1. Conceptual Model

In the following section, the description of the model will be given based on the Overview, Design concepts, and Details (ODD) protocol by Grimm *et al.* (2006). However, full implementation details, source code, model results and models for each of the experiments are available at www.css.gmu.edu/cholera.

1.1. Overview

1.1.1: Purpose

The purpose of the model is to explore the spatiotemporal dynamics of the spread of cholera, which is caused by the interaction of human (host) with their environment and significantly extends the basic model of Hailegiorgis and Crooks (2012). We utilize an ABM for this purpose as such an approach is most suitable for a developing an understanding of the system under investigation where assumptions about processes and interactions can be explored in a dynamic environment (Kelly *et al.*, 2013). As with all models, however, a number of simplifying assumptions have been made to covert the complexities of reality into a problem which can be modeled (Batty and Torrens, 2005), which we detail below.

1.1.2: State Variable and Scales

The model focuses predominantly on the spatial spread of cholera in the Dadaab refugee camps. We have tried to stylize agents' behavior by incorporating their daily routine that may happen within the Dadaab refugee camps so that we can have better understanding of the dynamics of cholera transmission. We represent the environment using geo-referenced spatial data. Figure 1 shows the Unified Modeling Language (UML) diagram of the model.

The main agent in the model is the refugee agent who represents an individual refugee who lives in the Dadaab refugee camps. A refugee agent has family and a fixed home location. Agents of the same family cooperate and share resources. Agents are instantiated with different attributes that contribute to their heterogeneity. Agents differ in their personal characteristics (e.g. age, sex), social ties (e.g. number of family members and friends), their body immunity type (symptomatic and asymptomatic), and goals and priorities. Behaviorally, agents are mobile and purpose-oriented. They determine a specific activity (goal) at a given time, depending on their priorities, and move towards it to fulfill their satisfaction. The properties of the refugee agents are shown in Figure 2.

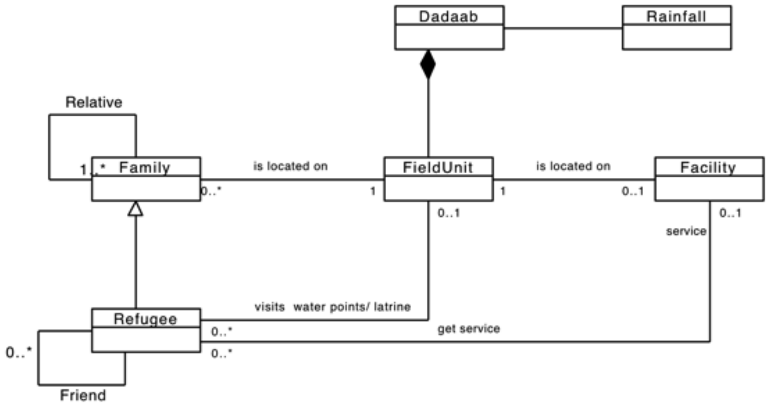


Figure 1: High level representation of the model using UML diagram.

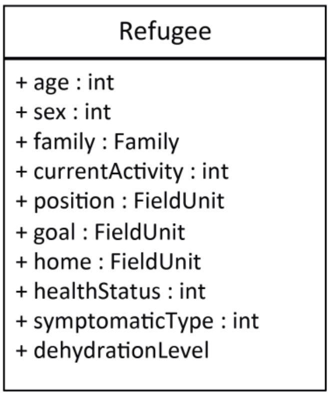


Figure 2: Refugee agent properties

In this model, we relate activities with facility locations. We consider nine types of activity locations that we assume are important in refugee contexts. These are: location of agents residence (i.e. home), school, water point (either borehole or river point), religious center (e.g. mosque), market, food distribution center, health center, latrine (either at home or on field), agent's friend or relative house within the same camp, and agent's friend or relative house in other camps.

Refugee agents are considered as susceptible hosts. They are myopic agents who do not have knowledge to differentiate between clean and contaminated water. Hence they can easily be exposed to cholera infection if they ingest contaminated water. For simplicity, we did not specifically model cholera bacteria as an agent rather we use water flows and contamination as a proxy to model the spread of cholera.

The other component of the model is the environment, which is a representation of the Dadaab refugee camps. It has a spatial extent of 13.5 kilometers by 25 kilometers with a spatial resolution of 90 meters by 90 meters. The spatial resolution is equivalent with an average distance that human can travel in a minute (humans can travels about 5 kilometers per hour on average, which is about 90 meters per minute). The spatial extent covers all of the three camps sites (Dagahaley, Ifo and Hagadera) located around the town of Dadaab. The environment encompasses field units (i.e. cells), camp boundaries, houses (e.g. tents), facilities, infrastructure, and elevation. The field unit is the main unit of the environment in which all processes of the model take place. A field unit may hold up to 100 houses but can only hold a single facility (e.g. a hospital or school).

The camps boundary represents the bounding box of the three camps. Each camp comprises of houses, facilities and infrastructure. Houses are located in compartments separated by roads. There are many kinds of facility units in each camp (e.g. schools, offices, places of worship, water points, marketplaces, food distribution centers, etc.). Although the spatial location of each facility in the model is reserved for visualization purposes and for future versions of the model, in the current version, agent interaction is limited to the following facilities: schools, health centers, markets, water points, and mosques.

The infrastructure represents the road networks of the Dadaab refugee camps. All types of roads (primary roads, secondary roads, feeder roads, and trails) are represented as the same type. There is no cost or preference on the types of the road the agents chose; however, roads have a capacity that constrains the flow of traffic (agents) in a given time. The level of road capacity is represented as 'crowd parameter'. A road can only be occupied by a given number of agents as it is set in the crowd parameter at a given time. If the road ahead is crowded, agents should stay where they are and wait until the road is cleared. The elevation dataset represents the topography of the Dadaab refugee camps. It has the same spatial resolution and extent with the other datasets and mainly used as an input for modeling the runoff of water.

There are four temporal resolutions within the model: minute, hour, day, and week. Each step of the simulation represents a minute. An hour has 60 minutes. One day represents 24 hours and a week represents 7 days. At each initialization of the model the clock starts at midnight (0 minute, 0 hour, 0 day, 0 week). Agents' activities can be constrained by time of day and agents give attention on these time divisions in their decision-making processes. For instance, every day agents should fulfill their activities and spend the night in their houses.

1.1.3: Process Overview and Scheduling

In each time step, each agent makes decision to stay where they are or move to their goal based on their priorities (e.g. needing food or water, to attend school)[1]. Agent movement from home to goal determined by the time and their success at the goal as shown in Figure 3. Each agent moves towards its goal by first selecting the nearest road and then plans its route using an A* algorithm. Since an agent can be constrained by time, the success of the agent to reach its goal is not guaranteed. If an agent is successful and reaches its goal, the agent might return to home if the agent accomplishes its activity or stays for a while until the agent is able to accomplish its activity before the agent returns to home (case 1). If the agent is too late to accomplish its goal, the agent will return home before reaching its goal (case 2). In some instants, an agent might consider choosing another goal before going back to its home (case 3). In the current version of the model, this will only happen when the agent fails to collect enough water from their nearest water point (e.g. the water point has run out of water) and are therefore forced to search for other water points to achieve its goal.

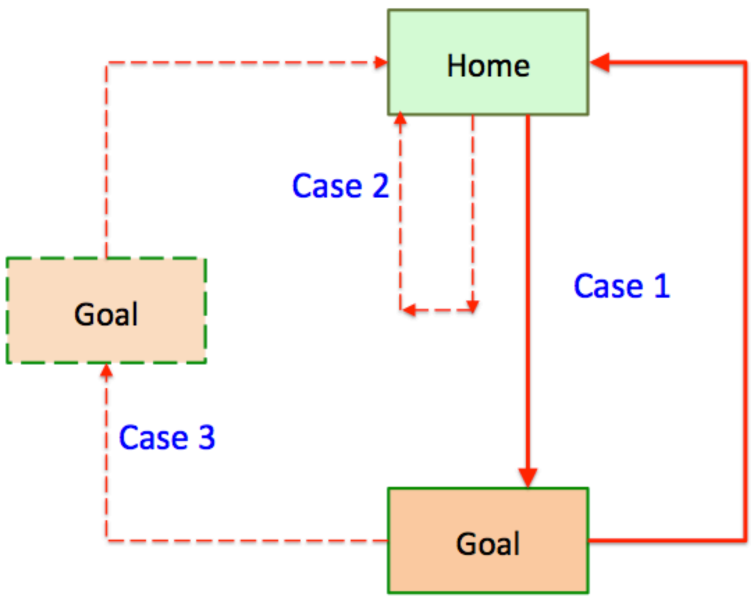


Figure 3: Agent's movement options between its home and its goal.

[1] It needs to be noted that data on daily activities within refugee camps is sparse, to say the least. Here we make several assumptions based on Maslow's "Hierarchy of Needs", in essence, first agents must satisfy their physiological needs such as food and water, before seeking out other activities (Maslow, 1943) and human mobility studies more generally (e.g. Gonzalez *et al*., 2008).

1.2: Design Concepts

Observation: The visualization window of the model as shown in Figure 4 portrays the extent and dynamics of cholera. We monitor the progress of cholera as agents portray different colors depending on their health status. At the global level, we monitor the following statistics: total number each activities, total number of susceptible agents, total number of exposed agents, total number of infected agents, total number of recovered agents, both at global level as well as at camp level. We also monitor the rainfall and surface runoff on the environment. These observations will be further elaborated in Section 4. With respect to *Interactions,* the dynamics of the model are driven by the interactions of the agents with their environment. Agents of the same family share resources (e.g. water), they also interact with other agents who are their relatives or friends, Relatives are given to them at the initialization of the model while they select friends randomly at runtime based on proximity.

Sensing: Agents are assumed to know the nearest facilities, their relatives and friends locations. They also know how to navigate along the road network towards their goal. The environment reacts to rainfall and creates surface runoff via the elevation gradient (as will be discussed in Section 1.3.3.2). With respect to *Emergence,* we anticipate the spatial distribution of cholera to be an emergent phenomenon. We class emergence as the large-scale outcome from simple interactions among individual agents (Kelly *et al.*, 2013). Although users of the model can infect a specific number of boreholes to observe the spread of cholera due to the probability of drinking contaminated water, the spatial distribution and spread of cholera from one camp to another emerges due to the simple interactions of individual agents with each other and their environment at the micro level as we will demonstrate in Section 4 of the main paper. *Stochasticity* within the model is seen in several processes. These include the agent selection of certain activities (e.g. going to market) or selecting a friend to visit which are drawn from a normal distribution. Assignment of some values to agents (e.g. to be a symptomatic or asymptomatic agent) at the initialization of the model is also carried out stochastically based on a parameter set by the user (see Section 1.3).

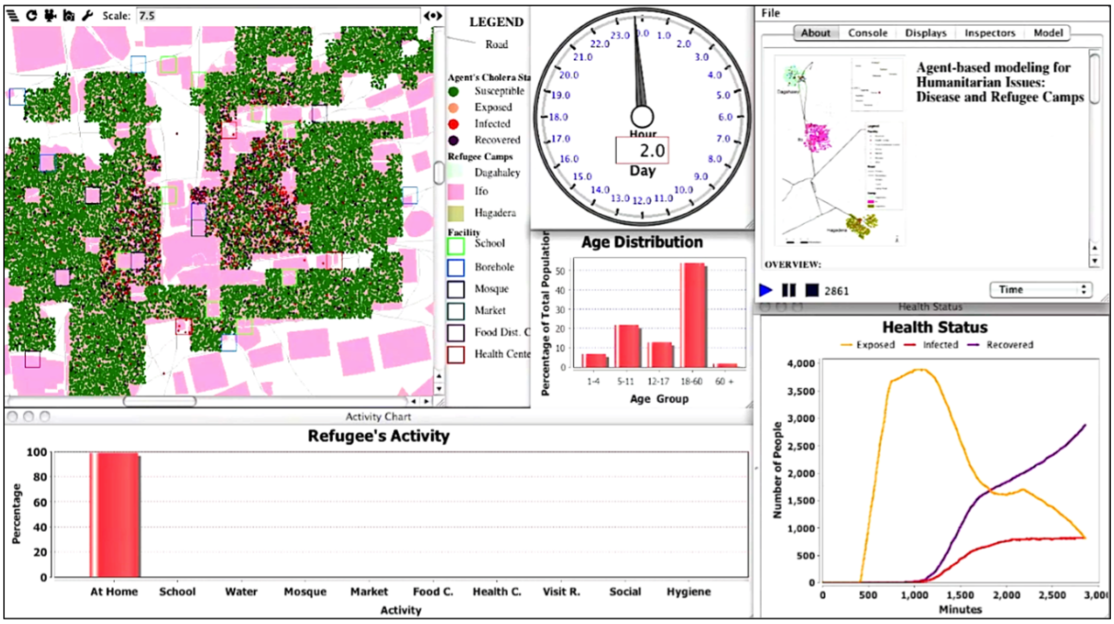


Figure 4: Graphical User Interface of the model. Clockwise from top left, spatial environment where agents are portrayed by health status, legend, clock showing simulation time passed, age distribution of agents, model console, charts for health status and refugee's activities

1.3: Details

1.3.1: Initialization

The initialization of the model relies on demographic and spatial data of the study area. Most of the parameter values were calibrated based on relevant literatures. The model gives the flexibility to users to run different experiments by changing some of the default values. For instance, users can experiment by increasing or decreasing the number of agents in the simulation. All the default values and the parameters are summarized in Table 1. Where possible all parameters were derived from existing literature. Where ranges are used, we apply a normal distribution. In cases where no such data exists, we estimate values by proxy. For example, dehydration rate is estimated by the daily consumption rate of the agent. We assume that agents need to fulfill their water demand daily, hence we calculate the dehydration rate by dividing the minimum daily water use by the minutes in a day (dehydration rate = Daily water consumption / (24 hours * 60 minutes). Similarly, the maximum occupancy threshold is estimated using the pixel area (90m by 90m) by assuming a family with average size of 4 to occupy 9m by 9m.

Table 1: Input parameters and variables.

Parameter	Default Values	Reference
Agent		
Initial number of agents	50,000-500,000	User settable
Daily water consumption	4- 15 liter/day	UNHCR (2011a)/CARE (2012)
Dehydration rate	0.003 liter/ minute	Authors estimation
Ratio of asymptomatic to symptomatic agent	3:100	King *et al*. (2008), Harris *et al*. (2008)
Rate of return to susceptible	0.0001%	Authors estimation
Maximum distance from home to open field latrine	2 km	Authors estimation
Maximum occupancy threshold	100 families per cell	Authors estimation
Road crowed threshold	1000 people per grid cell	Authors estimation
Maximum number of relatives	15 families	Authors estimation
Probability of guest contamination rate	0.5%	Authors estimation
Healthy person body resistance level	1.0	Authors estimation
Health depreciation rate	0.001/ minute	Nelson *et al*. (2009)
Clean water source preference probability	70%	Authors estimation
Overall ventilated improved pit latrine coverage	60%	UNHCR (2011a)/CARE (2012)
Mortality	Up to 50 % of infected (untreated). Up to 1% of infected (treated)	Nelson *et al*. (2009)
Minimum number of Vibrio to cause cholera infection	10000/ml	Franco *et al*. (1997) / Nelson *et al*. (2009)
Cholera infection duration	12-72 hours	Nelson *et al*. (2009)
Infected person fluid loss	1000ml/hr	Nelson *et al*. (2009) / Codeço, (2001)
Vibrio per gram of stool of infected person	10^7 -10^9/ml	Nelson *et al* (2009) / Franco *et al*. (1997)
Vibrio per gram of stool of uninfected person	10^2 to 10^5/ml	Codeço, (2001)
Minimum goal utility threshold	0.3	Authors estimation
Facilities		
Health facilities capacity	1000 patients /day	
Borehole maximum capacity	2 liter/people/day	UNHCR (2011a)/CARE (2012)
Borehole discharge rate	80% of the maximum borehole capacity	Authors estimation
Rainfall		
Rainfall absorption rate	10 mm/ minute	Authors estimation
Duration	25 minute/ day in a rainy day	Authors estimation
Rainfall amount	Daily rainfall (mm) from data	AccuWeather (2013)

1.3.2: Inputs

The input spatial dataset were generated from publicly available data sources. The camps information (camp boundaries, houses, facilities, and infrastructure) was processed from UNHCR (2011b) PDF maps and Google Earth (KLM format) from UNITAR, (2012). The PDF maps of each camp were converted into ESRI vector file using PDF Converter. The converted data was geo-referenced and edited in ArcGIS. The elevation dataset were generated using the 90m meters Digital Elevation Model (DEM), which is provided by CGIAR Consortium for Spatial Information (2012) GeoPortal. Demographic characteristics (e.g. age, sex, number of people per household etc.) for the agents is based on survey data from and UNHCR (2011b) and CARE (2012).

1.3.3: Submodels

1.3.3.1: Goal Selection

It is well known that agent-based decision making is a complex task (e.g. An, 2012; Kennedy, 2012). Therefore rather than taking an overly complicated approach. We make it as simple as possible. Agents determine their current activity based on their personal attributes (e.g. age, sex), and their current need. They also consider the time and distance into consideration when they make their goal choices. We make the assumption those agents who are less than 5 years old stay at home. They do not engage in any activities unless they are sick and need to visit the health facilities. They utilize water from the family reserve and use the home latrine. The other age groups can engage in any of the following activities depending on their needs (e.g. get water) or intentions (e.g. go to mosque).

While we model time in minutes, agents also have a planning step (i.e. the time it takes to carry out a specific goal) whereby an agent will select a specific goal from the set of goals. Agents will select a goal that gives them the highest utility. The next goal selection will be scheduled after the current goal has executed or aborted. The utility of a goal for an agent is given as follows:

$$U_i = \alpha_i . \beta(T)_i . \gamma_i \qquad \text{(eq. 1)}$$

Where U_i is the utility of goal i, α_i is set to a constant value for a specific goal selected from a uniformly distributed random number, β_i is a function of the time period T, which indicates the importance of the time period for a goal, and γ_i is the parameter indicating the importance of a goal. The time period T can be a specific time of a day, day/night cycle, or a specific day(s) in a week depending on the type of a goal. For instance, the selection of 'visit to relatives' goal will have a β value of 1 between 6:00 am to 12:00pm and 0 in other time (i.e. when the agent wants to sleep). The value of a goal G is as follows:

$$Gi = \frac{Ui}{\sum_{i-1}^{n} Ui} \qquad \text{(eq. 2)}$$

And the selected goal will be the one with the maximum G value as follows:

$$A_i = max_{1 \leq i \leq n}(G_i) \quad \text{(eq. 3)}$$

Where A_i is a goal with the maximum utility, G_i is the value of goal i, U_i is the utility of goal i, and $\sum_{i-1}^{n} Ui$ is the sum of all utilities for all of the goals. The A_i value should be greater than the 'Minimum Goal Utility Threshold' to be executed, otherwise the agent stays at home. The assignment of a facility location for the selected goal is based on proximity to the agent. In the model, agents give the highest priority to the nearest facility (i.e. minimum distance) when choosing between two facilities of the same kind. If two or more activities have the same utility, one of these activities will be chosen at random.

Goal Selection – School: Agents whose ages are between 5 and 18 are only considered as students. Out of the total number of agents in this age group, only 51% of them are attending school (CARE, 2012). Agents who are a student attend school between 8:00 am to 4:00 pm from the first day (Monday) to the fifth day (Friday) of the week. Agents will stay in the school until 4:00 pm before they consider returning to home or going to another location.

Goal Selection – Visiting Friends or Relative: Within the model, agents visit friends who live within the same camp (e.g. next-door neighbors). With respect to relatives, these can live in any of the three camps. In both cases the agent selects at random a specific agent's home location and makes the visit its priority and moves towards that agents' home location. If the agent reaches its goal and is dehydrated, the agent can drink water from its hosts' home location.

Goal Selection – Religious Center: In this model, we only consider one type of religious center. We make our assumption based on both the statistical and spatial data. According to the camp population statistics report from UNCHR (2012), about 95% of the total refugees in the Dadaab camps originated from Somalia who primarily practice the Islamic religion. The spatial information also indicates that there are only mosques in the refugee camps as opposed to other religious centers. Agents within our model only visit mosques during in the main prayer times. There are five prayer times in 24 hours for the Islamic religion: Fajr (5:30 am),, Dhuhr (1:00 pm), Asr (4:00 pm), Maghrib (7 pm), Isha (8:00 pm). This time is an approximation to Kenya's time of prayers. We also consider the fifth day (Friday) as the main communal worship day of the week. In these prayer times or day or both, agents more often choose to visit the mosques rather than carrying out other activities. We also assume that older agents are more likely to visit the mosques more frequently than younger agents.

However, agents who are young and currently attending school could only consider visiting the mosques when not in school.

Goal Selection – Market: The market is one of the integral parts of refugee life. Although the main source of food comes from food aid, refugees also engage in different income generating activities such as petty trading and weaving, trading and exchanging goods (see Werker, 2007). This is especially the case in refugee camps like Dadaab, which have existed for many years and many socio-economic activities have flourished as most of the refugees have been here for a long time. In the model, the priority of visiting the market depends on age and sex. We assume that most of agents in the mid age group (18-46) are more active in economic activities. This is especially the case for women who may likely have a higher tendency of engaging in economic activities than men (Werker, 2007).

Goal Selection – Food Distribution Centers: In Dadaab, food distribution is managed by CARE. According to CARE (2012), the food distribution in Dadaab refugee camps is scheduled by cycle. According to CARE (2012), the food distribution within the camp is scheduled in order for each family to receive food every 14 days (CARE, 2012). Each refugee family visits the food distribution centers twice in a month and collects about 9 kg of food per family member, which is equivalent 2,100-kcal/day/person. The model follows the same food distribution schedule. The date of distribution is randomly assigned to each of the agent families at the initialization of the model and each family knows when to visit the food distribution centers. Agents will only visit to the food distribution centers to collect their ration on their scheduled date. On that date the agent will give the highest priority for visiting the food distribution centre over all other activities. Any one of the agents within the household can visit the food centre. To simplify the model, we let the agents to satisfy all their food needs from food revived from food distribution centers.

Goal Selection – Health Center: Within the current version of the model, our agents only visit health centers when they are infected by cholera. Any infected or sick agent will place visiting the health facilities as its highest priority on its lists of daily activities. Health centers have a limited capacity with respect to the to treatment patients. Agents who get treatment will recover. Agents who are not treated on their first visit will return the next day until they are successful. However, agents whose health is deteriorating may end up dyeing before receiving the necessary treatment.

Goal Selection – Water Points: In this model, two types of water sources are considered. The first one is from boreholes or tanks, which are mainly delivered and administered by humanitarian organizations. In the current version of the model, we assume that water from this source is considered as clean unless pollution is introduced exogenously. The second source is rainfall. Agents can utilize surface water that might be accumulated in ditches or holes after the rain. We assume that water from this source can easily be contaminated by surface runoff, mainly due to feces. In which case, water pollution is taken place endogenously through surface runoff and feces accumulation.

Within the model water is collected by any member of the family and is equally shared with all family members. Agents fetch water from a water source and accumulate the water in the family bucket and utilize it from there. Agent should fulfill their water requirement each time they visit a water point. The maximum daily consumption of each agent is 15 liters per day (CARE, 2012). This amount includes all possible uses of water: drinking, cooking and cleaning. We consider two types of water utilization: water for drinking and water for all other use. This distinction could introduce complexity in the behavior of the agents and help us to explore the refugees' exposure to contaminated water and its consequences. There is a notion of dehydration in the model. Each time step, agents check both their body water level and their family water levels to make decision whether to fetch water or not. Agents fetch water from the nearest water sources and utilize the water. Agents can fetch up to 25 liters in a single visit depending on the availability of water from the sources. If the source is dried up or very crowded, agents will visit other nearby water points.

Goal Selection – Latrine: We consider two types of sanitation facilities in the model: ventilated improved pit latrine (VIPL) and open field latrine (OFL). The VIPL is viewed as safe as they can easily contain waste. The number of VIPL in the model is set as a parameter and the user can change the value. Agents may or may not have access to VIPL. If they have do, they utilize it each time depending on their need. However, if they do not have one, an agent will utilize the nearest open field as latrine. Their disposal (i.e. waste) will stay in the environment and can be taken up by runoff and could cause contamination of rainfall water. In the model, infected agents will visit any latrine more frequently than other agents.

1.3.3.2: Hydrology

Within our model we use rain as a proxy for climatic events as previous research has shown that there are strong correlations between seasons and outbreaks of cholera (e.g. Reiner *et al.*, 2012). Secondly, as noted above water is one of the main methods for cholera transmission (Codeço, 2001). In this model, we use rainfall both as source of water for the agents as well as a carrier of pollutants (disposal and feces). We utilize elevation surface data (DEM) as shown in Figure 5 to model the flow of rainfall over the ground in the model. Table 2 provides some statistics with respect to elevation and slope characteristics of each camp. We apply a simple hydrologic model that only consider elevation gradient to model surface runoff. Rainfall flows downhill according to the elevation gradient. As the water flows from uphill to downhill, it carries pollutants. The tration of pollutants in the water depends on the amount of pollutant per volume of water in the field unit (parcel). As the model has been purposely kept simple, we have not considered issues such as subsurface flows or soil moisture such as in the work of Bithell and Brasington (2009) or Beven and Freer, 2001. While these will probably impact the spread of water and pooling it was considered beyond the scope of this study.

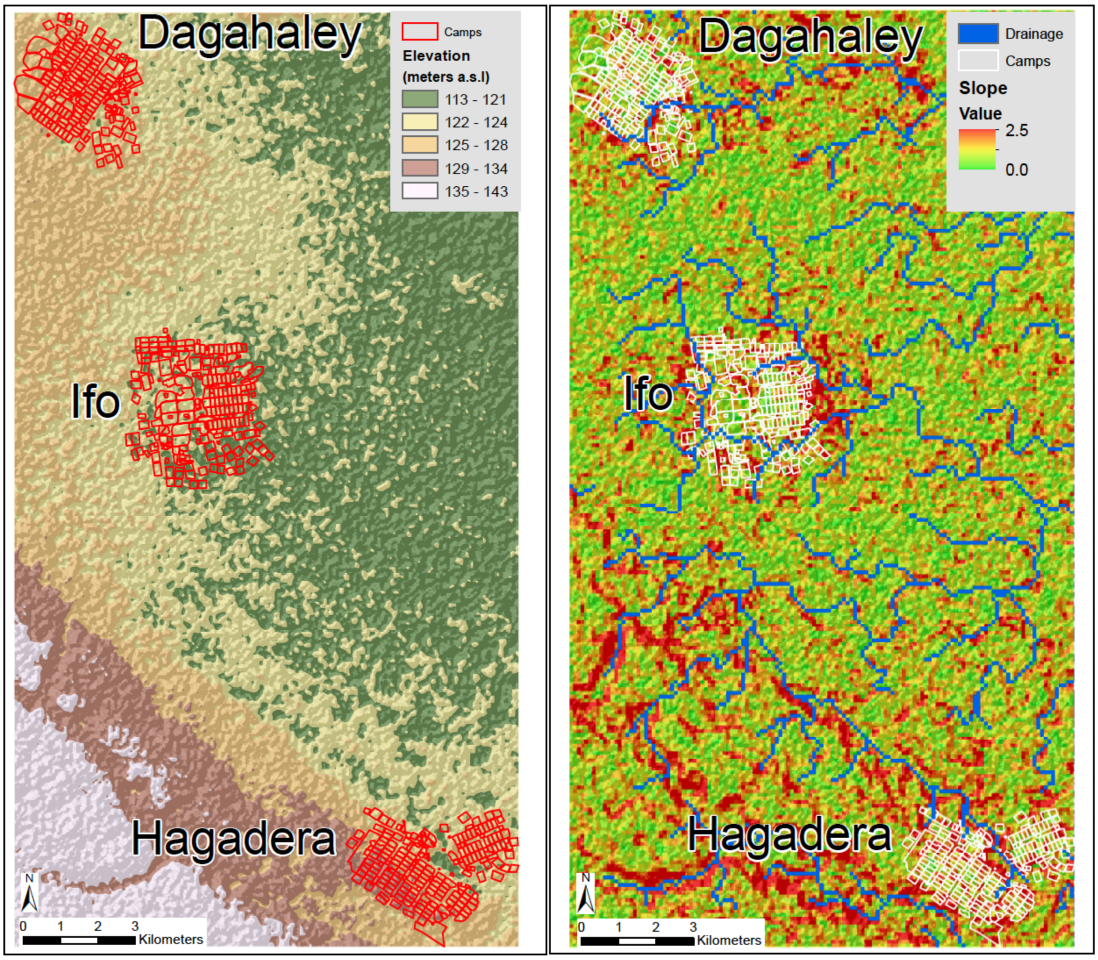
The dynamics of surface runoff is modeled using a cellular automata technique. If it's a rainy day, rain falls on the environment and each field unit (cell) gets equal amount of rainfall. At each time step, each cell will check if it has water to flow to its Moore neighbors. If the cell has water and the neighboring cell is a sink (i.e. at a lower elevation), it will give the water until it fills the sink depending on the elevation and water gradient. If the volume of water is less than the sink, all the water will flow to the sink. However, if the volume of water is greater than the sink, the water will flow until the two cells reach to equal level. Water flow is treated as follows:

$$V_t^c = \begin{cases} V_{t-1}^c + \sum_{i=1}^{n} \partial^i V_t^i \; - \boldsymbol{\zeta} + R_t^c, & if \;\; h_c < \; h_i \\ V_{t-1}^c - \mathrm{W}_t^c - \boldsymbol{\zeta} + R_t^c, & if \;\; h_c > \; h_i \\ V_{t-1}^c - \boldsymbol{\zeta} + R_t^c, & if \;\; h_c = \; h_i \end{cases} \quad \text{(eq. 4)}$$

Where V_t^c is the volume of water of central cell at time *t*, V_{t-1}^c is the volume of water at previous time step, V_t^i is the volume of water of the neighboring cell at time step, ∂_t^i is a parameter indicating the proportion for water flow from a neighboring cell, W_t^c is the outflow volume of water from the central cell. The model contains a consistency condition for water flow, by matching the outflow W_t^c to the total volume of water that is available in the central cell. ζ is the volume of water lost through absorption (which is a constant for all cells), R_t^c is the volume of rain at time t for the central cell, h_c and h_i indicates the height of central cell and neighboring cell respectively, n indicates the number of neighboring cells.

The height h includes both the elevation and the depth of the water level as:

$$h = elev + \frac{Volume}{Area} \quad \text{(eq. 5)}$$



A B

Figure 5: Digital elevation data used within the model (A) and the resulting slope differentials (B) with the refugee camps superimposed on top of them.

Table 2: Elevation and Slope statistics per camp.

Elevation (M)	Slope (Degrees)

Camp	MIN	MAX	RANGE	MEAN	STD	MIN	MAX	RANGE	MEAN	STD
Dagahaley	118.24	128.96	10.72	124.90	1.30	0.02	1.16	1.14	0.38	0.22
Hagadera	118.22	132.93	14.71	126.14	2.76	0.02	2.05	2.03	0.52	0.28
Ifo	115.96	127.93	11.98	121.64	1.31	0.01	2.02	2.01	0.42	0.26

Water that accumulates in the sinks can be considered as puddles or sources for drinking. As rainfall flows from cell to cell, the total concentration of pollutant in the water changes accordingly. The flow of pollutant also depends on the volume of water flow which is given as:

$$P_t^c = P_{t-1}^c + \sum_{i=1}^{n} \vartheta^i P_t^i \; - Q_t^c \text{ (eq. 6)}$$

Where P_t^c is the total amount of pollutant of the central cell at time t, P_{t-1}^c is the amout of pollutant at the previous time step, P_t^i is the amount of pollutant of the neighboring cell at time t, ϑ is a parameter indicating the proportion for water flow from a neighboring cell, and Q_t^c the pollutant loss through outflow. With respect to the surface boundary conditions, within the model we assume rainfall is evenly distributed throughout the study area and when water reaches the edges of the domain it does not accumulate but rather flows out and carries the pollutants with it. In addition, we assume that at the start of the simulation, the study area does not contain any bacteria in the soil.

1.3.3.3: Cholera SEIR Model

Cholera epidemic is mainly caused by *V. cholerae* bacteria and the bacteria can survive in aquatic reservoirs or the host's intestine. Nelson *et al.* (2009) described the dynamic nature of cholera by explicitly representing the interaction between host (human) and the environment (aquatic reserve), and the progress of the epidemic using Susceptible – Infected – Recovered (SIR) model (a common approach in many epidemiological studies e.g. Simoes, 2012; Tuite *et al.*, 2011; Augustijn-Beckers *et al.*, 2011). We extend their representation to Susceptible – Exposed – Infected – Recovered (SEIR) model to capture the time between ingestion of contaminated water and showing the symptom (i.e. the incorporation of exposed).

All the refugee agents are considered as susceptible hosts as shown in Figure 6. The infectious dose of *V. cholerae* in humans varies greatly depending on the bacterial strain and the host. In many cases, a bacterial cell concentration of 10^3/ml of water is necessary to infect the host (Nelson *et al.*, 2009). We assume that a susceptible agent who ingests contaminated water with a bacterial cell concentration of 10^3/ml of water or above will be become exposed to the cholera disease. An exposed agent stays as exposed for 12 to 17 hours before showing any sign or symptom (Nelson *et al.*, 2009[2]). The lag period depends on age, and body resistance of the agent. Infants show the symptom more quickly than adults (Harris *et al.*, 2008). After the lag period exposed agent will pass to infection phase.

We distinguished two types of infectious agents: symptomatic and asymptomatic. Symptomatic are agents who show symptoms of cholera and can die from the infection. Symptomatic infected agent spread *V. cholerae* through excretion of feces to the environment. Infected individual spread 10^9/ml of *V. cholerae* through excretion of feces to the environment and the bacteria can survive in the environment for long period of time (Franco *et al.*, 1997). Asymptomatic agents are 'silent shedders.' They shed 10^2 to 10^5/ml *V. cholerae* per stool to the environment without showing any sign of symptom (Franco *et al.*, 1997). In the model, asymptomatic agents immediately pass to the recovery stage while symptomatic agents stay in the infection phases until they get treatment and recovered. If they do not get treatment on time, they will die.

The extent to which the bacteria spreads to the environment depends on the type of latrine an agent uses. If the latrine is of OFL type, the bacteria are accumulated in the field and can be transported by rain. If the latrine is of VIP type, the bacteria will be contained in the latrine even during the rain. For simplicity, we did not consider contamination of water sources through seepage in this model unless it is exogenously introduced.

A recovered agent will stay as recovered for some time before becoming susceptible again (Nelson *et al.*, 2009). The rate of transition from recovered to susceptible can also be modeled by setting the parameters. A susceptible or recovered agent may also spread small amounts of *V. cholerae* through excretion of feces (Franco *et al.*, 1997).

V. cholerae is known for its remarkable versatility and resilience to stress. There is a growing literature which discusses how *V. cholerae* has evolved a complex stress management response system to survive in a hostile environment. When stressed by say starvation and low temperatures, *V. cholerae* enters into a Viable But NonCulturable (VBNC) state. *V. cholerae* cells in the VBNC state are living cells and can continue survive in the environment for more than 100 days (Li *et al.*, 2014; Chaiyanan *et al.*, 2001; Colwell, 2000; Wai *et al.*, 1999). They can be non-infectious in the VBNC state but regain virulence after

[2] However, the literature notes that infection rates can vary after ingestion and this is an area of debate. For further information see Hartley *et al.*, 2006.

resuscitation into culturable cells under suitable conditions. In our model, we applied this notion by assuming that if *V. cholerae* is excreted into the environment, it can stay alive in the environment by transforming into VBNC state until it gets into the host intestine through ingestion of contaminated water.

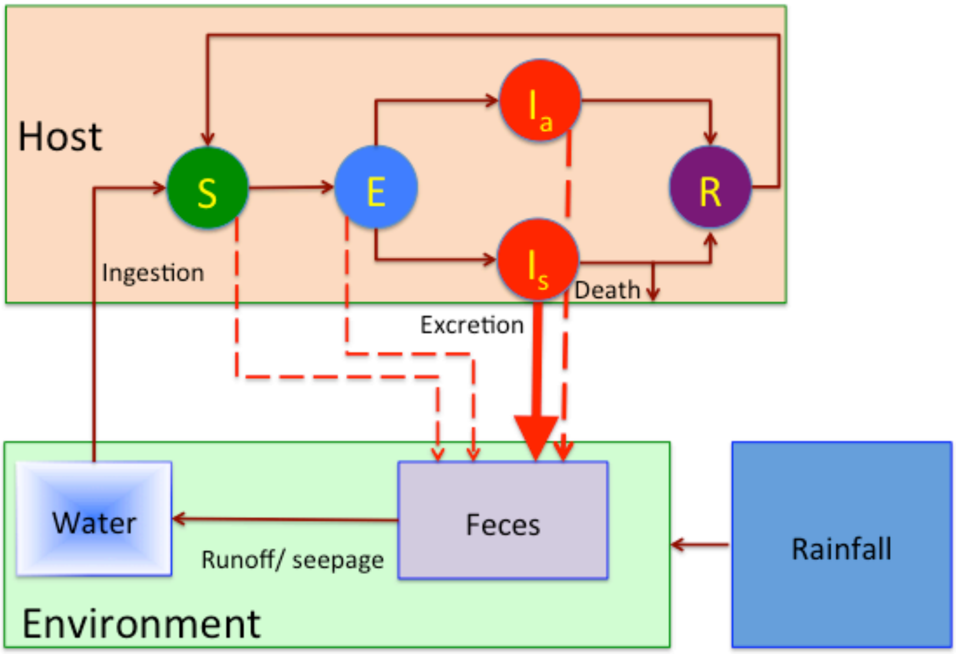


Figure 6: Cholera transmission through the interaction of host and the environment. The progress of cholera transmission is represented as SEIR model. S = susceptible, E=Exposed, Ia= Infected (asymptomatic), Is= Infected (symptomatic) R=Recovered.

1.3.4. Model Outputs:

The main outputs from the model include the number of people who are susceptible, exposed, infected and recovered per iteration of the model along with the spatial location of the individuals. Through such outputs we can trace the spread of cholera throughout the refugee camps.

2. References

AccuWeather (2013), Dadaab Weather, Available at http://www.accuweather.com/en/ke/dadaab/224830/month/224830?view=table [Accessed on July 19th, 2013].

An, L. (2012), 'Modeling Human Decisions in Coupled Human and Natural Systems: Review of Agent-based Models', *Ecological Modelling,* 229(25-36).

Augustijn-Beckers, E., Useya, J., Zurita-Milla, R. and Osei, F. (2011), 'Simulation of Cholera Diffusion to Compare Transmission Mechanisms', *Proceedings of the 11th International Conference on Geocomputation*, University College London, London, England, pp. 39-42.

Batty, M. and Torrens, P.M. (2005), 'Modelling and Prediction in a Complex World', *Futures,* 37(7): 745-766.

Beven, K. and Freer, J. (2001), 'A Dynamic Topmodel', *Hydrological Processes,* 15(10): 1993-2011.

Bithell, M. and Brasington, J. (2009), 'Coupling Agent-based Models of Subsistence Farming with Individual-based Forest models and Dynamic Models of Water Distribution', *Environmental Modelling & Software,* 24(2): 173-190.

CARE (2012), Horn of Africa Food Crisis, Available at http://www.care.org/emergency/Horn-of-Africa-food-poverty-crisis-Dadaab-2011/index.asp [Accessed on July 1st, 2012].

CGIAR Consortium for Spatial Information (2012), CGIAR Consortium for Spatial Information, Available at http://www.cgiar-csi.org/ [Accessed on July 1st, 2012].

Chaiyanan, S., Chaiyanan, S., Huq, A., Maugel, T. and Colwell, R.R. (2001), 'Viability of the Nonculturable Vibrio Cholerae O1 and O139', *Systematic and Applied Microbiology,* 24(3): 331-341.

Codeço, C.T. (2001), 'Endemic and Epidemic Dynamics of Cholera: The Role of the Aquatic Reservoir', *BMC Infectious Diseases,* 1(1), Available at http://www.biomedcentral.com/1471-2334/1/1.

Colwell, R.R. (2000), 'Viable But Nonculturable Bacteria: A Survival Strategy', *Journal of Infection and Chemotherapy,* 6(2): 121-125.

Franco, A.A., Fix, A.D., Prada, A., Paredes, E., Palomino, J.C., Wright, A.C., Johnson, J.A., McCarter, R., Guerra, H. and Morris, J.G. (1997), 'Cholera in Lima, Peru, Correlates with Prior Isolation of Vibrio Cholerae from the Environment', *American Journal of Epidemiology,* 146(12): 1067-1075.

Gonzalez, M.C., Hidalgo, C.A. and Barabasi, A.L. (2008), 'Understanding Individual Human Mobility Patterns', *Nature,* 453(7196): 779-782.

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S., Huse, G., Huth, A., Jepsen, J., Jorgensen, C., Mooij, W., Muller, B., Pe'er, G., Piou, C., Railsback, S., Robbins, A., Robbins, M., Rossmanith, E., Ruger, N., Strand, E., Souissi, S., Stillman, R., Vabo, R., Visser, U. and Deangelis, D. (2006), 'A Standard Protocol for Describing Individual-Based and Agent-Based Models', *Ecological Modelling,* 198(1-2): 115–126.

Hailegiorgis, A. and Crooks, A.T. (2012), 'Agent-based Modeling for Humanitarian Issues: Disease and Refugee Camps', *The Computational Social Science Society of America Conference*, Santa Fe, NM.

Harris, J.B., LaRocque, R.C., Chowdhury, F., Khan, A.I., Logvinenko, T., Faruque, A.S., Ryan, E.T., Qadri, F. and Calderwood, S.B. (2008), 'Susceptibility to Vibrio Cholerae Infection in a Cohort of Household Contacts of Patients with Cholera in Bangladesh', *PLoS Neglected Tropical Diseases,* 2(4): e221.

Hartley, D.M., Morris, J.G. and Smtih, D.L. (2006), 'Hyperinfectivity: A Critical Element in the Ability of V. Cholerae to Cause Epidemics?' *PLoS Medicine,* 3(1): 63-69, Available at doi:10.1371/journal.pmed.0030007.

Kelly, R.A., Jakeman, A.J., Barreteau, O., Borsuk, M.E., ElSawah, S., Hamilton, S.H., Henriksen, H.J., Kuikka, S., Maier, H.R., Rizzoli, A.E., van Delden, H. and Voinov, A.A. (2013), 'Selecting Among Five Common Modelling Approaches for Integrated Environmental Assessment and Management', *Environmental Modelling & Software,* 47: 159-181.

Kennedy, W. (2012), 'Modelling Human Behaviour in Agent-Based Models', in Heppenstall, A., Crooks, A.T., See, L.M. and Batty, M. (eds.), *Agent-based Models of Geographical Systems*, Springer, New York, NY, pp. 167-180.

King, A.A., Ionides, E.L., Pascual, M. and Bouma, M.J. (2008), 'Inapparent Infections and Cholera Dynamics', *Nature,* 454(7206): 877-880.

Li, L., Mendis, N., Trigui, H., Oliver, J.D. and Faucher, S.P. (2014), 'The Importance of the Viable But Non-culturable State in Human Bacterial Pathogens', *Microbial Physiology and Metabolism,* 5(Article 258): doi: 10.3389/fmicb.2014.00258.

Maslow, A.H. (1943), 'A Theory of Human Motivation', *Psychological Review,* 50(4): 370-396.

Nelson, E.J., Harris, J.B., Morris, J.G., Calderwood, S.B. and Camilli, A. (2009), 'Cholera Transmission: The Host, Pathogen and Bacteriophage Dynamic', *Nature Reviews Microbiology,* 7(10): 693-702.

Reiner, R.C., Kinga, A.A., Emch, M., Yunusd, M., Faruque, A.S.G. and Pascuala, M. (2012), 'Highly Localized Sensitivity to Climate Forcing Drives Endemic Cholera in a Megacity', *Proceedings of the National Academy of Sciences,* 109(6): 2033-2036.

Simoes, J. (2012), 'An Agent-Based/Network Approach to Spatial Epidemics', in Heppenstall, A.J., Crooks, A.T., See, L.M. and Batty, M. (eds.), *Agent-based Models of Geographical Systems*, Springer, New York, NY, pp. 591-610.

Tuite, A.R., Tien, J., Eisenberg, M., Earn, D.J., Ma, J. and Fisman, D.N. (2011), 'Cholera Epidemic in Haiti, 2010: Using a Transmission Model to Explain Spatial Spread of Disease and Identify Optimal Control Interventions', *Annals of of Internal Medicine,* 154(9): 593-601.

UNHCR (2011a), *Health And Nutrition Sector Update Dadaab Refugee Camps 15th to 21st October 2011, Epidemiological week 42 Update No. 14*, Nairobi, Kenya.

UNHCR (2011b), Kenya - Lagdera District Refugee Camp Overview, Available at http://data.unhcr.org/horn-of-africa/regional.php [Accessed on December 21st, 2011].

UNHCR (2012), Dadaab - World's Biggest Refugee Camp 20 Years Old, Available at http://www.unhcr.org/4f439dbb9.html [Accessed on July 1st, 2012].

UNITAR (2012), Horn of Africa Data, Available at http://www.unitar.org/unosat/horn-africa-data [Accessed on July 1st, 2012].

Wai, S.N., Mizunoe, Y. and Yoshida, S.I. (1999), 'How Vibrio Cholerae Survive During Starvation', *FEMS Microbiology Letters,* 180(2): 123-131.

Werker, E. (2007), 'Refugee Camp Economies', *Journal of Refugee Studies,* 20(3): 461-480.