higher rates of literacy, governmental stability, and wealth. The two rural population parameters, by state and by country, were both 0.3 weighted because a higher percent rural population is desirable but not totally necessary, because while an AguaClara plant does target the rural popluation and thus it will be more feasible that an ideal site is within a state with a higher rural population, there could be states with high urban populations and still have ideal rural villages. While the presence of a local partner is very, very important, we have incomplete information on the local partners in India (there is only available information from AguaClara LLC's current contacts), it is very likely that new local partners can be found. The team settled on this weighting system after many iterations. At first, Weights 1 and 2 were weighted the same, but the results were skewed because states with nearly nonexistent dependence upon surface water seemed to be optimal because they had very high poverty ratios. The current system reflects the team's decisions of order of importance.

The equations used are as follows:

$$\vec{Weight = Weight1, Weight2, Weight3, Weight4, Weight5, Weight6, Weight7} \quad (1)$$

$$Data_N = \frac{(Data_p - min(Data_N))}{max(Data_N) - min(Data_N)}. \quad (2)$$

$$Data_S = \sum Weight * Data_N \quad (3)$$

Equation 1 presents the seven weights joined into a vector of numbers used for weighting. This vector was used in the optimization algorithm. Equation 2 presents the normalization of each data point. The normalization of the data allowed the different parameters to be compared within the same scale. Equation 3 shows the simple method for multiplying each weight by its corresponding normalized data point. The result was a single number, the total sum. The total sum for each state was compared to the other states', and the highest number was our "optimal" state for this portion of the program. The results of the program are seen in a bar chart in figure 13.

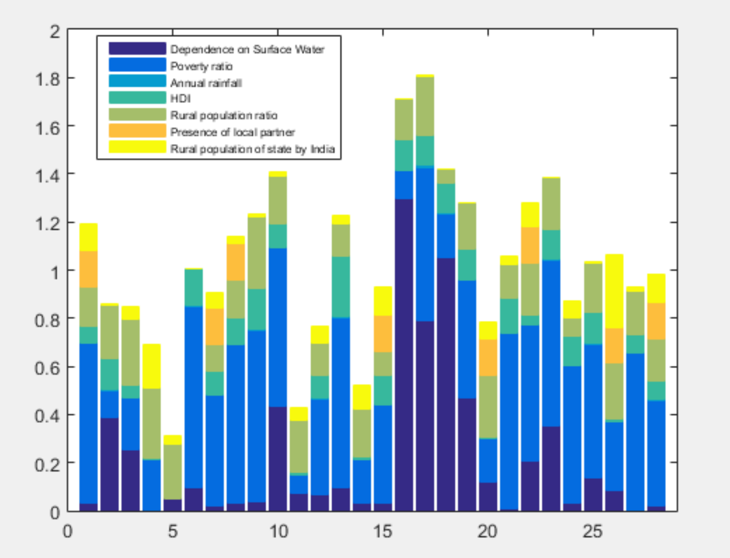
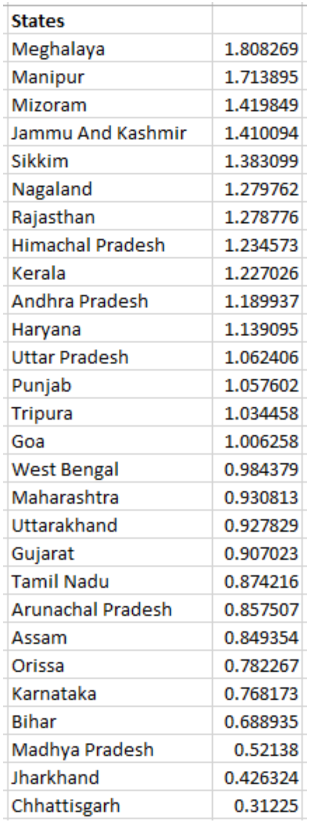


Figure 13: Most optimal states

The stacked bar chart represents one bar for each state with the different parameters aggregated to make up to the total. A stacked bar chart is the optimal visual representation because it shows the distribution of parameters that contribute to the total. The x-axis represents 28 different states and they are alphabetical. The list of states along with their relative place in our optimized results is below in Figure 13, Table 6, and Figure 14:



States	
Meghalaya	1.808269
Manipur	1.713895
Mizoram	1.419849
Jammu And Kashmir	1.410094
Sikkim	1.383099
Nagaland	1.279762
Rajasthan	1.278776
Himachal Pradesh	1.234573
Kerala	1.227026
Andhra Pradesh	1.189937
Haryana	1.139095
Uttar Pradesh	1.062406
Punjab	1.057602
Tripura	1.034458
Goa	1.006258
West Bengal	0.984379
Maharashtra	0.930813
Uttarakhand	0.927829
Gujarat	0.907023
Tamil Nadu	0.874216
Arunachal Pradesh	0.857507
Assam	0.849354
Orissa	0.782267
Karnataka	0.768173
Bihar	0.688935
Madhya Pradesh	0.52138
Jharkhand	0.426324
Chhattisgarh	0.31225

Table 6: Ranking the optimal states

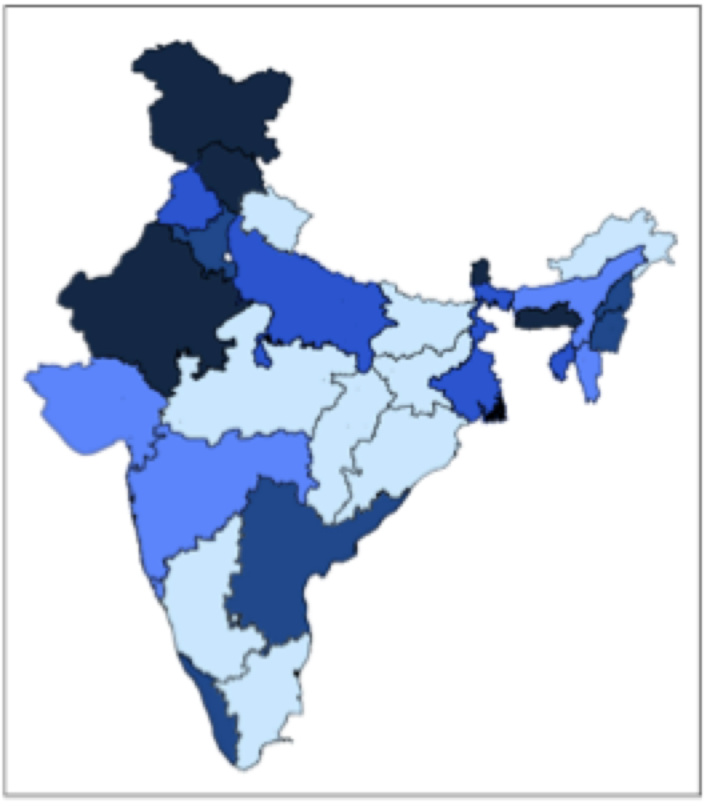


Figure 14: Most optimal states from applied weights

Once the optimization for the states was completed, the top five states were analyzed on a district level. The states of Meghalaya, Manipur, Mizoram, Jammu and Kashmir, and Sikkim were the top five of the optimization program. However, data for the Northeast states in India was difficult to procure, and due to feasibility and time constraints, the states of Manipur and Mizoram were not analyzed. The state of Jammu and Kashmir was also not analyzed because of well-known political conflict in the state. Therefore, the top five states became: Meghalaya, Sikkim, Rajasthan, Himachal Pradesh, and Kerala (in that order). The data on the district-wise level was obtained through similar research methods, largely through Human Development Reports on the various states of India throughout the years. The Census of India also provided district-wise data. The optimization program in Matlab was very similar for the district data as for the state data, and running the districts numbers through the code gave the results shown below in Figure 15.

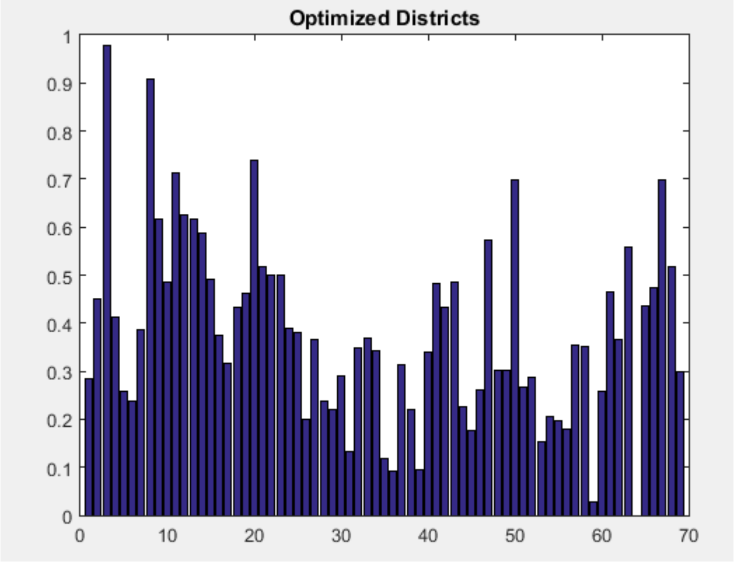


Figure 15: Most optimal districts

The program did not generate a stacked bar chart because there are over 70 districts and it was a very complicated and cognitively overwhelming result. The top ten districts are shown in Table 7 below, with their relative results. The graphical representation is shown in Figure 16.

Table 7: Ranking the optimal districts

Rank	**S.No**	**Optimization**
1	South Garo Hills	1.61
2	Chamba	1.56
3	Barmer	1.53
4	East Garo Hills	1.41
5	West Garo Hills	1.37
6	Ganganagar	1.36
7	Lahul and Spiti	1.33
8	West Khasi Hills	1.31
9	Ribhoi	1.30
10	Nagau	1.29

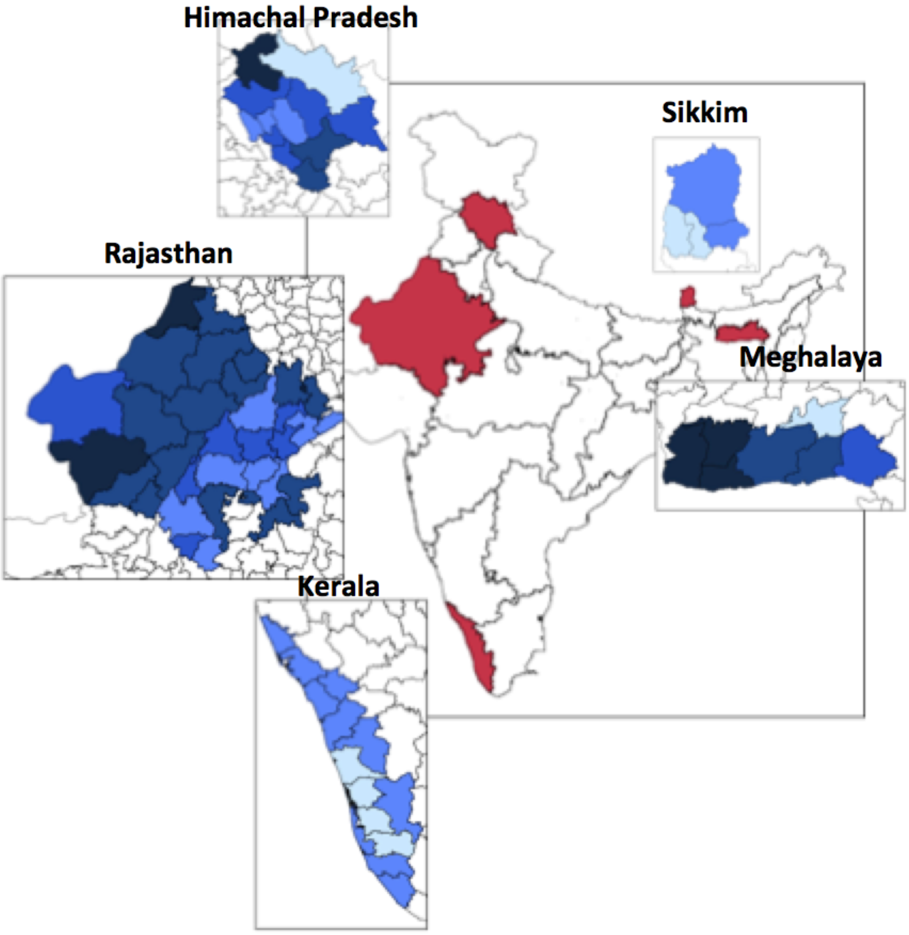


Figure 16: Most optimal districts from applied weights, red=optimal state, darkest blue=most optimal district

The code for the two optimizations can be found on the AguaClara Google Drive. The files are within the same folder labeled "Regional Planning 2015-2016". The folder also includes Excel sheets of data and results, all of which are labeled.

The code for mapping in Python can also be found on the AguaClara Google Drive. The files are within the same folder labeled "Regional planning 2015-2016". The folder also includes Excel sheets of data and results. The python code where written in ipython notebooks.

Discussion

The method for selecting favorable districts for AguaClara surface water treatment technologies was based on a weighting of selected requirements using an algorithm written in Matlab. The maps were created in Python: first, by using a shape file of each state and the coordinates of its districts, secondly by using the optimization numbers generated from the Matlab code. Python created the color-coded map coordinating the darkest color with the highest number. Using data gathered from the Ministry of Water Resources and from Census of India, the team created an algorithm based on a weighting system that the team determined. The weights reflected the importance of each constraint. For example, the data for surface water availability percentage had the highest weight because it is the most important consideration for an AguaClara plant. The final output was a color-shaded map, where the darkest shade was the most optimal solution.

The team calculated the optimized states and analyzed the highest scoring states for their performance in qualitative measures such as current political conflict and presence of local partners. The team then calculated the optimized districts within the top five states. The results for both the state analysis and district analysis are presented below, with state analysis first in Table 9.

Table 9: Data for top 5 states from results of Matlab Code

S.No	**State Name**	**Total % population dependent on surface water**	**poverty ratio**	**Annual Rainfall (mm)**	**HDI**	**% rural population in state**	**local partner**	**% rural population in India**
1	Meghalaya	31.93675	12.53	2354	0.573	0.799301558	0	0.002845
2	Sikkim	14.245967	9.58	1427.8	0.573	0.748470709	0	0.000548
3	Rajasthan	8.40089	16.05	180.1	0.434	0.751298706	1	0.061791
4	Himachal Pradesh	1.617122	8.48	417.5	0.652	0.899695277	0	0.00741
5	Kerala	4.146106	9.14	1647.4	0.79	0.613293279	0	0.02096
6	AVG(of all states)	9.13	22.49	698.22	0.52	0.6954084669	-	0.03566

The top five states from the results of the Matlab code were Meghalaya, Manipur, Mizoram, Jammu and Kashmir, and Sikkim. However, the team could not find data for the districts of the northeast states of India such as Manipur, Mizoram, Jammu and Kashmir because these are very small states with little public data. Due to feasibility and lack of resources, these states were discarded as sub-optimal. Therefore, the top five states the team decided to analyze at a district level were Meghalaya, Sikkim, Rajasthan, Himachal Pradesh, and Kerala.

Meghalaya has very high dependence on % population dependent on surface water and lower than average (average is about 22.5) poverty ratio, its rainfall is very high, HDI is above average (average = 0.518), %rural population in state is also above average (average = 0.695). The only thing that Meghalaya is lacking is availability of a known local partner and %rural population in India. However, the team put a relatively low weight on local partner availability since it was based on local partnership availability on places that have been researched by AguaClara LLC and there is a reasonably good possibility that local partners are available, just not yet contacted. The low % rural population in India was due to the fact that Meghalaya is a very small state (8,772 mi^2). Despite its small size, the data indicated that there was a good chance that there might be a village or town within Meghalaya that will be a good site for AguaClara to expand to.

Sikkim is similar to Meghalaya in that it has higher than average total % population dependent on surface water, annual-rainfall, HDI, % rural population in state, lower than average poverty ratio and has no local partners as of yet. The only difference is that unlike Meghalaya, Sikkim has a lower than average % rural population in India but this is only due to the fact that Sikkim is a smaller state. As mentioned previously team members are optimistic about chance of finding local partnership and have thus put a relatively low weight of 0.15 for local partnership availability. Even though it is a small state, the data seems to indicate that there is a good possibility that there will be many towns/villages for AguaClara to expand to in Sikkim.

Rajasthan and Kerala produced similar results in all aspects that the team put relatively high weight on. Both these states had less than average total % population dependent on surface water, which would be worrying. However, since the average without Jammu and Kashmir, Manipur, Meghalaya, Mizoram, and Nagaland is 4.01, both Rajasthan and Kerala are actually above average. They both had less than average poverty ratio. Both states had over 60% rural population. They showed differences in other areas but since they performed well on the constraints that the team thought was important, it made sense to analyze them at a district level. Kerala has a very low poverty ratio which is the second highest weighted constraint. Since most states in India have poverty ratios above 10 percent, this factor makes Kerala a viable site. Besides the low surface water dependence, Kerala performed well in the other constraints and so the team thought Kerala would also be worth analyzing at the district level.

Despite the fact that Himachal Pradesh came fourth in the results of the Matlab code, it has a low % population dependent on surface water which is strange since the team put the highest weight on this constraint. The state does, however, have a high percentage of surface water availability (Figure 3). The discrepancy might be because of the absence of surface water treatment plants in this region, and therefore, Himachal Pradesh is still a viable possibility for the expansion of an AguaClara plant.

Once the analysis for the state level optimization was completed for viability, the team began to research the results of the district level optimization. Unsurprising correlations were found through the results: three of the top five districts in the optimization are located in Meghalaya, which was the top state optimized on the state level program. This result speaks to the credibility of the program and the sources used for data extraction, because the top optimized districts are within the top optimized state. The top five results of the optimization are presented in Table 11.

Table 11: Data for top 5 districts from results of Matlab Code

S.No	**District Name**	**Total % population dependent on surface water**	**poverty ratio**	**Annual Rainfall (mm)**	**HDI**	**% rural population in state**
1	South Garo Hills	59	45	2459	0.48	0.89
2	Chamba	10.5	54.1	1117	0.43	0.93
3	Barmer	13.2	45	243	0.58	0.92
4	East Garo Hills	44	55.9	2554	0.41	0.85
5	West Garo Hills	30	53	2459	0.57	0.88
6	AVG(of all districts)	8.1	23.6	1179	0.61	0.78

The three top districts in Meghalaya are South Garo Hills, East Garo Hills, and West Garo Hills. South Garo Hills is the top optimized district. It is the least populated district in Meghalaya and its administrative district is divided into four blocks, which could be advantageous because its blocks will not compose many villages (there are on the order of 600 villages in this district). It is a viable choice for the top district in India because it has a very high percentage, 60 percent, of people dependent upon surface water; a poverty ratio around 50 percent, which is the highest in the area, and a relatively high HDI index. East Garo Hills and West Garo Hills are also obviously viable districts, though they have different characteristics. East Garo Hills and West Garo Hills both have very high population relative to the state: over 1000 villages each. East Garo Hills has a 30 percent population dependent upon surface water and a poverty ratio above 50 percent, while West Garo Hills has a 44 percent dependent upon surface water population and also has a high poverty ratio. These two most important of the characteristics speak to the viability of these two districts as well.

The other three districts in the top five are Chamba and Barmer. Chamba is located in the state of Himachal Pradesh. It is situated on the bank of the Ravi River, which is a major tributary of the Himalayan Indus River, making the location ideal for a water treatment plant. However, it does still have a low percentage dependent upon surface water, which may be for several reasons: chiefly, that the treatment is not available, and thus the nearby surface water cannot be safely used, which would make an AguaClara plant even more ideal for the district. Chamba also has a high poverty ratio and high rural population.

The best way to further investigate Chamba's viability is to determine exactly why the river is not the chief source of water for the district; if the reason is that they can't properly treat the water, then an AguaClara plant would be welcomed.

The last district in the top optimized five districts is Barmer, which is located in the state of Rajasthan. Barmer is the fifth largest district in India and includes Rajasthan's capital city, Barmer City. Despite its large population size, it has a high poverty ratio. Barmer recieved a high ranking in the optimization program because of its large percentage of rural population and high poverty ratio, but it does have a low percentage dependent upon surface water and its geography is primarily desert, with one or two large rivers running through. At first glance Barmer would not be the most ideal out of the top five list, though further investigations would determine if villages near the major rivers have strong ideal characteristics.

Conclusions

The semester has been spent gathering, analyzing, and processing large amounts of data sets. The optimization program is a solid deliverable because it can be used and edited by future teams. The program, for this semester, yielded a list of optimal states, and out of the top five optimal states, another list of optimal districts from those states. The top five states: Meghalaya, Sikkim, Himachal Pradesh, Kerala, and Rajasthan. Three of the top five districts are in Meghalaya, which was number one on the list of optimal states. Therefore, the team recommends Meghalaya as the first step in future work's analysis of India. The team recommends AguaClara LLC make contacts in Meghalaya and analyze feasibility from the ground, which is the missing link.

Future Work

An ideal step is to refine the optimization program written in Matlab so that it can be transferred to Python. Python is more ideal because it can produce a map automatically, whereas Matlab's program is only able to produce a bar chart and then a map must be generated in GIS or Python (the inputs to Python can be the results of the optimization program in Matlab). Alternatively, a programmer more skilled in Matlab may be able to generate a map directly from Matlab; however, the current team did not have that knowledge. In terms of choosing an AguaClara site in India, there are other future steps that must be taken to choose an AguaClara site. The optimization program is a good deliverable because it can be adapted and used in other areas, given the data on any area; however, the program was only able to narrow the search in India down to the district level, because it would be almost impossible to find enough data on the village or town level to generate usable results. Future teams must begin to look at the non-quantifiable factors of the districts presented in this report and analyze the viability of taking an AguaClara plant there. Future teams may also reach out to local partners in the optimized districts to analyze the feasibility of working in the area, and from there, narrow down the districts into the town and village levels. A very important area of future work is to obtain data for water turbidity in the surface water around these districts.

References

[1] Union territory. URL https://en.wikipedia.org/w/index.php?title=Union_territory&oldid=717195836. Page Version ID: 717195836.

[2] Governance models for community water systems: The case of AguaClara. URL https://confluence.cornell.edu/download/attachments/77267239/Governance%20Models%20for%20AguaClara.pdf?version=1&modificationDate=1342785623000&api=v2.

[3] AguaClara LLC. Community prerequisites. URL https://confluence.cornell.edu/display/AGUACLARA/Community+Prerequisites.

[4] Census of India. HH-6: Households by main source of drinking water and location. URL http://www.censusindia.gov.in/2011census/Hlo-series/HH06.html.

[5] Department of Drinking Water and, Sanitation, and Department of Drinking Water and. Strategic plan – 2011- 2022. URL http://mdws.gov.in/sites/default/files/StrategicPlan_2011_22_Water.pdf.

[6] Desai, Sonalde and Vanneman, Reeve. India human development survey-II (IHDS-II), 2011-12. URL http://dx.doi.org/10.3886/ICPSR36151.v3.

[7] Ankita Gandhi, Chandan Kumar, and Partha Saha. India human development report 2011: Towards social inclusion. URL http://www.iamrindia.gov.in/ihdr_book.pdf.

[8] Sonakshi Hudda. River pollution: Causes, actions and revival. URL http://www.janhitfoundation.in/pdf/booklet/river_pollution_causes_action_and_revival.pdf.

[9] National Rural Drinking Water Programme. Contamination wise report. URL http://indiawater.gov.in/IMISWeb/DataEntry/RWS/AAPFormat/AAPFormatReport.aspx?FN=1&Rep=Y.

[10] Office of the Registrar General. Households access to safe drinking water. URL http://planningcommission.nic.in/data/datatable/data_2312/DatabookDec2014%20222.pdf.

[11] Kenichi Victor and Nishikawa Chavez. Introducing AguaClara: The process of establishing a pilot plant in the state of chiapas. URL https://confluence.cornell.edu/download/attachments/77267239/Introducing%20AguaClara%203.0.docx?version=1&modificationDate=1454943630000&api=v2.

[12] World Bank. Willingness of households to pay for improved services and affordability. URL http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2008/07/23/000333037_20080723000750/Rendered/PDF/447900PP0P09411413B01PUBLIC10PAPER1.pdf.

Semester Schedule

Task Map

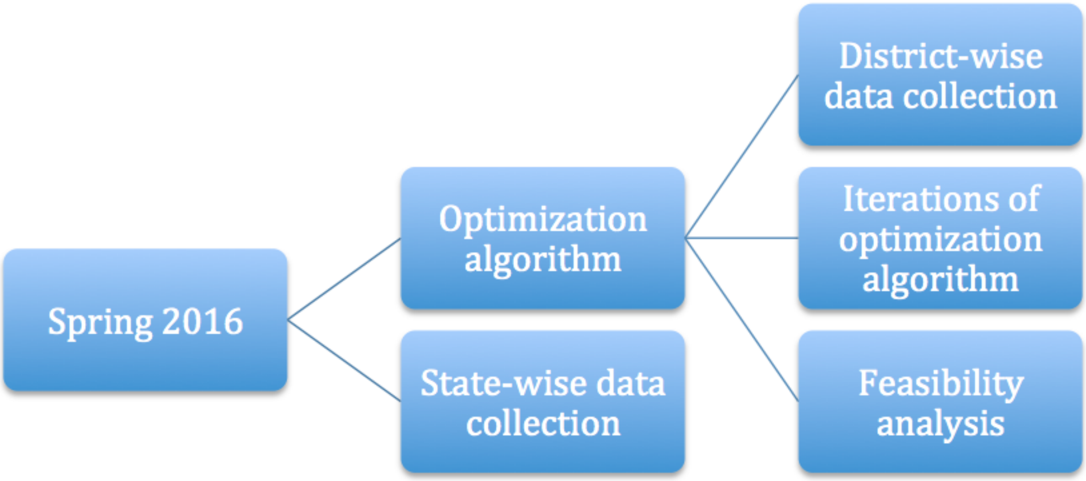


Figure 17: Regional Planning Task Map

Task List

- **Feb 26: Data collection (physical and social factors)-completed**
  1. Data Collection
     (a) Physical Factors
        i. Gather rainfall data and filter locations to only places where there is above 1000mm/year.
        ii. Look at map and look at proximity to rivers (see if there is a river nearby).
        iii. Gather data regarding surface water availability so that the team can focus on places where there is high surface water availability.
     (b) Social Factors
        i. Gather population data and narrow down to towns with population between 5,000 and 25,000, based on ease of implementation and past experience.
        ii. Gather poverty data. This provides insight into likely places where people will be able/willing to pay the extra water tariff.
        iii. Gather Projected Demand. Need a reasonable amount for demand for clean drinking water.

    iv. Gather data for percentage of population with access to safe drinking water. We are looking for a place with a relatively low percentage, but correlated with a high percentage of people who depend on surface water for drinking.

- **March 11: Visuals, Additional Data Collection, Intro to Political Conflicts-completed**

    1. Locate natural water bodies in India. Graph damns/river onto map. Also present summary of data on percentage use of surface water for drinking water vs. other sources (well water).
    2. Present annual rainfall summaries per state; in google map layout, or in written chart.
    3. Willingness to pay - present summary of data on percent access to drinking water, and percent poverty information, and median household income.

- **April 8: Introduce optimization program-completed**

    1. Decide the basic idea for code (Input data, output color coded map representing range from least likely to most likely).
    2. Decide coding language, program.
    3. Figure out basic algorithm- introduce idea of weights. Goal is for user to enter weights and program to output the most optimal areas based on the user's inputs.

- **April 22: Matlab optimization-completed**

    1. Simple matlab code for state level data analysis to get bar chart of results. The program will yield the most optimal states in India (out of 28).
    2. Introduce idea for normalizing data before adding weights, necessary because the scales are different.
    3. Water Tariff Problem research : $ may be too much.
    4. Experiment with the weighting system, explain why the weights we chose are logical.

- **May 10: District level optimization-completed**

- Using the top five states from the optimization, now do the same process on the district level.

- Consider non-quantifiable factors such as political conflict when analyzing the "most" optimal states (program only considers quantifiable data).

- Fine-tune code so that it may be easily understood and used by future teams.

Report Proofreader: Lauren