SIMULATING DIFFUSION OF CHOLERA IN GHANA

JULIANA USEYA
February, 2011

SUPERVISORS:

Mrs. Ir. Ellen-Wien P. W. M. Augustijn-Beckers
Dr. Raul R. Zurita-Milla

SIMULATING DIFFUSION OF CHOLERA IN GHANA

JULIANA USEYA
Enschede, The Netherlands, February, 2011

Thesis submitted to the Faculty of Geo-information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Geoinformatics

SUPERVISORS:

Mrs. Ir. Ellen-Wien P. W. M. Augustijn-Beckers
Dr. Raul R. Zurita-Milla

THESIS ASSESSMENT BOARD:

Dr. Ir. R. A. de By(chair)
Dr. Ir. P. R. van Oel (External Examiner)

ABSTRACT

Despite all the widely available cholera awareness information, people are still suffering and dying of cholera. People do not know how much contribution is made by the different mechanisms of cholera transmission during its propagation. The main objective of this research is to develop an Agent-Based Modelling (ABM) prototype that simulates the diffusion of cholera and evaluates the contributions from the different mechanisms during a cholera outbreak in Kumasi, Ghana. Four cholera transmission mechanisms, a hydrological model, a cholera disease model, daily activity model, housefly model and a synthetic population were integrated in an ABM simulation prototype to simulate the propagation of cholera and was implemented in NetLogo. People with blood type 'O' have a higher risk of contracting cholera than others, and low hygiene standards, unsafe water supply and low income levels are factors which perpetuate the spread of cholera. Human-to-environment-human transmission mechanism plays a major role in the diffusion of cholera and human-to-human transmission has the least contribution.

Keywords

Cholera, Transmission mechanism, Agent-Based Modelling

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I am genuinely and deeply indebted to a number of educationists who have contributed to this dissertation in various ways. I am especially grateful to my supervisors who took their valuable time to support, guide, assist and encourage me till the completion of this thesis, Mrs. Ir. Ellen-Wien P. W. M. Augustijn-Beckers and Dr. Raul R. Zurita-Milla. I appreciate all your constructive advice, comments, reviews and suggestions. This thesis would not have been possible and completed without you and I am heartily thankful to Ir. Bas V. Retsios and Dr. Ir. P. R. (Pieter) van Oel who offered support in the development of my prototype. To Dr. Frank Osei, you are my source of inspiration.

Special thanks to Catherine Linard, Eric Lambin, Nicolas Poncon and Floriana Gargiulo for responding to my e-mails even if I have never met them.

I would like also to express my appreciation for the moral support from family members (dad, mum, Fungayi, Julian and Fadzie), friends, classmates and all of those who supported me in any respect.

Last but not least, I pray and thank Almighty God for everything.

Chapter 1

Introduction

1.1 MOTIVATION AND PROBLEM STATEMENT

Global media sources are always disseminating information about emerging public health threats and disease outbreaks worldwide. All the nations and health organisations need to "keep abreast of the potential public health threats worldwide and undertake prompt prevention and control activities" [54]. Diseases are complex phenomenon hence epidemiologists implement some tools to assist understanding of their occurrence and transmission mechanisms [28]. Cholera is an epidemic disease which is of major global and public health significance and is listed as one of the three internationally quarantinable diseases by the World Health Organisation (WHO).

Cholera is an infectious disease caused by Vibrio cholerae (V. cholerae), and is transmitted through ingestion of contaminated water and/ or food [58, 63]. It is both a waterborne and food borne disease. Cholera is one of the most researched communicable diseases and has gained both global and public health attention. It is a healthy threat, especially in developing countries [24], where there is poor access to safe drinking water and sanitation infrastructure have deteriorated [84]. Cholera outbreaks within most developing cannot be easily avoided but its propagation can be and has to be contained because if untreated it can cause high mortality [43] and the deaths can occur within a few hours of contracting the cholera infection [84]. The effective control measures of cholera rely on its prevention, preparedness and response [84].

In Ghana, cholera has been a public burden and persistent since the beginning of the 1970s [58, 60]. It is a complex disease since V. cholerae can survive both within and outside an aquatic environment [58]. It has environmental, socio-economic and demographic factors which contribute and influence its propagation within a population. Drinking contaminated water has been identified as the (main) leading mode of cholera infection, followed by eating contaminated food. Poor sanitation practices and poor hygienic standards serve as the main catalysts for the increase in cholera transmission and fatality, they accelerate the cholera propagation.

Cholera is more prevalent in Kumasi Metropolitan, Ghana. This is due to high urban growth making it difficult for sanitation coverage to keep pace with the urban population growth due to urbanisation. The housing conditions have become overcrowded hence causing high population density. Factors such as overcrowding, proximity and density of dump refuse sites, low socio-economic status, poverty, proximity to water sources and human susceptibility among others influence cholera propagation.

Mechanisms of cholera transmission are human-environment-human (HEH) transmittance [48], environment-to-human (EH), human to human (HH) transmittance and vector transmission (VT) by houseflies (Musca domestica)[35, 65].

Despite all the widely available cholera awareness information, people are still suffering and dying of cholera. After several decades of health and biological research in Ghana, there are still many unanswered questions about the epidemiology and diffusion of cholera in Ghana. Even though cholera transmission mechanisms and the different factors contributing to its propagation are known, the degree (percentage) of contribution from these factors to the overall spread is unknown [18, 43]. Thus, the most contributing factor is not known. The risk factors influencing cholera propagation are poorly understood. The contribution from the influencing factors can be studied using epidemiological tools including ABM. ABM contains tools which are useful in helping to have better understanding of disease occurrence, its contributing factors [28] and the dynamics of the disease spread to a certain extent.

Currently, there are three main approaches to epidemiological modelling which are network theory, mathematical/statistical and agent based modelling (ABM) [23]. For this research, ABM approach is going to be implemented. It is a bottom-up approach [71], it allows modelling of disease (complex phenomenon) as dynamic systems of interacting agents and their environment [24]. ABM has the ability of representing the factors influencing disease transmission in either spatially explicit or non-spatially explicit environment [49]. Agent based models are simple but useful in simulating local disease outbreaks [62]. Apparently, it seems that there is/ are no existing agent based models for cholera that have be published, hence it is my main motivation to develop an ABM prototype for the simulation of cholera diffusion in Kumasi Metropolitan, Ghana.

1.2 RESEARCH IDENTIFICATION

Key concerns to be investigated in this research are identified in this section and these are the main objectives, sub-objectives and research questions. The innovation of this research is also included.

1.2.1 General objective

Is to design and implement a prototype using Agent-Based Modelling to simulate the spread of cholera and evaluate the most contributing mechanism of transmission during the diffusion of cholera in Kumasi, Ghana.

1.2.2 Specific objective

1. Identify existing agent based models with related transmission mechanisms as of cholera.
2. Develop conceptual models for each transmission mechanism, agent behaviours and cholera dynamics.
3. Design implementation of the prototype.
4. Evaluate the performance of the prototype.

1.2.3 Research questions

To achieve the above mentioned objectives, the following research questions have been formulated and have to be answered explicitly.

1. How is cholera spread in Kumasi, Ghana?
2. Which existing ABMs of other diseases have the same transmission mechanisms as cholera?
3. How to effectively adjust the existing ABMs of diseases with the same transmission mechanisms as cholera to simulate cholera propagation?
4. How to work with limited data?
5. How to develop the conceptual models for each transmission mechanism, agent behaviours, and cholera dynamics?

6. How to design and implement the prototype?
7. How to test the performance of the prototype?
8. How to evaluate the different transmission mechanism combinations and their contributions to the overall spread?

1.2.4 Methodology

1. Literature study.
2. Acquiring skills and familiarising with NetLogo.
3. Data collection and Data Preparation
    - Data collection
    - Prepare the synthetic spatial environment data (shapefiles (houses, river network, and refuse dump sites), elevation data and houseflies distribution map.
    - Generate the synthetic population.
4. Create conceptual design models for:
    - Each transmission mechanism.
    - Agent UML Class diagram.
      Identification of agents, their attributes and behavioural interactions between themselves and with the environment.
        - Individual drinking water behaviour.
        - Individual eating food behaviour.
        - Cooking behaviour (household level).
        - Faeces contamination behaviour (household level).
    - Cholera dynamic model.
      Identification and definition of the (cholera) disease characteristics and its development.
5. Integration of the agent behaviours design models, transmission mechanisms design models and cholera dynamic design model.
   The development of the main design of the ABM prototype.
6. Design implementation in NetLogo.
   The implementation will follow the following sequential steps:
    - Load the environment (spatial) data
    - Load the agents
    - Enable population with behaviour
    - Simulate the transmission mechanisms
    - Simulate the flow of water
    - Simulate flies
7. Run different combinations and evaluation of the factors' contributions.
8. Verification and evaluation of the performance of the prototype.

1.2.5 Thesis structure

Chapter 1 - Introduction
This chapter contains the motivation and problem statement, research objectives, research questions and methodology.

Chapter 2 - Literature review
This chapter contains literature about cholera, the study area (Kumasi), agent-based modelling, conceptual designing for ABM and related work.

Chapter 3 - Conceptual design modelling
This chapter explains the conceptual design models for the static view (UML class diagram) and dynamic view processes (Flow charts and sequence diagram).

Chapter 4 - Data preparation
This chapter consists of the data collection, preparation of the spatial data and generation of synthetic population. It also contains well detailed steps which were followed in the data preparation.

Chapter 5 - Implementation of Cholera Simulation Prototype
This chapter explains the how the different models and processes were implemented in NetLogo.

Chapter 6 - Verification, Results and Discussion

Chapter 7 - Conclusion, limitations and recommendation

Chapter 2

Literature review

2.1 CHOLERA

Cholera is an acute/severe intestinal infectious disease caused by bacterium called vibrio cholerae (V. cholerae) O139 and O1 serotypes [26, 60, 24, 22]. It primarily affects the small intestines after ingestion of sufficient dose of the V. cholerae bacterium through water and/or food that is contaminated [60, 24, 63]. The main symptoms of cholera include sudden onset of profuse painless watery diarrhoea which increase frequency over the following hours and vomiting clear, watery and alkaline fluid [32]. The purging usually accelerates and its volume increases, the stools are rice-water in appearance and grey in colour, slightly cloudy with white flecks of mucus [32]. As the body fluids are lost, the patient becomes severely dehydrated, hence becomes weak [32]. Cholera causes intense thirst, loss of skin turgor, wrinkled skin (hands and feet), sunken eyes, pinched facial expression, muscle cramps (especially in the extremities), thready or absent peripheral pulses, falling blood pressure, and inaudible, hypoactive bowel sounds [73]. Without immediate oral rehydration therapy and proper medical treatment, it leads to fatality [43, 20] within hours of the onset of the purging. However, most patients who recover are free from the infection after about 2 weeks [73].

Humans are the specific hosts and victims of V. cholerae [85, 73, 48]. Once it enters the body, it has an incubation period ranging from several hours to five days [85, 73]. It cannot survive the acidic conditions in the stomach, hence "it reaches the small intestines and will propel themselves through the thick mucus that lines the small intestines to get to the intestinal wall where they can thrive. V. cholerae bacteria start up production of the hollow cylindrical protein flagellin to make flagella, the curly whip-like tails that they rotate to propel themselves through the mucus of the small intestine,"[86]. In the intestines, according to Osei and Duker [58] citing Carpenter CCJ, [15], V. cholera "produces an exotoxin that stimulates the mucosal cells to secrete large quantities of isotonic fluid or increases the permeability of the vascular endothelium thus allowing isotonic fluid to pass through abnormal amount resulting in the watery diarrhoea."

Cholera bacteria survive in alkaline conditions (above pH 4.5), temperatures below 70°C, moisture [84, 48] and moderate salinity conditions [50]. These are abiotic conditions which favour the growth of V. cholerae [50]. It can also survive freezing conditions [26]. Cholera is a complex disease and difficult to exterminate from the environment since V. cholerae can survive both within and outside aquatic environment [58]. V. cholerae is often found in the aquatic environment and it naturally occur as part of the normal flora of streams, riverine, estuarine, brackish water and coastal waters. Its presence is not associated with faecal contamination from cholera patients [25, 84, 55, 69]. V. cholerae attach to copepods, species of zooplankton that comprise the aquatic flora and fauna of rivers and bays serving as host for the bacterium [25]. Certain types of algae and plankton grow explosively when water temperatures rise hence the more the algae or plankton, the more the cholera bacteria [20, 25, 50]. Stagnant water and slow flowing water may lead an increased exposure to the organism [4]. According to Islam et al. [45], V. cholerae can survive in almost all in all kinds of aquatic environments including fresh water sources such as rivers, lakes, ponds and tanks since there can exist some spots with nutritious susbstances necessary for flourishing the bacteria.

Vibrio cholera can survive in both aquatic and non-aquatic environments. For the aquatic environment, it occurs in both natural surface water and fresh water sources as long as these reservoirs or sites provide enough nutritious substances for the bacteria to flourish in them, Islam et al. [45] quoted Robert Koch in 1884. Islam et al. [45] argued that V. cholerae can reproduce or multiply itself outside the human body. Non-aquatic environment include refuse dump sites, food (e.g. fresh vegetables, fruits, meats and cooked food) [33], human and animal faecal waste, inadequately treated sewage, plants near water source openings. An argument made by Wachsmuth et al. [81], is that V. cholerae can remain viable in aquatic environment for about two months, unless if there is recontamination. Hence my assumption is the bacteria can also live outside the aquatic environment.

Cholera bacteria survives well in faecal specimen if kept moist [58]. Stool passed out by cholera patients contain V. cholerae and people can get infected through faecal-oral route (ingesting water or food that is contaminated by the bacteria from stools excreted by an infected person) [20, 35]. In an epidemic, the source of the contamination is usually the feces of an infected person.

A question which remains unanswered is whether the V. cholerae O1 survive in all the aquatic environments as free-living organism or do they prefer an epibiotic host as a suitable ecological niche [45].

2.1.1 Cholera transmission

Sources of cholera infection

Cholera is both a waterborne and foodborne disease [68, 22]. Most common sources of cholera infection include surface or well water, raw or undercooked seafood, grains and raw fruits and vegetables [20].

1. Surface or well water
   V. cholera can lie dormant in water for long periods and contaminated public wells are frequent sources of large scale cholera outbreaks [20].

2. Raw or undercooked seafood
   Fish has been hypthesised as an essential vehicle for cholera transmission [69]. Eating raw or undercooked seafood especially shellfish emanating from water sources that are contaminated by V. cholera make people at risk of contracting cholera [20, 69, 27].

3. Grains
   In regions where cholera is widespread, grains of rice, millet and many others can be contaminated after cooking and allowed to remain at room temperature for several hours become a medium for the growth of cholera bacteria [20].

4. Raw fruits and vegetables
   Raw, unpeeled and unwashed fruits and vegetables are frequent source of cholera infection in areas where cholera is endemic [20]. In most developing countries, uncomposted manure fertilizers or irrigation water containing raw sewage can contaminate produce in the field. Fruits and vegetables may also become tainted with cholera bacteria during harvesting or processing [20].

Mechanisms of cholera transmission

There are two routes of cholera transmission namely primary transmission and secondary transmission [57, 11]. Primary transmission occurs when humans are exposed to contaminated water sources regardless of previously infected people or faecal contamination and is responsible for the beginning of initial outbreaks [57]. Primary transmitted cases should be located relatively close to water sources [57]. Secondary transmission occurs through faecal-oral route i.e. exposure to faecally contaminated water sources [57] hence aquatic environment is essential for the transmission and spreading of cholera [11].

However, secondary transmission route (faecal contamination) can occur through a number of ways which are hypothetically drinking faecally contaminated water, eating faecally contaminated food and direct contact with infected person. Therefore, my hypothesis is that cholera has four main mechanisms of transmission which are human-environment-human (HEH) transmittance, environment-human (EH) transmittance, human to human (HH) transmittance and vector transmission (VT) by houseflies.

1. **environment-human (EH) transmittance**

   EH transmittance (primary transmission) occurs from a natural reservoir of V. cholerae in an aquatic environment to human hosts [11]. In developing countries, access to safe drinking water is not a gurantee hence most people use public water sources. Without proper hygiene, people drink the water without treating it first or boil it to destroy the bacteria.

   Cholera mainly spread to developing countries through seafood mainly due to consumption of raw or under cooked smaller fish [69, 27] that feed on plankton where V. cholerae attach. It was discovered that the micro-organism can survive for weeks in the stomach of fish, even when refrigerated [69]. Low-lying riverine and estuarine environment provide the mild saline and alkali conditions favouring a permanent reservoir of V. cholerae [22].

2. **human-environment-human (HEH) transmittance**
   HEH transmission occurs mainly due to the faecal-oral route [22] and is called the secondary transmission. It was identified as cholera's leading spreading mechanism [48].

   - During rainy season, surface run-off containing V. cholerae will contaminated surface water sources and unprotected wells [58]. Drinking water from the contaminated surface water sources or wells perpetuates the transmission of cholera [4].
   - Uncomposted manure fertilizers or irrigation water containing raw sewage can contaminate produce in the field, fruits and vegetables maybe tainted during harvesting or processing [20].

3. **human to human (HH) transmittance**
   Cholera is rarely spread from one person to another (human contact) hence HH is very unlikely [27]. After being in contact with stool or vomitus from cholera infected person and then use those hands to eat without washing them with soap and clean water can cause cholera infection. Hands easily get contaminated by faeces in the environment [22]. In large gatherings of people such as funerals, festivals and pilgrimages, particularly if access to clean water and sanitation is limited, then HH transmission may occur [26].

4. **vector transmission (VT)**
   VT is still environment to human transmission of cholera being spread by houseflies called Musca domestica [35, 84]. The Musca domestica is "linked to the human habitat, and a chief offender among the filth breeding and filth feeding flies worldwide," [58, 35]. These filth flies breed and feed on matter in refuse dump sites [58] and will later fly and can contaminate food or drinking water that will be consumed by people.

It is important to note that during an epidemic, at most all the transmission mechanisms play a part in the spreading of the disease. Primary transmission mechanism (EH) does not fully explain the exponential growth of cholera incidences during outbreaks [57].

2.1.2 Factors influencing the spread

Cholera rapidly spreads in areas with poor sanitary practices, inadequate supply of safe drinking water and poor food hygienic practices [69, 84, 48, 68]. Once it is introduced within a population in a specific region, there are numerous complex factors which decisively influence its propagation and may lead to prolonged transmission [58, 48]. The factors include socio-economic, demographic, climatic and environmental factors. They play a role in facilitating the cholera transmission [22]. These factors include the following:

1. Poverty and low income level
   Usually, poor diet, malnutrition, poor housing facilities and lack of access to education are a result of poverty and low income levels [22]. Therefore people with any of these attributes are more likely to become infected by cholera.

2. Sanitation
   Cholera is hypothesised as a disease of deficient sanitation [48, 58]. The insufficient facilities for excreta disposal and drainage increase the risk of transmission of V. cholerae through the defecation of patients in areas surrounding their houses and its spread because of rain water floods [68]. The type of latrine used and inadequate treatment of sewage play roles in the cholera transmission.

3. Cooking practices [48]
   Light cooking of food which is contaminated may lead to the spread of cholera.

4. Poor personal hygienic standards
   Poor hygiene accelerate the cholera propagation within a given area [48, 22]. Unhygienic food preparation play a major role in the cholera transmission [69].

5. Unsafe water supply
   Most of the cholera outbreaks are caused by drinking contaminated water [48].

6. Human migration
   Migration of people is a risk factor since it plays a role in introducing cholera into new populations [4].

7. Overcrowding/ (high population density)
   Overcrowding will strain the existing sanitation resources (systems) hence putting the people at high risk of contaminating cholera [4, 58, 59, 48]. Usually in slum areas, there are poor sanitation practices and overcrowding.

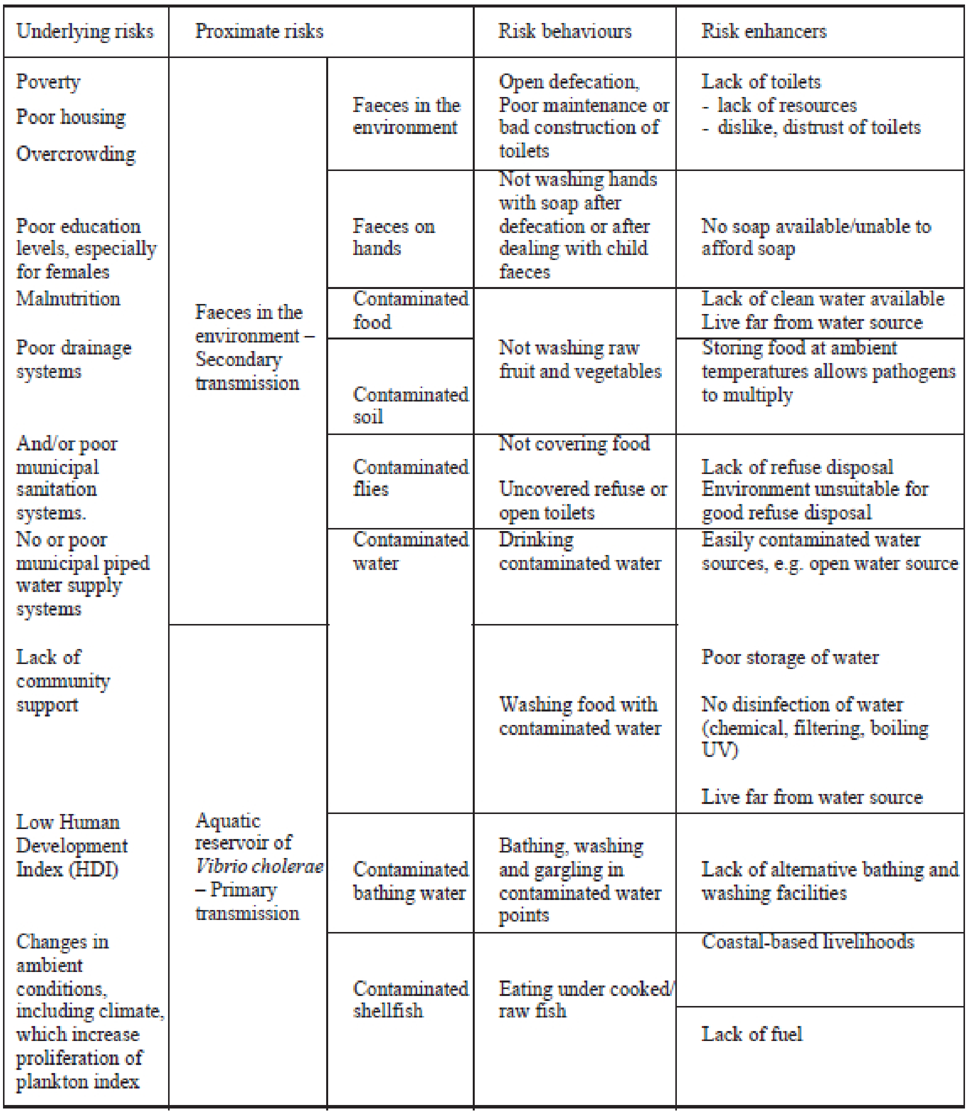
8. Low educational status
   Educational level is hypothesized to be an indicator of mortality because it indirectly determines health care seeking behavior [4, 5].

9. Heavy rains
   Heavy precipitation influence cholera transmission by feeding surface runoff into streams and rivers flooding to water supply, hence contaminating the drinking water sources [25, 68]. The rains usually trigger the epidemics although outbreak periods are irregular [22].

10. Proximity and density of dump refuse
    Cholera prevalence is directly proportional to the density of refuse dumps and inversely proportional to proximity of refuse dumps within a community [58].

11. Proximity to surface water sources
    Close proximity to drinking water bodies (which are contaminated) make inhabitants more prevalent to cholera [60, 4].

    Fig 1.1 is a table of a summary of cholera risk factors taken from Collins et al. [22]



Underlying risks	Proximate risks		Risk behaviours	Risk enhancers
Poverty Poor housing Overcrowding	Faeces in the environment – Secondary transmission	Faeces in the environment	Open defecation, Poor maintenance or bad construction of toilets	Lack of toilets - lack of resources - dislike, distrust of toilets
Poor education levels, especially for females		Faeces on hands	Not washing hands with soap after defecation or after dealing with child faeces	No soap available/unable to afford soap
Malnutrition		Contaminated food	Not washing raw fruit and vegetables	Lack of clean water available Live far from water source
Poor drainage systems		Contaminated soil		Storing food at ambient temperatures allows pathogens to multiply
And/or poor municipal sanitation systems.		Contaminated flies	Not covering food Uncovered refuse or open toilets	Lack of refuse disposal Environment unsuitable for good refuse disposal
No or poor municipal piped water supply systems		Contaminated water	Drinking contaminated water	Easily contaminated water sources, e.g. open water source
Lack of community support	Aquatic reservoir of *Vibrio cholerae* – Primary transmission		Washing food with contaminated water	Poor storage of water No disinfection of water (chemical, filtering, boiling UV) Live far from water source
Low Human Development Index (HDI)		Contaminated bathing water	Bathing, washing and gargling in contaminated water points	Lack of alternative bathing and washing facilities
Changes in ambient conditions, including climate, which increase proliferation of plankton index		Contaminated shellfish	Eating under cooked/ raw fish	Coastal-based livelihoods
				Lack of fuel

Figure 2.1: Cholera risk factors

Size and age structure of the population influence the cholera case incidences in high risk areas [5]. These risk areas can be defined based on household level case incidence and they are influenced by the status of each household [5].

2.1.3 Cholera mortality

Death due to cholera can occur from a few hours after onset of the purging until 5 days. Accessibility to treatment centers has been found to be an important determinant of diarrhea-related mortality [22]. In an unprepared community, case-fatality rate or death can be as high as 50% of severe cases.

2.1.4 Susceptibility

People staying in slum areas where there is overcrowding, poor sanitation practices, inadequate safe water supply and have poor hygiene are more susceptible to cholera.

Everyone is biologically susceptible to cholera except for infants who derive immunity from nursing mothers who had cholera previously [20]. Some people are more vulnerable to cholera than others [26]. Individuals' susceptibility to cholera (and other diarrheal infections) is affected by their blood type, those with O blood type are the most susceptible, thus twice as likely to develop cholera as are people with other blood types [76, 20, 26], while those with blood types A and B are the most resistant. Between these two extremes are the A and B blood types, with type A being more resistant than type B and the most resistant people have AB blood type. People who are gastric achlorhydia (hypoacidic) are at risk of cholera [26].

2.1.5 Cholera in Ghana

Cholera has been a public health burden in Ghana since the early 1970s when it was introduced [60, 58]. Between 1999 and 2005, Ghana Ministry of Health reported to WHO approximately 26,000 cholera case and 620 deaths [58, 60]. Ghana is in West Africa, and West Africa is recognised a virgin territory or a receptive area for cholera. In general, the origins of cholera incidence in Africa are socio-economic and environmental [22]. Cholera outbreaks in Ghana are caused by V. cholerae O1- El Tor biotype. However, cholera outbreaks in Ghana can be explained by the role of flooding on disease transmission [24]. Most epidemics are initiated by the EH transmission, then HEH transmission will play a major role. At times, cholera is introduced into a place by large migrations of people.

2.2 KUMASI METROPOLITAN, GHANA

2.2.1 Study area

The study area is Kumasi metropolitan, and is the second largest urban center (city) of Ghana situated in the southern central of Ashanti Region [88]. Ashanti Region which is located in the middle belt of Ghana. Kumasi metropolis lies at the intersection of 6.04°N latitude and 1.28°W longitude, covering an area of about 220 square kilometers [58].

Fig 2.2 is a map showing the location of Kumasi area, Ghana. Source is Osei et al. [60]

Kumasi has a population of approximately 1.2 million "which accounts for just under a third (i.e. 32.4%)"of Ashanti region's total population [60, 58]. For the population age and sex structure, the highest proportions of the population are in the age cohorts 0 - 4 years (13.2%) and 5 - 9 years (12.4%). Cumulatively, 39.9 per cent of the population is below 15 years. There are more males (50.2%) than females (48.8%) in the metropolis [6]. Kumasi

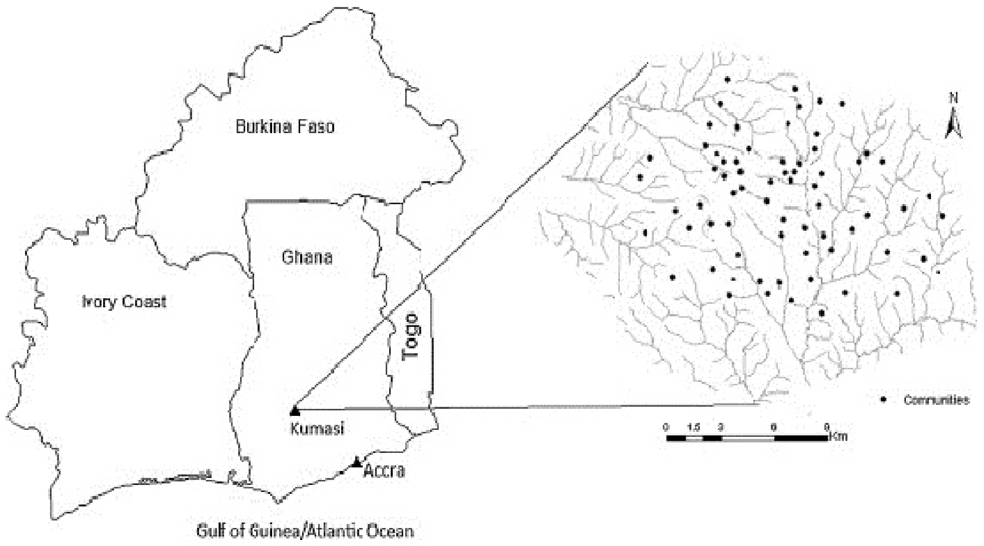


Figure 2.2: Kumasi, Ghana

is the Ashanti's regional capital and the fastest growing metropolis since it has attracted a large population. This is mainly because it is the most commercialized city in the region [60]. Another reason is that it is nodal city linking major road arteries to the northern and southern parts of Ghana [60, 47]. There are two main seasons, the rainy season and the dry season. The rainfall pattern is bimodal with long rainy season is from April to July with peaks in May/June and a short season between September and mid-November with peaks in September/October [60, 47]. According to The Chronicle [19] and Victoria Beddow [10], significant flooding occur quite often in the Kumasi in areas where the terrain is flat and areas near the rivers and streams.

2.2.2 Characteristics of Kumasi area

Since Kumasi is a center of diverse culture and commerce, it indeed reflects the socio-economic scenarios of the country. The metropolitan area comprises of settlements in the urban core and urban peripheries [14]. In the urban core, there is greater concentration and number of housing units whereas in the periphery settlements, there is less concentration of housing units [14]. The settlements are also referred to as communities or suburbs. A larger number of the population stays in low-income settlements [10].

Living/ housing conditions in many of these communities are overcrowded [83]. This is mainly because of rapid urban (population) growth and is due to urbanization. Poor housing conditions reflect high probability of low incomes [83]. The housing costs are high [60], hence approximately 75% of Kumasi residents stay in multi-family buildings (single or multi-storey compounds) [83, 72]. The houses in Kumasi metropolitan can be categorized into five main classes which are single storey traditional compound houses, multi-storey compound houses, government built detached or semi-detached for low-income households, large single household houses built on relatively large plots and block of flats [6]. About half of Kumasi's residents live in the traditional compound houses [72]. The on-going urbanisation has also resulted in growth/ development of slums (informal settlements) [38].

According to Graham Tipple (1984) four main housing sectors were identified and these are:

(a) High Cost Sector (Low density high class area) with 1 to 3 houses per hectare and 50 persons per hectare.
(b) Tenement Sector (High density medium class area) with 5 to 7 houses per hectare and 200 persons per hectare.
(c) Indigenous Sector (High density low class area)
(d) Government Built Sector (Medium density medium class area) with 3 to 5 houses per hectare and 100 persons per hectare.

According to the Kumasi Metropolitan Website [6], housing tenure in Kumasi can be categorized into three, namely house owners, renters and occupants. Occupants, here represents a group of people who neither rent nor own the house, but live there. About 74% of households rely on rental housing and about 70% of all households in the city occupy single rooms. Average room occupancy in the city is very high (2.7) depicting shortage of housing. The current housing situation in the Kumasi city is on average, a household size is 5.1, average number of persons per house is 25.7 and average number of rooms per house is 9.8 [6].

Water supply situation
In 1993, according to Whittington et al. [83], about 58% of the households in Kumasi had access to a private connection to the municipal water supply system. In the low income communities about 30% of the population have in-plot or communal standpipe access to the piped water supply [10]. According to Osei et al. [60], approximately 82% of the inhabitants now have access to pipe-born water. Unfortunately, the water supply is inadequate [79], there are rampant water shortages due to the low power voltage [29]. During these water shortages, most households use surface water from rivers and streams for drinking, cooking, bathing and washing of utensils [60]. Only a few people who are rich can afford to buy bottled water for drinking. Some households have dug shallow and unprotected wells where they fetch their water for day-to-day use [83]. During rainy season, they seldom collect the rainwater [83].

Sanitation system
There are three main types of excreta disposal systems that are used in Kumasi. These are public toilets, water closets and bucket latrines [83]. 86% of the households use public toilets which mainly are pit latrines [10] and a few bucket latrines [79]. Most of the public toilet users complain about the toilets' cleanliness and also their distance from the houses [10]. Usage of the public toilets has to be paid per each visit. Due to high demand for accommodation and high housing costs, most landlords (house owners) have converted the sanitation rooms into sleeping rooms for the renters hence leaving most of the inhabitants without access to a household sanitation system [60]. A few households have access to water closets in their buildings [83] with septic tanks. There are no sewers in most parts of the city despite them having been planned since the early 1970s [79]. However, due to the absence of both household and public sanitation systems in most highly populated areas and slum settlements, many people end up using the bush system whereas others just defecate on open space and refuse dump sites [60, 83].

Waste Disposal Facilities
The collection system of the waste management in the city is based on two systems which are house-to-house waste collection and communal solid waste collection [89]. At least,

5 out of 10 households dispose of their waste right beside their houses, instead of finding the nearest waste dumping site. Generally, safe house-to-house waste collection is still very low [60]. In some cases, the waste containers provided by the various waste management companies are not enough for the communities, due to the large populations. People who live beside bushes or near rubbish dump sites, defecate on these places.

2.2.3 Characteristics of communities

The largest ethnic group is the Ashanti, but other ethnic groups are growing in size. Approximately 80% of the population is Christian and 5% Muslim, with a smaller number of adherents to traditional beliefs [88].

Communities in Kumasi are not homogeneous since they consist of people belonging to 1st class, 2nd class, 3rd class and some slums called “zongos”. Majority of the people staying in the zongos are Muslims and they are poor and uneducated. They have migrated into Kumasi coming from other countries.

2.2.4 Behavior of people in Kumasi

Behaviour of people is controlled partly by socio-economic context, which determines the decisions such as to invest in safe house and water facilities, buy soap, and prioritize purchase of low risk food [22]. Majority of the risky behavior of people is related to cultural influences and lack of appropriate education [22].

Farming and trading are the main income generating activities in Kumasi. Most of female headed households and households with many people headed by older and less educated people, use their houses for income generation [72]. Some women cook food to sell at market or street corners for income generation [72].

Since almost 75% of the people live in compound houses, the central courtyard is used by all the households for cooking, pounding maize meal, washing clothes and many other activities [38].

Fish is usually bought whole (or in cross-sectional slices), either at the market, from roaming fish-sellers walking through the city, or directly from fishermen. The fish are not cleaned before being sold, but taken home to be gutted by the household members themselves and majority of them do not wash their hands with soap [69].

2.2.5 Limitations of data acquisition on Kumasi

Demographic data available for Kumasi area is at community level. It is very difficult to get individual data (micro data) about household structures since the institutes managing statistics usually provide aggregated datasets, describing the global properties of the households and individuals. This is mainly to maintain confidentiality. Data collection is often very expensive. However, artificial datasets of individuals and households “synthetic population”can be created from the aggregated data to represent a realistic population. Synthetic population consists of individuals described by their age and are gathered in households respecting a variety of statistical constraints (distribution of household types, sizes, age differences among parents and children) [36].

Synthetic population

Synthetic data needs to be as accurate, precise and realistic as possible, hence closely representing the real population data. There are different existing approaches of developing synthetic populations which include stratified sampling, geodemographic profiling, data fusion, data merging, iterative proportional fitting and reweighting [2].

Synthetic environments

Communities in Kumasi are spatially represented as points, and their boundaries have not been dermacated, hence not available. However, "synthetic environments" can be used. In general, there are many reasons why it is often necessary to use these "synthetic environments" which include [9]:

(a) practical reasons (e.g. costs of experiments or data acquisition, observation of systems characterized by a very slow evolution)
(b) the simulated system cannot actually be observed (since it is actually being designed)
(c) for ethical (e.g. the safety of humans would be involved)

2.3 AGENT-BASED MODELLING

Agent-based modelling (ABM) is a relatively new, powerful computational simulation technique for modelling phenomenon as dynamic systems of interacting agents living in an environment [16, 12]. In ABM, a system (or the phenomenon) is modelled as a collection of autonomous decision making entities called agents [44] which live in an environment. These individual agents assess their situations and make decisions based on their set of rules [12], hence their activities emerge to give the general behaviour of the system [78]. These sets of rules define the behaviour of agents and their interactions amongst themselves and with their environments.

ABM has gained popularity in the past few years and is being applied in many different fields [44] including epidemiology. The simulations models have great potential as tools which lead to a better understanding of the world's processes [7], hence epidemiologists have now started to implement agent-based models in their studies of epidemic disease outbreaks. The different interactions between agents and their environment determine the rate of spread/ diffusion of epidemic disease. However, the agent-based simulation models are tools which can be employed to study the disease's spread (due to the interactions), the spread of infective agent and risk factors [62] influencing its spread. They also take into consideration, the agents' susceptibility due to their differences in age.

As the simulation would be running, the outputs/ results are displayed on most probably a user-friendly interface. Therefore, one can visual observe the movements of agents and the infection state with time together with graphical displays of the population dynamics [31], thus one can explore the consequences of the specified individual-level rules on the level of population as a whole, using the simulation runs' results [74].

2.3.1 Elements of an Agent-based simulation model

An agent-based model consists of the following elements [70, 34, 9]:

(a) A set of agents (part of the user-defined model)

(b) Spatially explicit or non-spatially explicit simulation environment

(c) A set of agent and/or environment relationships (part of the user-defined model)

(d) A framework for simulating agent behaviours and interactions provided by an ABMs toolkit or other implementation).

2.3.2 Environment

The environment of an agent is a discrete space or part of the virtual world (or computational system) where agents inhabit, operate and serving to support their interaction with the environment and other agents [17, 51]. A simulation environment can either be spatially-explicit or non-spatially explict. It can consits of several clearly distinct sets of cells or layers that define different landscapes and can either be environment is static or dynamic. These multiple environments may interact amongst themselves and their interactions may influence the behaviour of the agents living on them. The environment can be a Euclidian space (non-spatially explicit), regular tesellations, network, TIN or vector polygons (GIS).

Agent

An agent is a computer system situated in some environment and is capable of autonomous action in this environment in order to meet its design objectives [46].
There is no universal agreement on the precise definition of an agent, but most definitions tend to agree on more points than disagree [1, 17]

Agent-based simulation models consist of discrete, purposeful, autonomous agents who interact in discrete space (environment) and time whose micro level interactions create emergent patterns [61].

An ideal and intelligent agent is a computational object with elementary properties which describe their behaviour and interactions [61, 91, 74]. There are emergent properties which arise from the interactions of the agents that cannot be deduced by mere aggregation of the elementary properties of the single agents [8, 74, 34].

Elementary properties

(a) Autonomous [46, 91, 34, 17]
Agents are autonomous (self-organisation). Agents are able to control and perform actions, thus some patterns emerge with coordination not from centralised authorities or institutions or humans, instead, by the local interactions amongst the autonomous decision makers. An agent is able to determine how to best solve a given problem and act on its own initiatives. Agents are self-organised.

(b) Behaviour rules
The behaviour of an agent determines what it is supposed to do during its life in the simulation. Some actions by agents are goal-driven but some are not. Therefore, goal-driven behaviours are achieved by agents by following simplest set of programmable rules which are aimed at achieving a certain goal, hence they enable generation of macro pattern of explanatory interest [34]. The behaviour rules explicitly express the interaction patterns emanating from the interactions of the agents either with each other or with their environment. Each rule consists of two parts which are a condition

part (specifies when the rules is to fire) and the action part (what is to happen) [91]. According to Tang W [77], behaviour rules are an explicit representation of decision-making processes of the agents. Some of the rules of agents may compose of their memory.

- Memory and Adaptive
  Agents are adaptive and backward-looking, they have a memory. They adapt by moving, imitating, replicating, or learning user's preferences, but not by calculating the most efficient action. Agents can adapt at the individual and the population level although not all agents are capable of adapting their behaviour [34, 46, 91, 1, 37]. In a simulation, the environment undergoes changes, it is the same with the status and behaviours of agents. However, in order to make decisions of any actions, agents need to have enough information about the environment and themselves. Current information/data and proposed potential action plans are stored in the agent's memory [91].

For instance, in a family, the mothers fetch water for household use. They know where they get the water from always (be it a tape or well or river) and in times of water shortages, they remember where the once fetched water before, hence the decision to collect water from such a place. The simple behaviour rule will be "move to fetch water".

(c) Interdependency
Agents are interdependent (each agent's decisions depend in part on the choices of others). Agents influence others in response to the influence that they receive. Interdependence may also be indirect, as when agents' behaviours change some aspect of the environment, which in turn affects the behaviour of other agents [34, 1].

(d) Attributes
Each agent is identifiable and has attributes such as unique identifier, location (x,y coordinates), age, gender, current health status and may more [1]. Agents vary by their attributes and/or levels, however, different levels of agents may exist in a simulation model and will have their own attributes, whereas the lower level agents can inherit attibutes from the higher level agents. An example is a household will have its own attributes but will inherit the attributes of the community in which they are situated. (In object-oriented programming (OOP) approach, a subclass inherits all attributes/-variables of its super class).

(e) Goal-oriented
Every agent has goals which it strives to fulfill or achieve [1]. An agent will act in a way that will enable it to a particular goal, which can be either a relative or extremal value [37].

(f) Communicative
Agents have the ability to communicate extensively with other agents using an expressive communication language. Ideally, this language should resemble human-like speech acts [91, 17]. For example, agents can query other agents and/or the environment within a neighbourhood, searching for specific attibutes with the ability to disregard an input which does not match a desirable threshold [17].

(g) Heterogenous
The notion of mean-individuals is redundant, the agents permit the development of autonomous individuals. Groups of agents can exist but they are spawned from bottom-up, amalgamations of similar autonomous individuals [17]. Heterogeneity originates

due to the differences in location, knowledge, preferences, experiences, wealth and many other aspects.

(h) Flexible
Intelligent agents should have *flexible* autonomous actions in order to meet its design objectives. Flexibility means -

- Responsive (or reactive)
  An agent is able to conceive its environment, sense their surroundings and be capable of reacting in timely fashion to changes that occur in the environment based on perceptions [46, 91, 17].
- Proactive
  Agents should not simply act in response to their environment but be able to perform opportunistic, goal-directed behaviour and take the initiative where appropriate [46, 91].
- Social
  Agents should be able to interact and cooperate with other agents or humans in order to carry out complex tasks they cannot handle as an individual and should help others with their activities [46, 91].

The following properties are regarded as less important when referring to intelligent agents [46, 91].

(i) Mobility
Some agents are capable of moving from one system to another to access remote resources or meet other agents. If agents can independently roam around the model space, then the mobility vastly improves the interactivity of agents with environment [37, 17].

(j) Personality
Refers to the capability of showing a believable human-like character such as emotions.

(k) Individual world view
Each agent has its own model of the external world.

(l) Reproductive
Some agent are capable of reproducing, hence they can evolve and adapt to changing situations, increasing their survival skills.

Agent relationships (interactions)

Agent-based models are comprised of multiple, interacting agents within multiple interacting simulation environment. A relationship between agents is specified, linking agents to other agents and/ or to other entities in the same system [17]. Interaction is a key aspect in ABM [9]. Agents have protocols to follow such that they are able to interact with their environment and/ with other agents (entities) in order to solve problems or simply reach their goals according to coordination, cooperation or competition schemes [9, 1]. Interactions include agent-agent, agent-environment and environment-environment interactions. The mechanisms of interaction among agents can either involve a direct exchange of information among the involved entities or an indirect one, realized through the possible perception by an agent of the effects of another agent's action.

ABM simulation platforms/toolkits

Castle and Crooks [17], define ABM toolkit or software as simulation/ modelling software that provide conceptual framework for organizing and designing agent-based models. They provide libraries of software functionality that include pre-defined routine/ functionalities specifically designed for ABMs.

Bandini et al. [9], define ABM simulating toolkit as a software framework or environment that provides abstractions and mechanisms for the definition of agents and their environments to support their interaction, but also additional functionalities like the management of the simulation (e.g. set-up, configuration, turn management), its visualization, monitoring and the acquisition of data about simulated dynamics. There are different toolkits available for agent simulations. Most of the commonly used ABM platforms follow the framework and library paradigm, providing a framework, set of standard concepts for designing and describing ABMs along with a library of software implementing the framework and providing simulation tools [64]. Following a library based paradigm means, a host programming language can be extended through generic procedures to facilitate ABM development. For a standard approach, the library will include a framework (a set of concepts for designing ABMs) along with tools for simulation and visualisation [52].

The development of agent-based models can be greatly facilitated by the utilisation of simulation/ modelling toolkits. They provide reliable templates for the design implementation and visualisation of agent-based models, allowing the modellers to focus on research (i.e. building models), rather than building fundamental tools necessary to run a computer simulation [17].

ABM has gained popularity since it is considered to be easy to implement, hence anyone who hears about it can easily program an agent-based model [9]. However, it is wrong to think that concepts behind ABM are ease to master, instead ABM is conceptually deep [9]. Having the wrong perception and conceptual mind about the ABM, then often it leads to improper use of ABM [9]. Therefore, it is crucial for the researcher to get some substantial amount of time to understand how to design and implement agent-based models [17].

Some platforms allow a researcher with limited expertise in software development to develop a model for his/ her research. The use toolkits can reduce the burden modellers face on programming parts of a simulation that are not content-specific (e.g. Graphical User Interface, GUI, data export-import, visualisation/ display of the model [17].

Simulation toolkits [66, 37, 64, 17, 53]

There are a lot of ABM toolkits available these days, and they are distinguished from each other in terms of the scale of the software that one can apply according to the following continuum:

(a) Desktop computing for ABMs Application Development

- Spreadsheets: Excel using the macro programming language VBA
- Dedicated Agent-Based Prototype environments: NetLogo, Repast Simphony
  NetLogo (http://ccl.northwestern.edu/netlogo/)
  NetLogo is the highest-level ABM platform/toolkit which is easy to install and to operate. It uses its own programming language, Logo programming language and it is easy to learn hence requires limited expertise in programming. Users can explore models without any technical knowledge. It includes an extensible “models library ”which user can explore and see the different types of models that have

been made using the toolkit. Generally, they are comprehensively documented and the documentation is easy to access. NetLogo contains many sophisticated capabilities (behaviours, agent lists, built-in graphical interfaces and many more) and is likely the most widely used platform for ABMs. An animation display automatically linked to the program and optional graphic controls and charts. NetLogo does not distribute its source code. NetLogo was developed for more specific models implementing mobile agents acting concurrently on a grid space with behaviour dominated by local interactions over short times. It is designed primarily for ABMs f mobile agents with local interactions in a grid space.
  In NetLogo, agents are represented as "turtles"and the environment as "patches". It provides a graphical environment to create programs that control graphic "turtles"that reside in a world of "patches"that is monitored by an "observer". Models dveloped in NetLogo are difficult in extend in NetLogo but can be easily extended in other ABM plaforms such as Repast. NetLogo is designed specifically for the deployment of models over the internet. NetLogo is highly recommended, even for prototyping complex models.

(b) Large scale (Scalable) Agent Development Environments:
  - **Repast** (Recursive Porous Agent Simulation Toolkit)
    It is pure Java implementation hence requires relatively good Java programming skills. Repast is a free and open source software and has been widely or extensively used in social simulation applications. It also provides an extremely wide set of "library "for developing models. Repast Simphony is the latest version of Repast and it was designed to provide point-and click tools for model deign, model behaviour specifications, model specifications and results exercusion.
  - **Swarm** (http://www.swarm.org)
    Is one of the oldest ABM toolkit and was originally written in Objective C then later fitted with a Java interface. It is a open source simulation system designed specifically for the development of multi-agent simulating of complex adaptive systems. It reqires strong programming skills. Swarm was designed to study biological systems.
  - **MASON** (Multi-agent Simulator of Neighbourhoods/Networks)
    Is a multi-agent simulation library in Java, designed to serve as the base class structure for custom Java simulations. It also includes a model library and suite of 2D and 3D visualization tools and is developed with an emphasis on speed and portability. It also requires strong programming skills but unfortunately it does not provide functionality for dynamichartingmodel output during simulations. Little technical documentation is available.
  - **AnyLogic**
    Is a proprietary toolkit which incorporates a wide range of functionality for the development of agent-base models. AnyLogic models can only be created on Microsoft operating systems, although a simulation can be run on any Java-enabled operating system once compiled. Source codes and documentation of example models are no available.

(c) General programming languages
  Majority of simulation syatems require the user to develop their models in an object-oriented language [17]. The programming languages are low-level computer languages and include the following:
  - Python

- C++
- Java
- Visual Basic

2.4 CONCEPTUAL DESIGN MODELLING FOR ABM

The process of building an agent-based models begins with a conceptual model [17]. A conceptual model is layout or representation of concepts or entities and the relationships between them. For ABM, basic questions or goals, elements of the system and the measurable outcomes of interest are identified [17]. It is very crucial to identify the purpose of the simulation model whether it is going to be a explanatory or predictive model. Iba et al. [44], described the purpose of a conceptual modelling phase as a phase to specify what the target system is: what are the agents and their attributes, what the rules of agent interaction and behaviour are and what the model environment is. The activities of the agents and the sequence of communications are also described during this conceptual designing phase.

Conceptual modelling can be done by either drawing paper, described in words or can remain imaginary in the mind. However, the bottom line is conceptual models are used to assist in the understanding of the subject matter to be represented by the model. Due to the evolving technology, there now exist software which can be implemented in the conceptual designing phase which use some standard notations. Unified Modelling Language (UML) for object modelling is commonly used, and Information Engineering (IE) is used for entity relationship (ER) modelling [87].

“In UML notation, the conceptual model is often described with a class diagram in which classes represent concepts, associations represent relationships between concepts and role types of an association represent role types taken by instances of the modelled concepts in various situations. In ER notation, the conceptual model is described with an ER Diagram in which entities represent concepts, cardinality and optionality represent relationships between concepts, ”[87].

When using (agent) UML to describe a model, the products developed in this phase are [44]:

(a) Class Diagram for Static view
(b) Activity diagram for Dynamic view (similar to flow chart)
(c) Communication-Sequence Diagram for Dynamic view and iteration
(d) Object state diagram for Dynamic view

(a) **Conceptual-Model Class Diagram** for Static view
A class diagram represents the static structure of a system. Generally a class diagram encompasses

- classes
- class component parts
- the way in which classes are related to one another

A class is a descriptor for a set of objects that share the same attributes, operations, methods, relationships and behaviour. A class can be defined as a template for a set of objects where their similarities are emphasised and at the same time ignoring their differences. For ABM, the first step is to specify/ identify the different agents and the environment of the system, then the modeler can define and represent them in a class diagram. The characteristics or attributes of the objects are clearly shown and presented in the class diagram. The associations representing the static relationships between or among the classes are also clearly presented.

(b) **Activity diagram** for Dynamic view similar to flow chart
An activity diagram displays task distribution and control flow. Activity diagram is a sophisticated flow but it is very much similar to a flow chart. Activity diagram enables the modeler to describe the procedure of agent's behaviour that is the activity analysis [44]. It assists to visualise and exploit the opportunity of parralel and concurrent processes.

(c) **Communication-Sequence Diagram** for Dynamic view and iteration
A sequence diagram displays the time sequence of objects participating in an interaction. The sequence diagrams are used primarily to show the interaction between objects in a sequential order in which the interactions occur. In ABM, the modeler describes the sequence of the communication among agents. The agent's behaviour is often done in cooperation with other behaviours.

(d) **Object state diagram** for Dynamic view
An object state diagram displays an object lifecycle including the responses and actions.

ABM approaches

There exist various agent-based modelling approaches which include Pattern-oriented modelling and Overview Design concepts and Details (ODD) protocol.

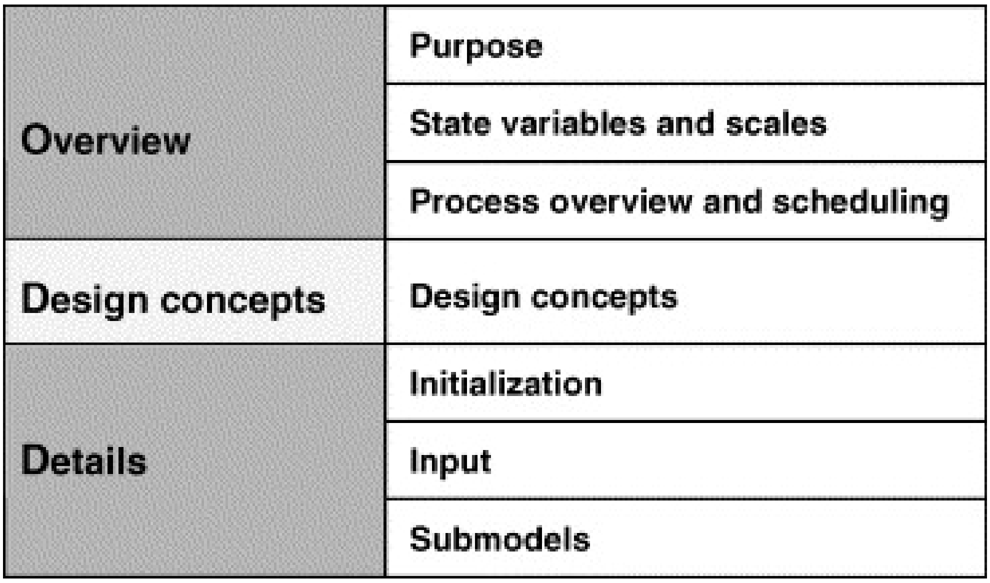
(a) **Pattern-oriented modelling (POM)** [40]

Pattern oriented modelling is a strategy that provides a unifying framework for decoding the internal organisation of agent-based complex systems and may lead toward unifying algorithmic theories of the relation between adoptive behaviour and system complexity. POM was mainly formulated in ecology.

Modelling starts with specific questions which lead to the formulation of conceptual model that helps to decide which elements and processes of the real system to include and ignore in the system. For complex systems, the questions addressed may not be enough to locate "Medawar Zone", since agent-based complex systems include many degrees of freedom. The key principle for POM is to first utilise the multiple patterns observed in real systems at different hierarchical levels and different spatial and temporal scales to guide the design of the model structure. First and foremost, modellers need to observe the patterns that seem to characterise the system and its dynamics, and the variables and processes that must be in the model so that these patterns could, in principle, emerge.

(b) **Overview Design concepts and Details (ODD) protocol** [39]

ODD is a standard protocol for describing agent-based models and consists of three blocks (Overview Design concepts and Details), which are subdivided into seven elements which are Purpose, State variables and scales, Process overview and scheduling, Design concepts, Initialisation, Input and Submodels. The basic idea of ODD protocol is to structure the information about ABM in sequence of the seven elements mentioned above.



Overview	**Purpose**
	State variables and scales
	Process overview and scheduling
Design concepts	**Design concepts**
Details	**Initialization**
	Input
	Submodels

Figure 2.3: Seven elements of the ODD protocol grouped into 3 blocks

Overview (Purpose, State variables and scales, Process overview and scheduling) provides an overview of the overall purpose and structure of the model. After reading the overview, it should be possible to write in an object oriented-programming language, the skeleton of a program that implements the ABM described. The skeleton includes the declaration of all objects (classes) describing the models entities (different types of individuals or environments) and the scheduling of the model's processes. For the Process overview and scheduling, flow charts are convenient to use for the visualisation.

Design concepts describe the general concepts underlying the design of the model. They provide a common framework for designing and communicating ABMs. A short checklist of questions regarding the design concepts is provided and has to be followed when designing an agent-based model.

The Details block present details that were omitted in the overview. Initialisation emphasises on the initial values of the state variables. The environmental conditions are inputs that impose dynamics of certain state variables. The submodels represent all the processes listed in the process overview and scale including the parameterization of the model.

The logic behind the ODD sequence is context and general information is provided first (Overview), followed by more strategic considerations (Design concepts) and finally more technical details (Details).

Examples of ABM development from conceptual designs

(a) **Model Driven Development (MDD)** [44]

MDD is a new paradigm of software development for Agent-based simulations. It was proposed that tools such as Model Driven Architecture (MDA) and Executable UML be used for MDD process to build the conceptual models and simulation models. MDA is a systematic design approach for the development of software systems. It provides a set of guidelines for the structuring of specifications, which are expressed as models. MDA is based on modelling separately technology-independent and technology-specific aspects of a system. An essential feature of the MDA framework is to make transformations from design models to program codes, automatic program code generation.

MDD allows modellers to describe their models conceptually in high-level language, hence concentrating only on modelling without considering the software implementation since program code will be automatically generated and they will be exact translation of their designs. The codes generated use low-level programming languages.

According to Iba et al. [44], conceptual modelling phase is followed by the simulation design phase then verification phase. This is an iterative process until the best simulation model is achieved eventually. In the simulation design phase, the modeler designs and impliments the simulation model, the conceptual models are translated into simulation model. The verification phase, is when the modeler runs the simulation and inspects whether the simulation is coded correctly. If necessary, the modeler returns to the first or second phase and modifies the models.

(b) **Structure-Behaviour-Function (SBF)**
Vattam et al. [80] described a computational technique that takes Structure-Behaviour-Function (SBF) models of complex systems and simulates the behaviour of the modelled system in NetLogo. HYPERION was identified as a computational architecture which takes a SBF model of a system and transform into NetLogo programs through a SBF-NetLogo compiler. Nevertheless, the process of compiling is complicated by two factors [80]. “First, the mapping rules for the compilation process are context dependent and cannot be completely determined a priori. Second, the compilation process has to deal with an impedance mismatch arising from how the changes that a complex system undergoes are represented in the SBF and NetLogo frameworks” [80].

2.4.1 How to build a flexible and extensible prototype simulation model

Conceptual framework

The prototype simulation model consists of static or data model, dynamic behaviour model and dynamic interactive model (simulation) [7]. **The static model** includes mobile and non-mobile objects, for this research, vibrio cholerae, river network, dump sites, households and community are non-moving objects, whereby people are mobile objects. The static model will be conceptually modelled in a **class diagram**.

The behaviour model defines the agent’s behaviour. The behaviour production system consists of behaviour rules, actions, conditions, memory rule interpreter and arbiters [7]. Augustijn-Beckers and de By [7] explained in detail that behaviour rules explicitly store the interaction pattern by defining what-if scenarios. Each rule consists of four parts namely

(a) a class name
(b) a condition
(c) a characteristic
(d) a nameless function

This dynamic behaviour model will be conceptually modelled as **flow chart or activity diagram**.

The dynamic interactive model is based on the sequential order of the events and interaction between objects. Each event occurs at a time instant and marks a relevant change of state in the system [7]. Such dynamic interactive model consists of conceptual components such as a clock that keeps track of the simulation time and an event list [7]. The dynamic interactive model will be conceptually represented in a **sequence diagram**.

Guidelines in building an ABM simulation prototype

Steps in developing the simulation are:

(a) Select the boundaries of the system
(b) Define the objective of the system
(c) Identify the needs of the end-user
(d) Collect real system/ behaviour data (input)
(e) Formulate and develop the simulation model
(f) Verification (ensure that the model satisfies the specification)
(g) Validate the model (check against the real phenomenon)
(h) Test the model
(i) Document

2.5 USEFUL RELATED WORK

(a) *"The geosimulation of West Nile virus propagation: a multi-agent and climate sensitive tool for risk management in public health"* by Bouden et al. [13]

This reserach is about the development of a multi-agent geosimulation system that simulates the interactions of populations of mosquitoes and birds over space and time in relation to the spread and transmission of West Nile Virus (WNV).

They modelled the populations of mosquitoes distributed over the geographical environment using 'Intelligent density map" since the numbers of the mosquitoes are very large. The density map used cellular automata principle to simulate the the evolution of the different categories of the mosquitoes.

From this research, the idea of creating an intelligent density distribution map can be applied to develop a density distribution map of houseflies.

(b) *"Spatially explicit agent-based modelling for Schistosomiasis transmission: human-environment interaction simulation and control strategy assessment"* by Hu et al. [42].

Schistosomiasis (or Bilharzia) is a parasitic disease spread by Schistosoma japonicum through human contact with infected (contaminated) water. Hu et al. [42], simulated the human-environment interaction of schistosomiasis in China. In this research, they have modelled the HEH transmission which is the almost similar to the main transmission mechanism of cholera. For the human and environment interactions, they modelled the human water contact behaviour that is through working on the farm, swimming or bathing, or washing clothes or vegetables. Of much importance, they considered the faeces contamination behaviour. The faeces contaminate the environment through two ways which are manure and wild faeces. The agent-based model of the disease was implemented in NetLogo.

(c) *"Multi agent based simulation for Typhoid fever with complications: An epidemic analysis"* by Sutiono et al. [75].

Typhoid is an infectious disease spread by Salmonella typhi organism. Its main mode of transmission is by faecal-oral route or urine-oral route, either directly through hands soiled with faeces or urine of cases or carriers or indirectly by ingestion of contaminated water, milk, food or through flies [75]. Typhoid fever is also hypothesised to be an indication of poor environmental sanitation and hence occurs in areas where water supply is not safe and sanitation is substandard. The agent-based model of the disease was implemented in NetLogo. The main mechanism of transmission for typhoid implemented in the model is HH and the human-environment interactions are not considered.

(d) *"A multi-agent simulation to assess the risk of malaria re-emergence in southern France"* by Linard et al. [49]

Malaria is spread by the Anopheles mosquitoes. In this research, the population of mosquitos on one cell were modelled as a single agent. The mechanism of transmission of is VT. The agent-based model of the disease was implemented in NetLogo.

By studying, adopting, modifying and integrating all the ideas in the above mentioned agent-based models of other diseases and other agent-based models, an agent-based model for cholera can be built by implementing the different ideas from the different researches.

Chapter 3

Conceptual Design Modelling

3.1 OVERVIEW

3.1.1 Purpose

The main purpose of this ABM prototype is to simulate the diffusion of cholera in Kumasi, Ghana. The broader objective is to evaluate contributions that are made by each of the transmission mechanism to the overall spread during a cholera outbreak.

The cholera ABM simulation model should enable users to visualize the diffusion of the disease and should be able to switch on and off the different transmission mechanisms, in order to evaluate their different contributions. However, the cholera simulation model shall be a tool for testing the hypotheses that cholera is spread in four different ways namely EH, HEH, VT and HH and these transmission mechanisms have contributions towards the spread of cholera during an outbreak. When a cholera outbreak occurs, at least all the four different mechanisms play roles in the cholera propagation.

3.1.2 State variables and scales

Schematic overview

Important features to be modelled are individuals, houses where they stay, open space refuse dumpsites where people dump their waste, flies which breed at dumpsites and are potential vectors for spreading cholera and rivers where vibrio cholerae may naturally exist. There are different processes that will to be modelled including the flow of water (run-off) from dumpsites to the river and the diffusion of cholera amongst community members. There are various models that have to be integrated into the main cholera simulation model. These include the disease model, each transmission mechanism's model.

Fig 3.1 shows a schematic overview or conceptual framework of a Cholera Simulation Prototype. Spatial environment and a synthetic population are datasets which are the inputs. They are prepared independently then imported into the prototype. Daily activities and the interaction model enables the individuals to have activities. Rainfall data, bacteria diffusion model and run-off model will be integrated, therefore becomes an input to to the prototype. Transmission mechanisms are independent model which can work separately but each mechanism is coupled with the disease model.

UML Class Diagram

The ABM cholera prototype comprises mainly of three super classes which are agent, object and environment. The agent super class contains static agents (community and households)

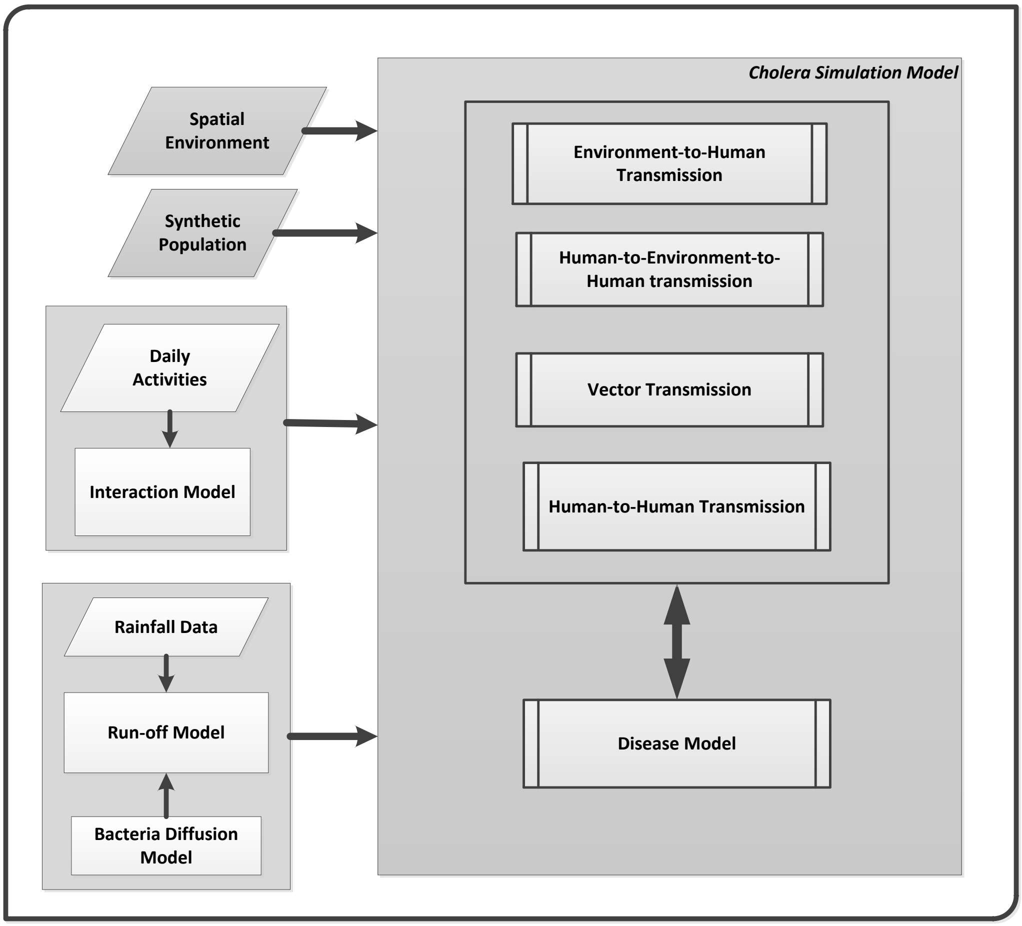


Figure 3.1: Conceptual Framework

and mobile agents, the humans (individuals), whereas the object super class has refuse dump sites, shallow wells, river network and houses subclasses. The environment super class consists of the elevation, houseflies and cholera bacteria subclasses. Fig. 3.2 presents the UML Class Diagram which represents the static model (view) for the cholera simulation prototype for all the classes containing objects with their attributes respectively.

Each individual has attributes which include age, gender, personal ID, health status, blood type and household ID where s/he belongs. The individual class inherits the attributes from household class, whose attributes are household ID, income level, level of hygiene, type of toilet they use and water source they use at a particular time. Each individual dispose their waste (faeces/vomitus) at a toilet or a refuse dump site. Each and every household stays at a house, whereby the house class belongs to the environment super class.

Musca domestica breed on the refuse dumpsites, and carry the bacteria on their sticky mouths and legs from the infected dumpsites to (uncovered) food in the house to be consumed. The food to be consumed has a contamination status which is either contaminated or not. The same applies to the water which people drink, it is either contaminated or not depending on their source, where it was collected.

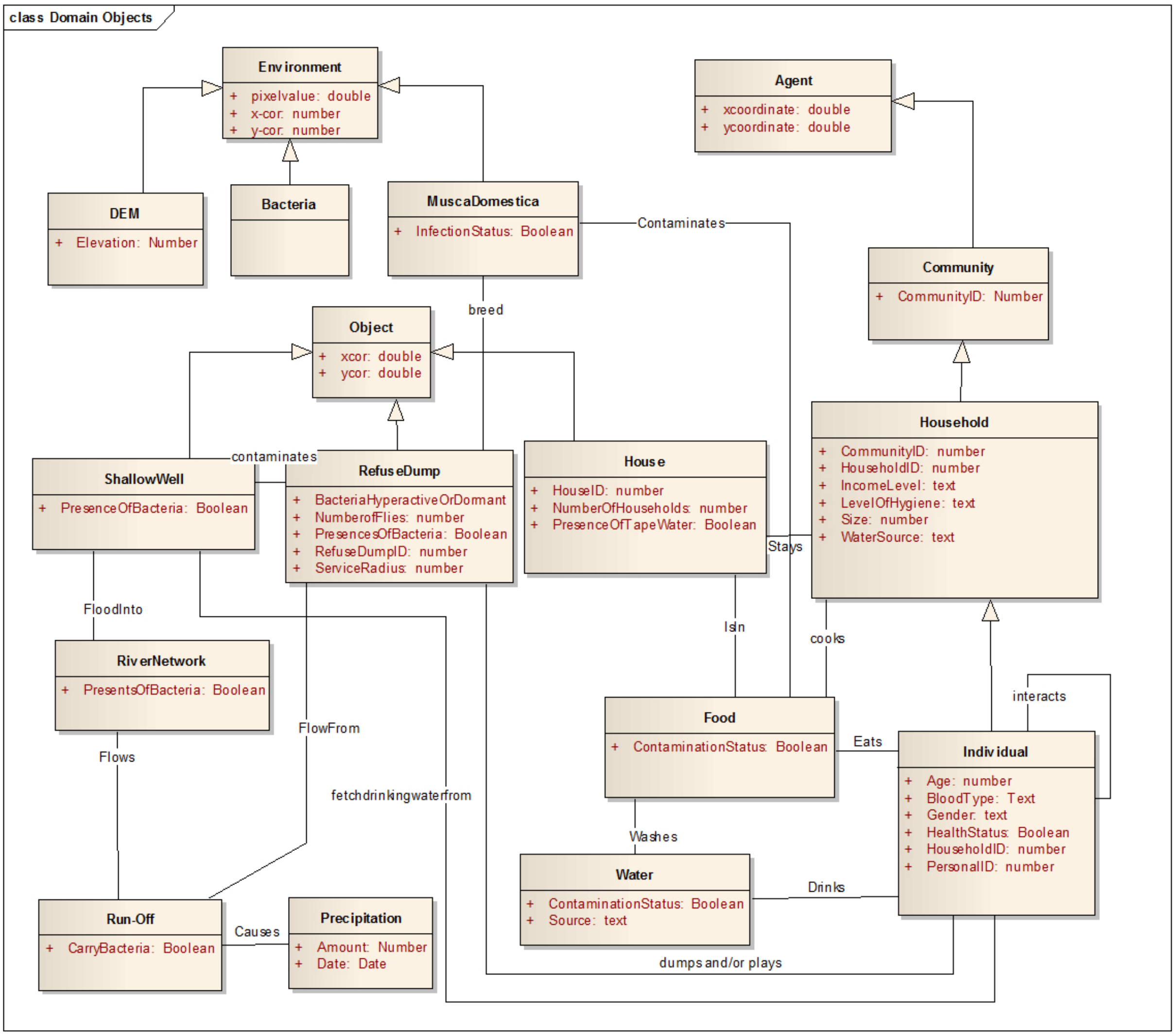


Figure 3.2: Class Diagram

Heavy precipitation triggers cholera since its run-off can pass through infected dumpsites and contaminate water bodies (rivers). The run-off serves as a pathway for faecal matter from dumpsites containing vibrio cholerae to the nearest rivers/streams where people collect their water from during water shortages. There are shallow wells located in close proximity to the rivers which become contaminated from contaminated from the river during a river flood.

3.2 PROCESS OVERVIEW AND SCHEDULING

This section will highlight the sequential processes that are essential for this cholera ABM prototype.

(a) **Disease Dynamic Model**

Cholera is a highly communicable disease spread through the ingestion of food and/or water that is contaminated by V. cholerae, hence everyone is susceptible to cholera. Individual's susceptibility is dependent on the blood type, where people belonging to the O blood group being more susceptible than others.

After ingestion of sufficient infectious dose of the bacteria, the bacteria have incubation period ranging from a few hours to five days. The main symptoms for cholera include onset of purging rice-watery stools, vomiting and dehydration. A cholera patient has to immediately seek health attention from the clinic or hospital otherwise one can die within a few hours of being infected. The incubation period of V. cholerae is as slow as 3-4 hours. An infected person fully recovers from cholera after at least two weeks. Fig 3.3 represents a flow chart of the dynamic disease model.

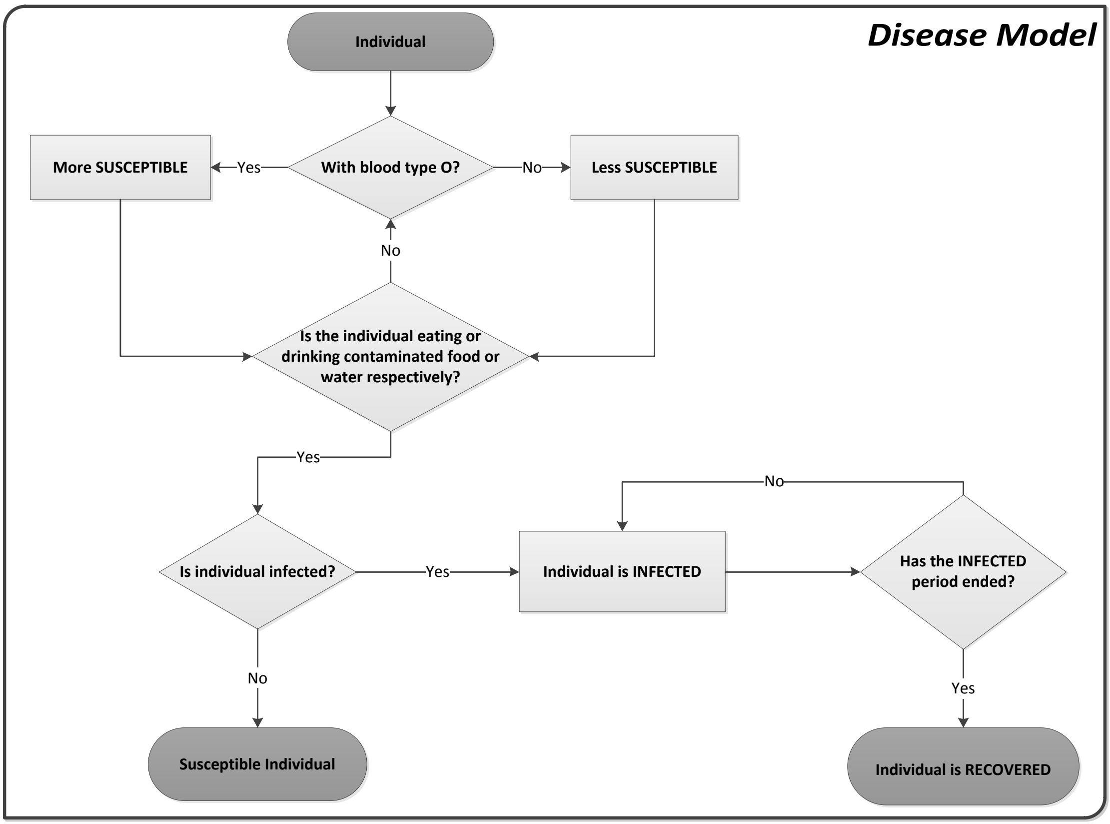


Figure 3.3: Disease Model

For over a century, epidemics have been modelled mathematically, from Louis Pasteur's work on cholera epidemics, a system of partial differential equations were developed which modelled the change in percentage of a population over time [30]. From this early model, we get the standard categories used to describe an epidemic. For this research, the cholera prototype adopts SIR (Susceptible-Infected-Removed) model. Individuals pass from their different susceptibility levels to infected phase then to the removal phase. On the Removal stage, either the cholera patient recovers or dies.

Assumptions made for this cholera (disease) model are:

- There are no cholera vaccinations.
  This is mainly because in Ghana, cholera vaccination is not common.
- Population is fixed (no entries/births or departures/deaths)

- Latent period is zero
- Infectious period = disease duration
- After recovery, individuals are immune

The time line for cholera illness is very short as compared to other infectious diseases such as Pertussis. Within a few hours of acquiring the infection, the symptoms will start showing. One can be sick for a maximum of five days before dying if s/he does not seek medical help. The disease dynamic model will reflect how the numbers of people in each compartment fluctuate with time.

(b) **Mechanisms of cholera transmission**

The mechanisms of cholera transmission identified are at least three, which are EH, HEH and HH. VT is still EH transmission, but through the flies which carry the bacteria onto uncovered food to be consumed by people. The minimal cycle includes the environment and the humans.

- **Environment to Human Transmission (EH)**

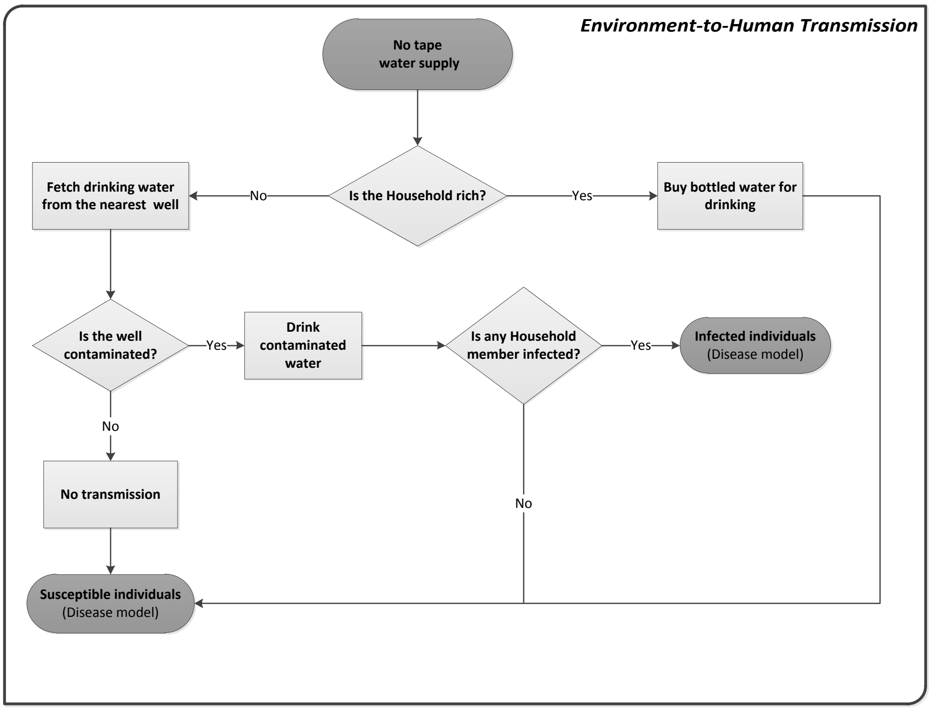


Figure 3.4: Environment to Human Transmission

EH is the primary mechanism of cholera transmission whereby V. cholerae naturally occurs in the water bodies without the presence of faecal matter. Majority of the population in Kumasi have access to tape water and a minority uses water from shallow wells which are either protected or not protected. In the event that there is water shortage, poor households fetch water from these shallow wells. Depending on the contamination status of the wells, hygiene level and income level of the households, some people may get cholera infection from drinking contaminated water. Usually, rich people can afford buying bottled water for drinking purpose. The people with high hygienic standards, will always boil the water before drinking or disinfect their drinking water.

- **Human to Environment to Human Transmission (HEH) or Faecal-Oral-Route**

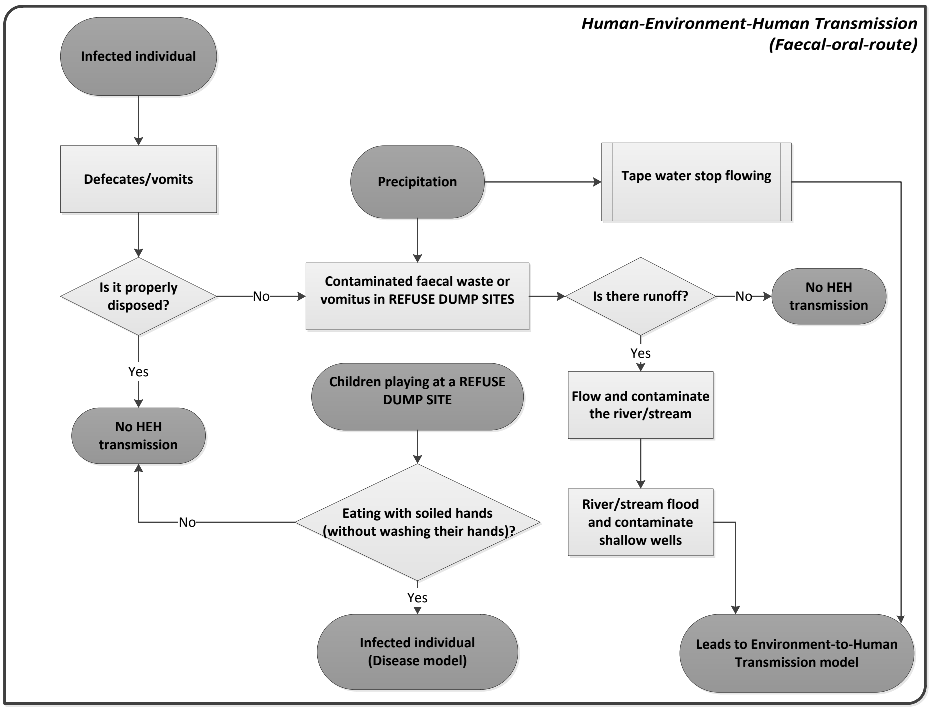


Figure 3.5: Human to Environment to Human Transmission

HEH is considered the secondary mechanism of cholera transmission. Faecal matter is dumped either in a toilet or refuse dump sites. In the event that faecal waste that is contaminated is disposed onto a refuse dumpsite, it automatically becomes infected and depending on the availability of favourable breeding conditions for bacteria, it can survive or even multiply. Young children usually play at refuse dumpsites, hence if they use their soiled hands to eat without washing them with soap, they are at risk of contracting cholera. Hands can easily get contaminated by bacteria in the environment.

There are a lot of water supply shortages whereby water stops running from the taps for some days (even upto two weeks), hence people who would not have collected enough water from the rain will fetch water from the wells. Most cholera outbreaks occur during the rain seasons hence the assumption that rain triggers the cholera epidemics. Runoff from the contaminated dump sites may carry the bacteria into the rivers/streams nearby hence contaminating them. Heavy precipitation cause rivers/ streams to flood hence overflow onto the wells nearby. The proximity of these wells to the river network is next to nothing. The water in the wells may get contaminated, hence when people drink water from them, they are at risk of contracting cholera infection.

- **Vector Transmission (VT)**
  Musca domestica breed at open space refuse dumpsites. If a dumpsite is contam-

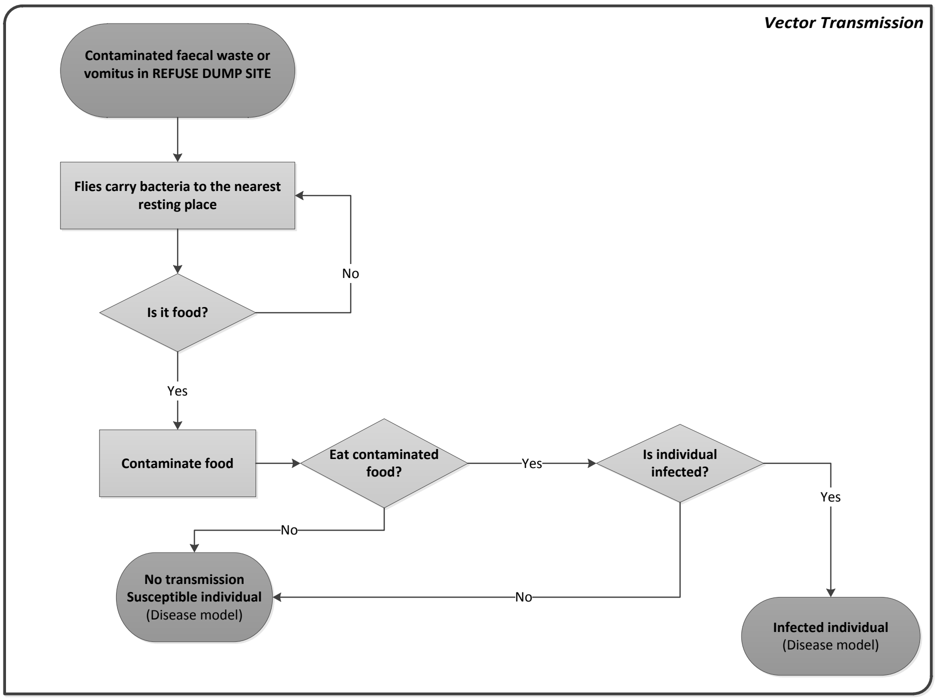


Figure 3.6: Vector Transmission

inated, as these flies fly away to the nearest resting place, possibly to a house onto food, there is high risk of the food getting contaminated. They are usually attracted by the smell of food in the house. They are found in their greatest numbers in summer months. Musca domestica can fly as a group to surrounding areas within a distance of about 500m and within this buffer distance, cholera transmission can occur.

- **Human to Human Transmission (HH)**
  When one is suffering from cholera, s/he can hardly do things for herself/himself, hence one of the family member is more likely to assist. If this care giver does not wash her/his hands with soap and clean water, and prepares some food or gets in contact with food, there is high risk of the food becoming contaminated with cholera bacteria. Whoever eats the contaminated food can get infected. Fig 3.7 shows a flow chart of Human-to-Human Transmission

(c) Hydrological Model

The hydrological model's design is based on the Grand Canyon Model found in NetLogo Models Library which simulates rainfall on a patch of terrain. This model will be extended such that it includes the bacteria diffusion model.

Conceptual description of Grand Canyon

Elevation was prepared as a raster layer with pixels of size 32m, and the whole area represented is 9.7 km by 9.7 km. Turtles (moving agents) are raindrops and water. Raindrops fall randomly and will flow to patches of lower elevation. The raindrops always flow downhill unless if there is no neighbouring patch with lower elevation. Raindrops have a height, therefore a raindrop can only move to a patch of lower elevation if that patch does not have other raindrops with height exceeding its height

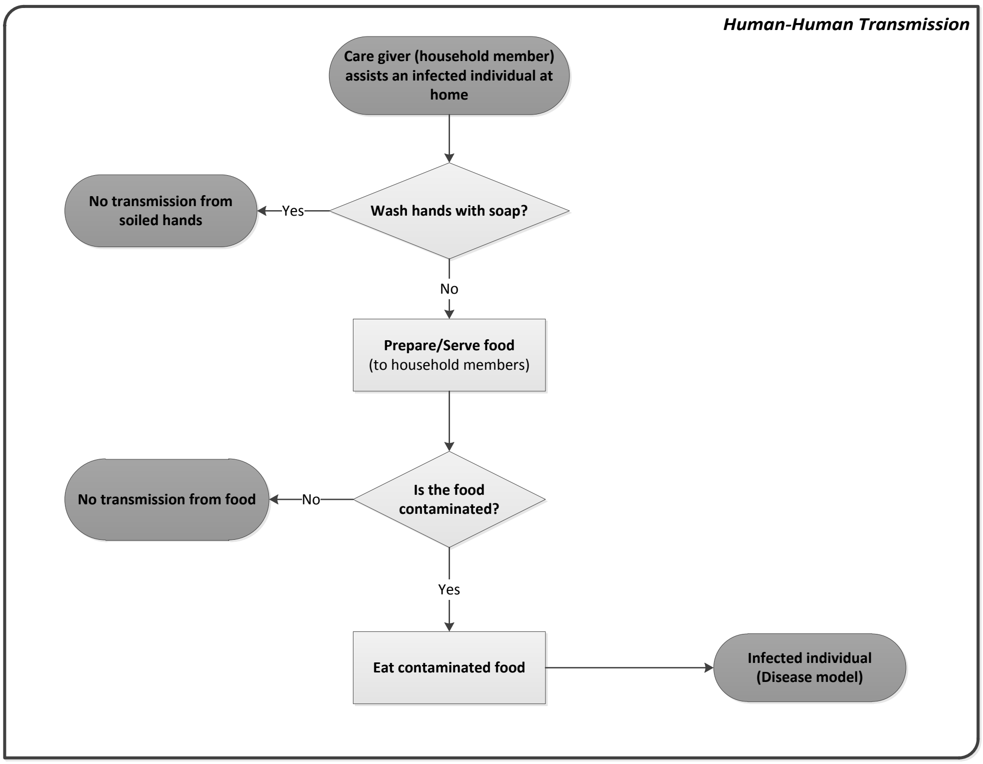


Figure 3.7: Human to Human Transmission

together with its patch's height.
[HeightofPatchA + HeightOfRaindrop(s)onPatchA] >
[HeightofPatchB + HeightOfRaindrop(s)onPatchB]
Raindrop(s) on patch A can only move to patch B if patch B's total height (together with the raindrops) is less than patch A's total height.

Extension part to be implemented in the Cholera Simulation Model

Additions to Grand Canyon model will be for the rain drops that fall onto or that flow through a dumpsite that is infected with vibrio cholerae, there is higher chance of the water drops to become contaminated. Water particles surrounding a contaminated particle will become contaminated. As the run-off flow to the river, it should go round the the buildings (houses).

(d) **Drinking behaviour**

Household members drink water inside the house which would have been fetched by some of the household members. If they have a tap, then generally they will fetch it from the tap unless if there is no water running from the tap, then most likely it is fetched from the shallow wells nearby. On average everyone is supposed to drink about 2 litres of water per day, therefore if the drinking water is contaminated then there is high risk of getting infected.

(e) **Eating behaviour**

Food is prepared and served a plate per household individual. Depending on the availability of soap (determined by the income level of the household), individuals may eat without washing their hands, this applies to children who can come from playing outside the house and eat with soiled hands. Depending on the hygiene standards of a household, some food can be contaminated by houseflies and some people wash fruits and vegetables before serving them, whereas others just eat them unwashed due to ignorance.

(f) **Cooking behaviour**

Households have different ways of cooking/preparing different types of foods. Some foods may be served raw, some half cooked and some fully cooked. For some foods that might be contaminated may be half cooked hence not destroying the bacteria. Availability of cooking energy influence they way some people prepare their food. Households with low income will not utilise a lot of energy (electricity), therefore income level of the household is a factor that influences how people prepare their food. Half cooking food using contaminated water may not kill the vibrio cholerae, hence increasing the risk of getting infected. Food should always be covered to avoid contamination from infected houseflies. People with high hygiene standards will not leave food uncovered.

3.2.1 Activity diagram

The activity diagram below represents the daily activities of the individuals, and a day has been divide into ten time steps. Depending on the ages of the people, they have been assigned to different tasks, the only common tasks are eating and drinking which are done by everyone, both sick and not sick. Fig 3.8 represents the daily activity diagram

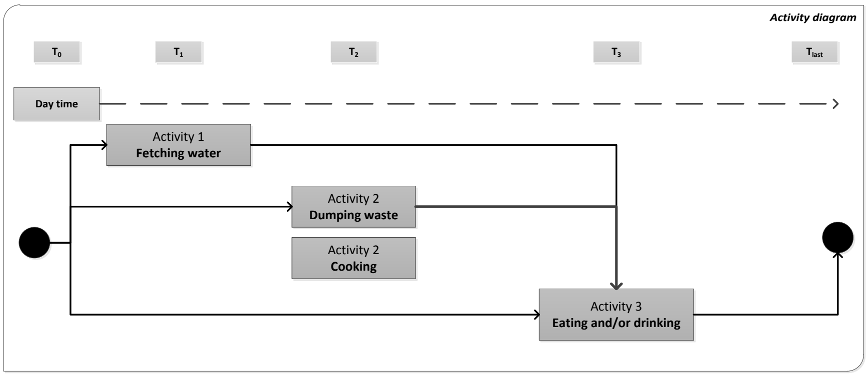


Figure 3.8: Daily Activity Diagram

3.2.2 Sequence Diagram

UML Sequence diagrams are interaction diagrams that show the sequence of events, they show how processes operate with one another and in what order. Global variables are updated at each time step, daily every family has to fetch water for household use, dump refuse at a dumpsite, cook, eat and drink water. These processes make the basic cycle.

Time is modelled as discrete time-steps, (for this thesis, the prototype will use ten steps per day).

Everyday agents have a routine of daily activities which are fetching water and disposing waste at an open space refuse dumpsite unless when the individual is ill. Those who are ill remain at home until they have recovered. Food is cooked everyday, and the individuals eat and drink daily irrespective of their health status. Children under 13 years occasionally play at dumpsites. When there is a sick patient in the household, someone assists him/her. At every time steps the model checks and updates all the global variables.

Rain falls once in two days and the rain drops will form run-off which will flow to patches of lower elevation till they get to the river. If there is raindrop that is contaminated, the other particles nearby will also change their status to contaminated.

Fig 3.8 represents the Sequence diagram.

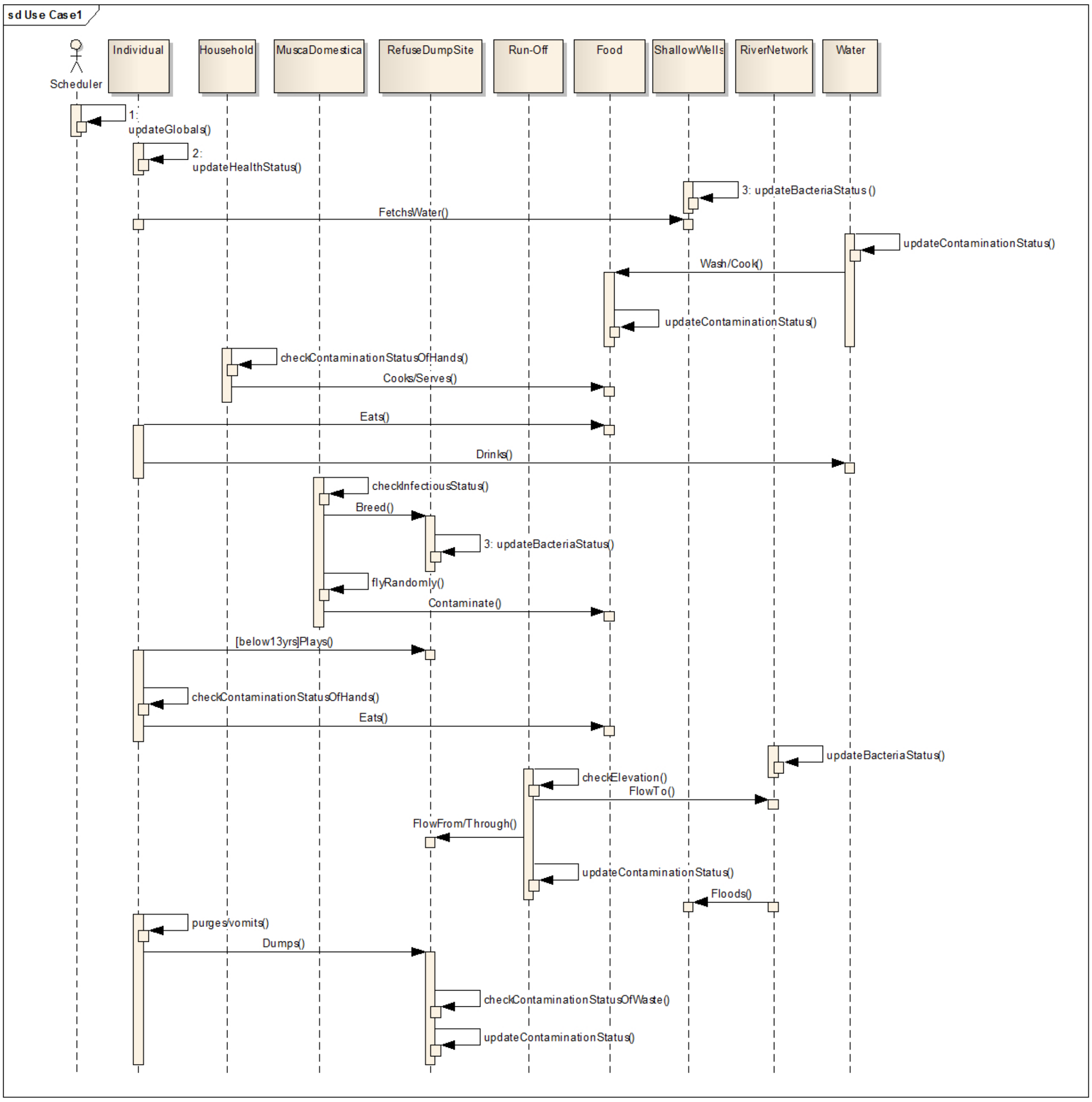


Figure 3.9: Seqeunce diagram

Chapter 4

Data preparation

This chapter focuses on data collection, preparation of the spatial environment and generation of a synthetic population. The process of conceptual designing is crucial to the data collection and preparation since it serves as a guideline to which datasets are essential and can be processed, for their output is necessary for the development of the prototype. There are two categories for data required and these are spatial datasets and non-spatial datasets which are further divided into demographic data and rainfall data. The spatial datasets are required the generation of buildings layer and elevation data (for hydrology model). Whereas, demographic data is crucial for the generation of a synthetic population.

4.1 DATA COLLECTION

Spatial data

Spatial datasets are necessary for the creation of spatial environment. The spatial environment will consist of raster layers for elevation data and shapefiles for spatial objects. Spatial objects needed are buildings (houses) where people stay, river/streams where people go to fetch water and open space refuse dumpsites where people dump their wastes. The shapefiles for rivers/streams and open refuse dumpsites were adopted from a PhD research titled “Spatial statistics of epidemic data: the case of cholera epidemiology in Ghana” by Frank Osei. (Rivers/streams were digitised from a topographic map whereas the locations of refuse dumpsites were collected by a GPS). Both the feature layers cover the whole of Kumasi Metropolitan.

Buildings/houses shapefile had to be digitised from an image, therefore a google earth image (of 2006) for Aboabo community (Fig. 4.1) was downloaded from Google Earth. Prior to the downloading of the image, the rivers/streams shapefile and refuse dumpsites shapefile were converted kml files and the kml files were launched onto Google Earth. The whole of Kumasi is too large an area to simulate for this project, therefore a single community that has a river network and streams running through, and that is neither too upstream nor downstream was chosen. Open refuse dumpsites are also supposed to be visible within this community of choice. Another crucial factor considered was the visibility of houses on the image, the image should have very minimal cloud cover. For the visual appearance of the houses, their house structures should be different, which would imply that there are different residential density suburbs, that is with families of high, medium and low income levels. The assumption is that people of the same income levels tend to stay in the same area with almost similar characteristics. The community that satisfied the mentioned prerequisites is the Aboabo community, hence the google image downloaded is for this community.

Figure 4.1: Google Image

For the elevation data, a DEM was downloaded from CGIAR (srtm) website as a Geotiff image. The DEM covered the whole of Kumasi area and has pixel sizes of 90 metres.

Non-spatial

Demographic data and Cholera statistical data
In order to generate a synthetic population, demographic statistics are required. A Census report for 2000 census was collected from the Frank Osei's research datasets.

In 2005, there was a cholera outbreak in Kumasi, and aggregated statistics for cholera cases that were reported per community.

Rainfall data
Since the cholera outbreak occurred from September to December, monthly total rainfall data (July - December 2005) for Kumasi was downloaded from NOAA (ncdc) website.

4.2 DATA PREPARATION

The datasets that were collected need to be processed such that they are ready for intake in the simulation model. Fig 4.2 is a flow chart showing summary of sequence of steps followed in the data preparation. Spatial environment was prepared first, then a synthetic population was generated then was distributed over the geographic space.

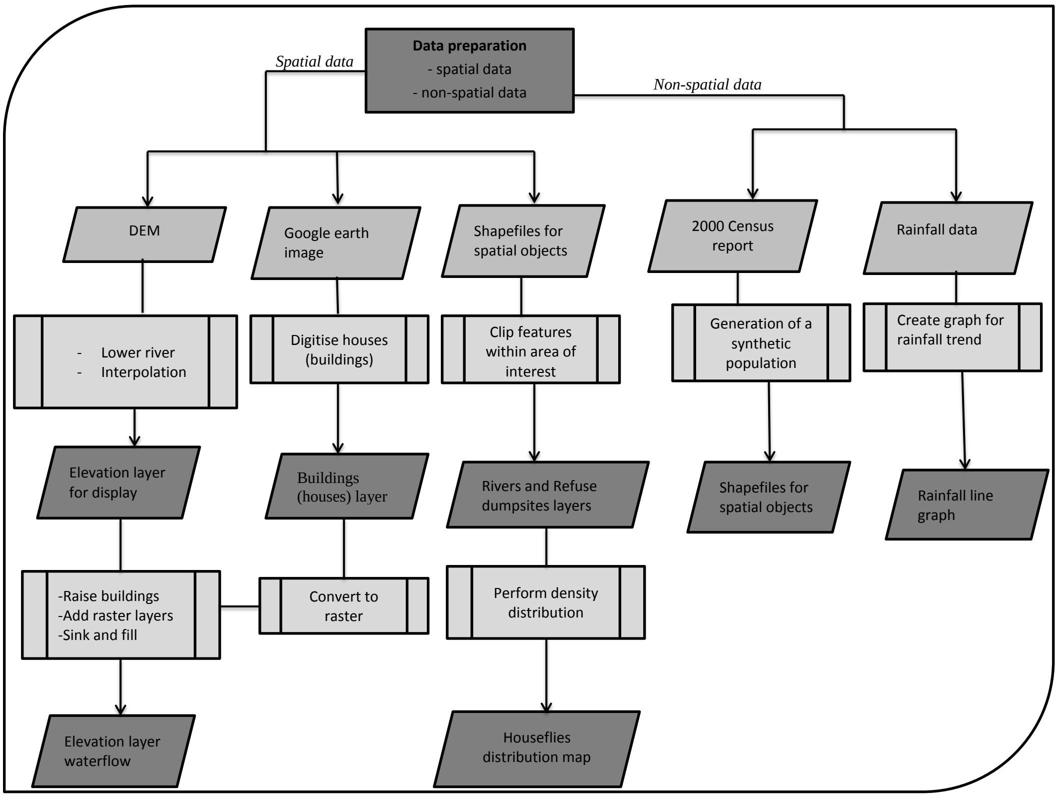


Figure 4.2: Flow chart for data preparation

4.3 SPATIAL ENVIRONMENT

All the environment datasets were prepared in ArcGIS 10, ArcMap and ArcCatalog.

4.3.1 Preparation of shapefiles for spatial objects

The shapefiles for the rivers/streams and open space refuse dumpsites that were collected had no metadata about their coordinate system nor a reference ellipsoid. When loaded in ArcMap, their units were in metres and written documents revealed that they were using Transverse Mercator coordinate system. They were then assigned to Accra Ghana Grid (Ghana coordinate system). Later, the shapefiles' coordinate systems were converted to CGS_WGS 1984, and then they were changed to kml files which were launched onto google earth. These kml file assisted in choosing the appropriate community to use in the simulation. The google image for Aboabo community was downloaded.

There are no boundaries to demarcate where a community starts nor ends. The communities are represented as points on google earth image. A bounding box was created in ArcMap and was used to clip refuse dumpsites and river within the bounding box covering Aboabo community.

Preparation of buildings layer
Since there was no digital layer for buildings, a buildings layer had to be created. Google earth image for Aboabo settlement was georeferenced and houses (buildings) were digitised from the image, but the image was very haze, hence it was difficult to accurately digitise the buildings. From the image it was difficult to tell by naked eyes whether a building is a house where people reside or the building is for any other use. Each building has a building ID and a boolean showing either presence or absence of tape water.

Fig 4.3 shows rivers and communities in Kumasi and the position of Aboabo community. There is an inserted map showing the map of Aboabo community, the rivers, open refuse dumpsite and the buildings.

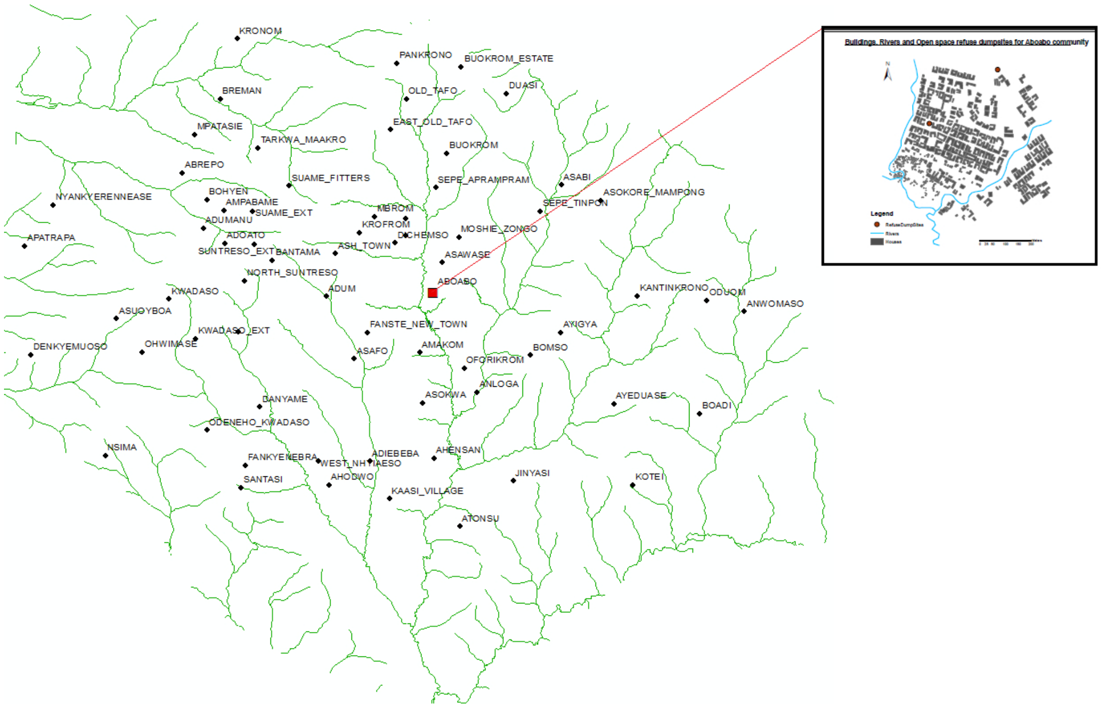
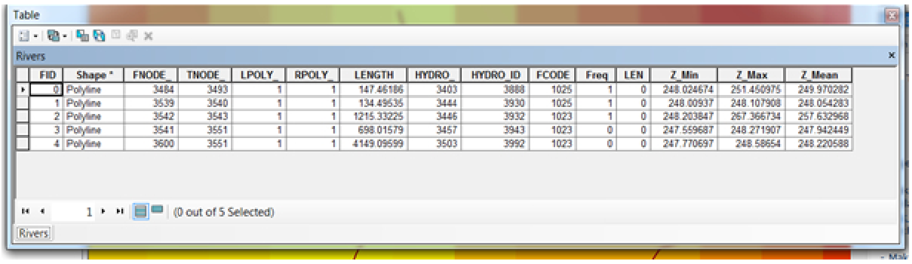


Figure 4.3: Aboabo suburb in Kumasi

4.3.2 Preparation of elevation data

Determining heights of each river segment
The river layer was made up of five segments. Using 3D Analyst extension, and the DEM as the height (elevation) base layer, the maximum and minimum heights of each river segments were determined. This helped to identify the flow of the river.

Extraction of points from river feature
Points from the river layer were extracted using the Feature Vertices to Point tool. Each line segment has its own number of vertices.



FID	Shape *	FNODE_	TNODE_	LPOLY_	RPOLY_	LENGTH	HYDRO_	HYDRO_ID	FCODE	Freq	LEN	Z_Min	Z_Max	Z_Mean
0	Polyline	3484	3493	1	1	147.46186	3403	3888	1025	1	0	248.024674	251.450975	249.970282
1	Polyline	3539	3540	1	1	134.49535	3444	3930	1025	1	0	248.00937	248.107908	248.054283
2	Polyline	3542	3543	1	1	1215.33225	3446	3932	1023	1	0	248.203847	267.366734	257.632968
3	Polyline	3541	3551	1	1	698.01579	3457	3943	1023	0	0	247.559687	248.271907	247.942449
4	Polyline	3600	3551	1	1	4149.09599	3503	3992	1023	0	0	247.770697	248.58654	248.220588

1 (0 out of 5 Selected)

Figure 4.4: Attribute Table

Making the heights of river points to be in descending order such that the water flows down
A new field was added to the river points attribute table and I artificially lowered the points of each segment using the formula New Height = ([FID] * 0.0669)
All the points from the five segments were loaded into one point shapefile.

Making the river line to have lowest elevation than the surroundings
Subtracted 2 meter in height from all of the height values of the river points.

Extraction of points from the DEM
Raster to Points Tool was used to extract points from the DEM (base) layer srtm 36 11.tif.

Combining points
Using the Append tool, the points extracted from raster (DEM) and the points from river were combined into a single point shapefile.

Interpolation - Creating new raster layer from points
Using the Spatial analyst extension, the points were interpolated using the Inverse Distance Weighting method (IDW)and reduced the output pixel size.

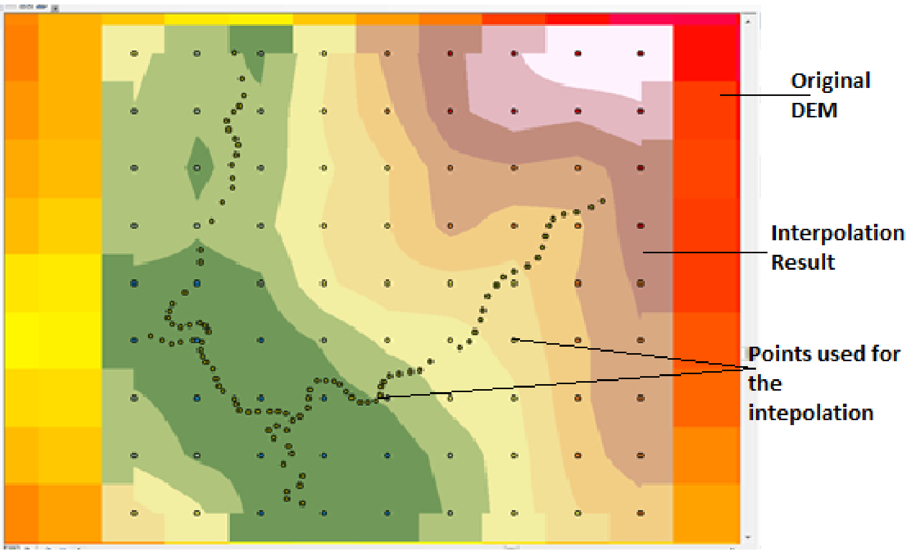


Figure 4.5: Elevation Raster

As shown by Fig 4.5, the course of the river is not fully described by the elevation raster layer, therefore I further decreased the values of the river points by 3 metres and performed the interpolation process again. Fig 4.6 shows the output and overlayed the river shapefile.

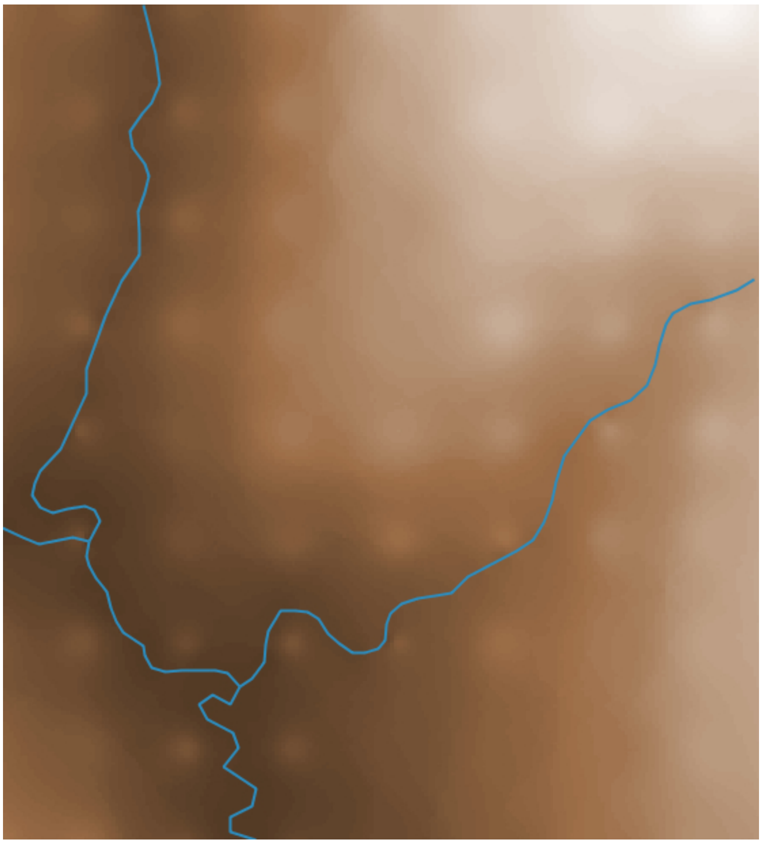


Figure 4.6: Elevation Raster for display

An attempt was made to check whether the river runs from higher elevation to lower elevation, river profiles were created.

Creating a river profile

- A new river layer is created in ArcCatalog and specifying that it should contain Z-values.
- Using 3D Analyst and newly created elevation layer was used as the reference of obtaining the Z-values of the river, an interpolate line tool was used to digitise the river in two segments.
- Drawing profile graphs of the river
  Two profiles were created using Create profile graph tool on the 3D Analyst tool-bar, one for each segment.

Conversion of raster layer to ASCII format
Using the Raster to ASCII tool, I converted the elevation raster layer to ASCII format which is a format that is acceptable in NetLogo. The above step produced a text file and a prj file. I then wrote down the number of columns and rows and the cell size from header of the text file, then deleted the header. The text file was ready for input in NetLogo.

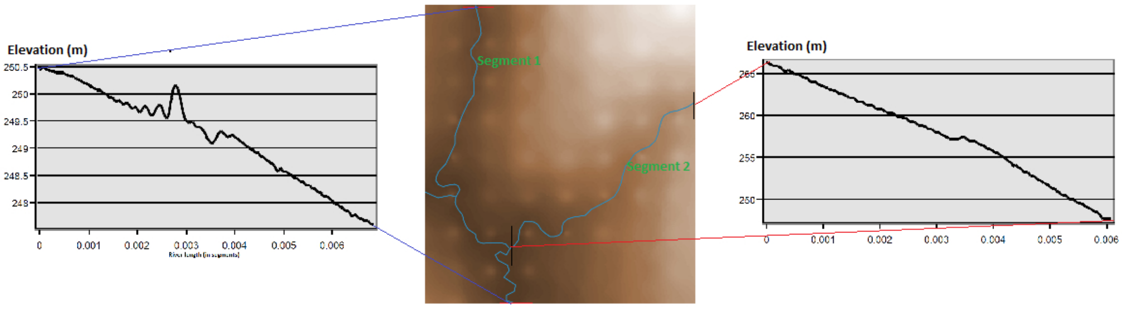


Figure 4.7: River profile

4.3.3 Preparation of raster layer for hydrological model

This is a step further to the above procedure of preparing the elevation data. In the procedure, areas with buildings were raised (increased their elevation) such that when water flows, it does not flow thorough buildings and also to avoid natural dips in the terrain which may hold and keep the water particles.

Preventing water from flowing through buildings

Raising areas with buildings
Polygon to raster tool was used to convert the buildings' polygon shapefile to a raster layer and fixed the extent of the raster to be the same as the extent of our study area map (bounding box).

Classifying the raster layer of the houses
Reclassification was done to the output raster, where pixels with no data were assigned a zero value, and the buildings pixels were assigned a value of 3. Therefore all buildings were raised by 3 meters.

Adding two raster layers
Addition of raster layers was done using a raster calculator. The new classified raster layer was added to the elevation raster layer (used for the display). The resulting layer does not allow water to flow through buildings.

Preventing water from flowing through buldings

Identifying sinks and filling the dips
In order to prevent water from accumulating in terrain dips, the dips have to be filled up hence the following steps were carried out using tools in hydrology toolbox.

Creating flow direction
A flow direction raster layer was created using Flow direction tool. This flow direction raster layer indicated the flow direction of water and it enabled the detection of sinks or dips in the elevation data.

Identification of dips
Using the sink tool, all the dips in the elevation layer were identified.

Filling dips

Using the Fill tool, all the dips in the elevation layer were filled accordingly such that no sinks will exist in the terrain, therefore all run-off should flow without accumulating at some areas. This makes the area relatively flat. Fig 4.8 (output layer) is a raster layer that will be used for the water flow, areas where there are buildings have a very light color and areas along the river channel have the darkest color.

Figure 4.8: Raster layer for water flow

Conversion to ASCII file

The raster layer for the water flow was converted to ASCII file using the raster to ASCII tool. This is to prepare for input into NetLogo.

4.3.4 Preparation of Density Distribution Map for Houseflies

Houseflies' populations are large and difficult to model as discrete individual agents. Largest number of flies are found at the open refuse dumpsites where they breed, and they fly randomly to all other places. However, for this research, the author has assumed the population of houseflies to be constant and was represented as a raster layer (each pixel with a certain number of the houseflies).

Some people dump waste at open space refuse dump sites whereas some are not hygienic enough they dump waste at any open space, hence another assumption is that the populations of houseflies are directly proportional to the amount of waste at that particular place. At dumpsites, a large amount of waste is disposed hence we expect a large number of houseflies to breed there.

This research will utilize a static "Density distribution map" for the distribution of flies over different reference areas depending on the locations of favourable habitats of houseflies and assumptions mentioned above.

Points were first extracted from the rivers, and assigned random numbers for the population of houseflies but the population are lower than at dumpsites. The two point layers were loaded into a single point feature layer. Using the Kernel density tool in Spatial Analyst Toolbox, the raster layer for the distribution was created. Output map is represented in Fig 4.9

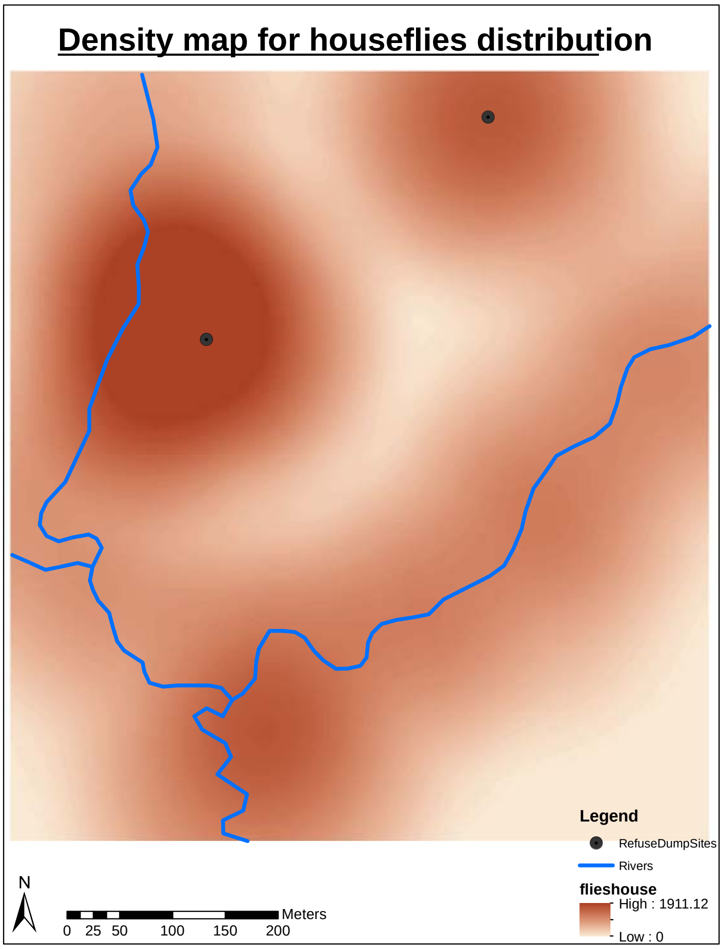


Figure 4.9: Density Distribution Map of Houseflies

4.4 SYNTHETIC POPULATION

In order to try to accurately simulate how diseases spread at local level over a given population at a particular area, demographic data for that area is required. Real population data are typically limited to census and/ or register data. Statistical agencies are prevented from releasing detailed geographical identifiers in public-use data sets due to privacy and confidentiality concerns [67]. Indeed, it is very difficult to get individual data (micro data) about household structures since the institutes managing statistics usually provide aggregated datasets, describing the global properties of the households and individuals. Data collection is often very expensive. However, the remedy of this problem is to generate synthetic populations, which artificial datasets of individuals and households "synthetic population"can be created from the aggregated data to represent a realistic population. Synthetic population consists of individuals described by their age and are gathered in households respecting a variety of statistical constraints (distribution of household types, sizes, age differences among parents and children, age of household head and many more) [36]. Attributes of individuals within a generated synthetic population include their age, gender and many more. These synthetic populations can be created from existing survey data or aggregated statistics from a census.

There are different existing approaches of developing synthetic populations which include stratified sampling, geodemographic profiling, data fusion, data merging, iterative proportional fitting, reweighting [2], disaggregation of zonal data to raster data and many more. The choice of a method to use should depend upon the data available and also the quality desired. For more accurate artificial populations, different methods are implemented as an integral that is, an output of one method will an input of another method. Examples include Synthetic reconstruction approach which combines IPF and Monte-Carlo sampling.

Examples of methods for creating synthetic population

Iterative proportional fitting (IPF) is a method used to create a synthetic population when aggregated statistics from census data describe properties of individuals and households are integrated with individual data extracted from surveys on samples that can be bigger or smaller than the size of the desired artificial population [36].

Monte-Carlo sampling method allows generation of multidimensional artificial populations with many (infinite number of) different characteristics (attributes). The input demographic statistics can be from census data or administrative registers. A household is generated by first selecting age of household head, and depending on that information, a household size is selected [56].

Disaggregation of zonal data to raster data This is when "GIS is used to generate a synthetic disaggregate spatial micro database which corresponds to all known statistical distributions. This is achieved by when raster cells are used as addresses for the microsimulation" [56].

A combination of the three above mentioned methods were used by Moeckel et al. [56] to generate synthetic population for Netanya, Israel.

4.4.1 Generating a synthetic population for Aboabo community

Synthetic Population Generator

The following procedure (steps) from Monte-Carlo Sampling procedure [56] was created to allow development of the synthetic generator.

(a) Generation of a population of individuals with age, gender and blood type, into a list.
(b) Randomly pick a household size
(c) Random choose a female figure head aged from 15 years to 60 years.
(d) Random pick the other household members from main list of individuals.
(e) A complete household will then randomly pick an income level for its family

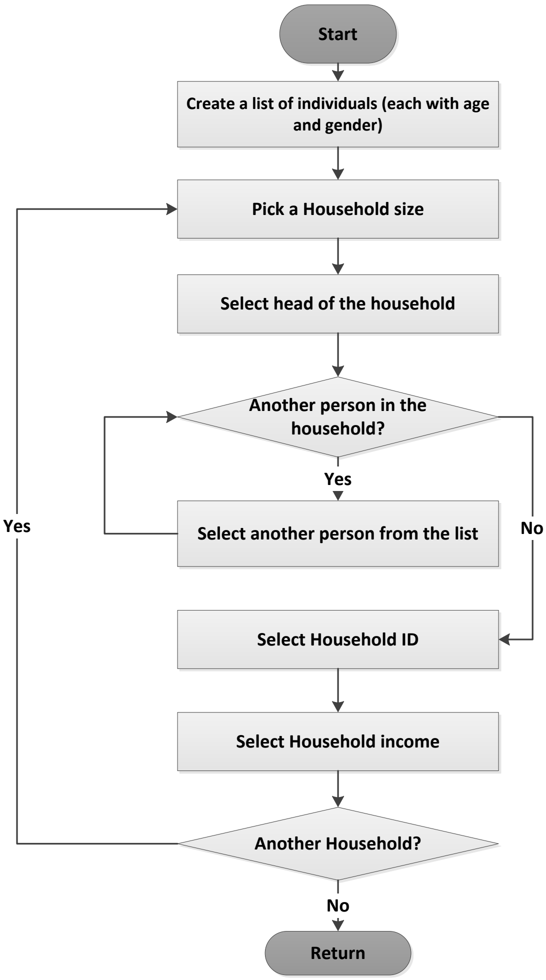


Figure 4.10: Synthetic population generator

4.4.2 Creating a list of individuals with different age groups

Due to data limitations, for this research, Monte-Carlo Sampling method was adopted to generate the statistically realistic populations of households for Aboabo community respecting the following aggregated statistical constraints from 2000 Census report.

Table 4.1: Census statistics

Age (yrs)	Percentage (%)
0-4	13.2
5-9	12.4
10-14	11.4
15-19	11.3
20-24	10.9
25-29	9.3
30-34	7.1
35-39	5.5
40-44	4.5
45-49	3.4
50-54	2.5
55-59	1.8
60-64	1.6
65 and above	5.1

3 501 people with ages were created into a list. Fig 4.9 is a bar graph above showing the total statistics of people in each age group and is compared to the age structures from census report.

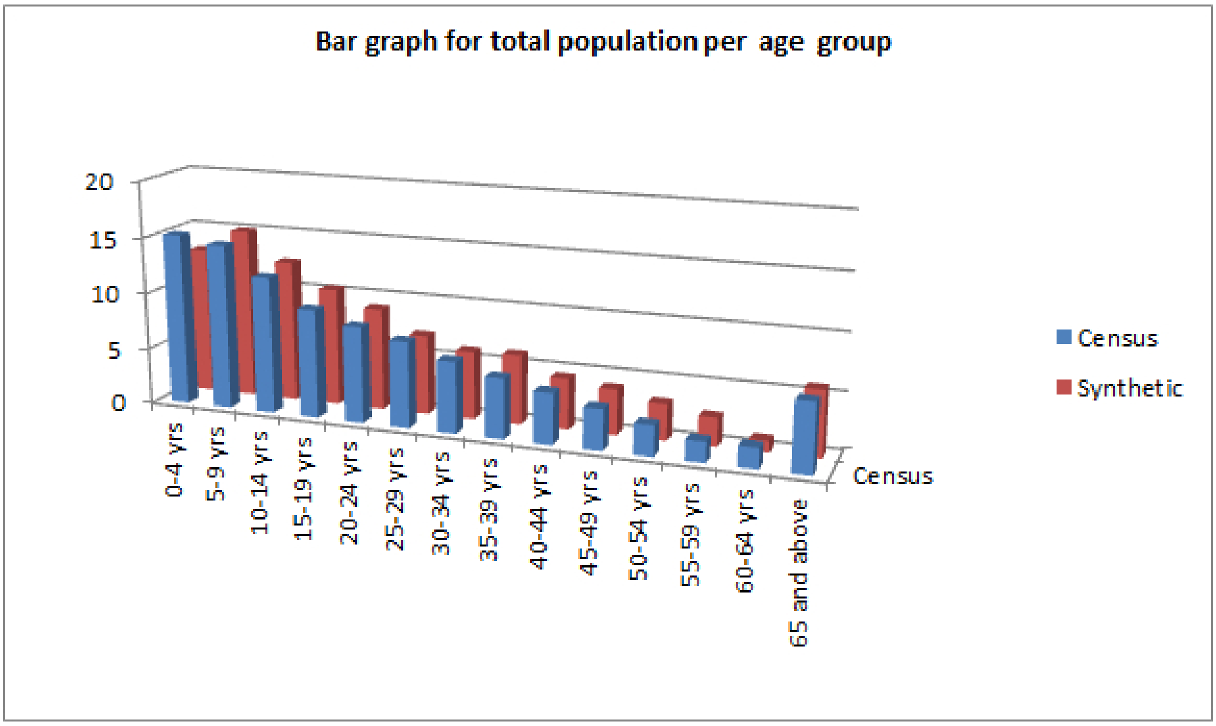


Figure 4.11: Bar graph

According to different literature on blood types (groups), from national blood services and sample surveys done in different countries it has been found that approximately between 40% and 50% of the total populations are of blood type O (that is both O+ and O-). However, for this research 45% of the total population has been attributed to belong to the blood group O.

Generation of families

From the 2000 census report, an average household size for Kumasi Metropolitan is 5.1. Starting from the age of 15, they become economically active, therefore I have chosen a female figure starting from 15 years, depending on the household size, other family members are randomly picked from the list of individuals. If a household is complete, it is assigned a household ID, income level and level of hygiene. The assignment is random. The attributes of a household are inherited by its household members.

The synthetic population generated is presented in an excel sheet, an example is shown in Table 4.2. Each individual has attributes of age, gender, blood type and belongs to a family (household).

Table 4.2: Synthetic Population

PersonID	Age	Gender	BloodType	HouseholdId
1589	9	M	other	6
2236	3	F	O	6
2863	48	F	O	6
1664	19	M	other	7
1823	36	F	O	7
2125	68	M	O	8
2928	34	F	other	8
3282	24	M	other	8
2046	17	F	O	9
2544	17	F	other	9
2706	7	M	other	9
2837	49	F	O	9
2941	12	M	O	9
2461	46	F	other	10
2522	17	M	other	10
2761	26	M	O	10

The total numbers of each household were also calculated and are presented in Table 4.3

Table 4.3: Households' characteristics

HouseholdID	Total (Frequency)	Hygiene Level	Income Level
1	6	middle	low
2	1	high	aver
3	2	low	high
4	4	low	aver
5	5	middle	low
6	3	high	aver
7	2	middle	aver
8	3	low	high
9	5	high	high
10	3	high	low
11	3	low	aver
12	6	middle	aver
13	1	high	low
14	5	middle	high
15	6	high	low
16	3	low	low
17	6	low	middle

4.5 SPATIAL DISTRIBUTION OF HOUSEHOLDS

After the generation of the synthetic population, the population was manually distributed over the existing house (building) layer that was prepared from digitising. The Aboabo residential area was divided into three distinct regions, assumed as low, medium and high income levels. The delineation was done based on the visual appearance of the building structures. All houses that are intersecting low income level class were allocated to families with low income status the same was done for high income level. This was mainly done based on the idea that people or families with similar characteristics tend to live close to each other.

The houses (buildings) in the low income area are very small, this makes it difficult and unrealistic to allocate more than a single household per house. Exceptions were buildings which looked a bit bigger, in such a case two or three households were allocated per house. This is because most people (families) generate income from rentals.

More households were allocated per house in the middle income area than in the high income area. According to Kumasi Metropolitan Website [6], 74% of households rely on rental housing. This explains why many households stay at the same house (building).

Fig 4.11 shows the distribution of individuals who are grouped into households per house.

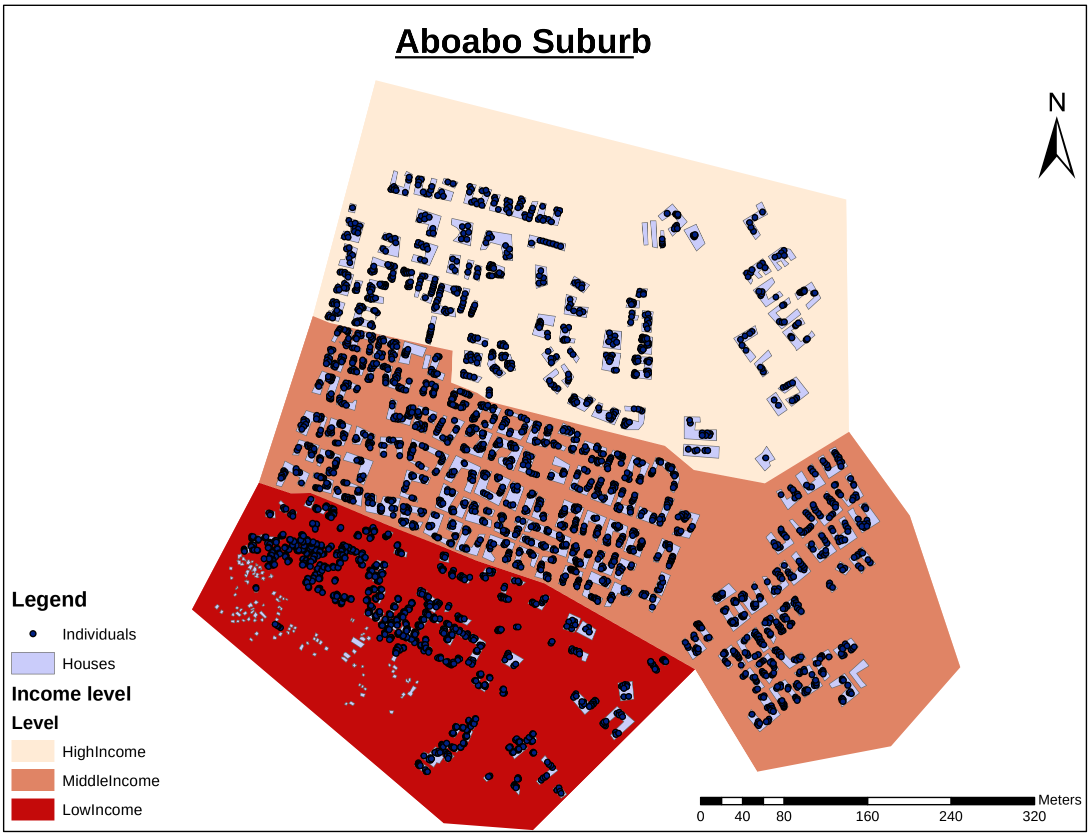


Figure 4.12: Population distribution over the houses

4.6 PREPARATION OF RAINFALL DATA

The period when 2005 cholera outbreak occurred, coincides with the rainfall season for Kumasi, Ghana which reaches its peak in September/October. From the rainfall data collected, and the author extracted data from September to December. The total amounts of rainfall per day were plotted against the date when it rained. Average amount of rainfall received during rainfall days is 80mm.

The first cholera case was reported around the 29th of September, a day after seven consecutive days of raining. After that it literally rained at least once a day.

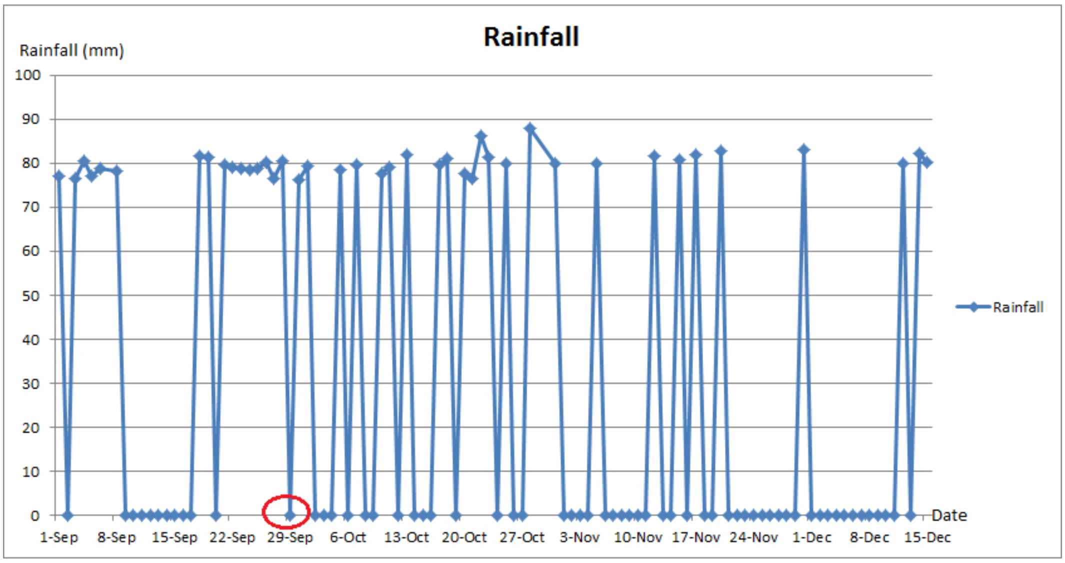


Figure 4.13: Rainfall

Chapter 5

Implementation of Cholera Simulation Prototype

5.1 NETLOGO

Implementation of the prototype was done in NetLogo 4.1.1.
NetLogo is a multi-agent programmable modeling environment for simulating natural and social phenomena and was authored by Uri Wilensky and is in continuous development at the CCL (Center for Connected Learning and Computer-Based Modeling) [90]. It uses Logo programming language. NetLogo world is made up of four types of agents which are turtles, patches, links and the observer. Turtles are mobile agents that move around in the NetLogo world. The world can either be two dimensional (2D) space that is divided into a grid of patches which are square pieces of ground over which turtles can move or three dimensional (3D). Links are agents that connect two turtles whereas the observer is an overseer of the whole world of turtles and patches. Both patches and turtles have coordinates, but coordinates of patches are only integers whereas turtles' coordinates can have decimals [90]. Different types of turtles can be defined in NetLogo and they are called turtle breeds.

In order to develop a new model, one has to design a user interface.

5.2 DESIGN CONCEPTS

Designing of the user interface

A user interface is the space or system where people and a digital machine (information systems) interact, allowing input and/or output of information. For every application development, the user interface is one of the most important parts of the program since it determines how easily one can make the program do what the developer wants. Depending on the users of the program, the complexity or simplicity of how to operate the user interface must be taken into consideration.

For this cholera simulation model in NetLogo, the author designed the user interface as presented in Fig 5.1. The prototype's user interface was designed in such a way that the user can switch on and off cholera transmission mechanisms, it can be extended thereby allowing intake of other input files with a different projection or coordinate system. Different files for initial population are also acceptable. Therefore, there are two "Choosers", one for projection and the other for initial population.

There are four buttons, one for loading the projection (Load Projection). The second button is the "Setup display" button which loads and displays (initialise) the environment, the agents and the objects. The "Go" button runs the prototype. The last button is the "Infect" button which randomly infect twenty individuals. Its purpose is to introduce the illness to the population.

There are four switches one for each of the four transmission mechanisms. They can be switched on and off, depending on what the user wants to be active/non-active during the runs of the model. The transmission mechanism will enable the disease to be transfered from an infected person or environment to people through the specific different ways (EH, HEH, VT or HH).

There are also four "Sliders" which enable users to set values (parameters) prior to the simulation runs. The variables to be set are:

(a) Risk of children getting infected whilst playing at a dumpsite.
(b) Risk of flies getting/carrying the infection from a dumpsite.
(c) Risk of human to human transmission.
(d) Number of ticks to represent a day.

For visualization as the prototype will be running, the users can see the movements of the individuals on the graphics window (which is 2-D viewing/display area). The histories of variable's values are visualized as plots or graphs. The numbers of patients infected through either of the mechanisms will be reported on monitors. Behind the interface, there are procedures which have been coded using logo programming language to perform the required tasks.

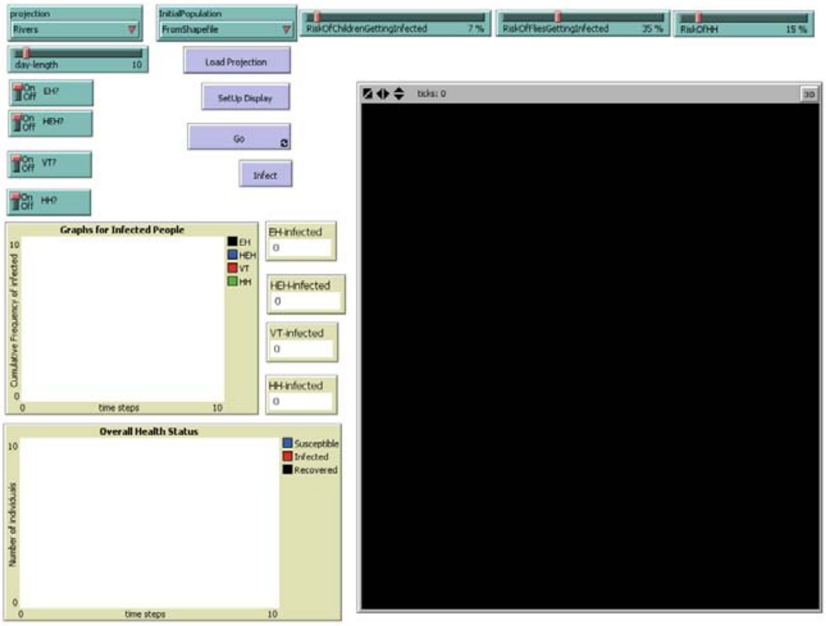
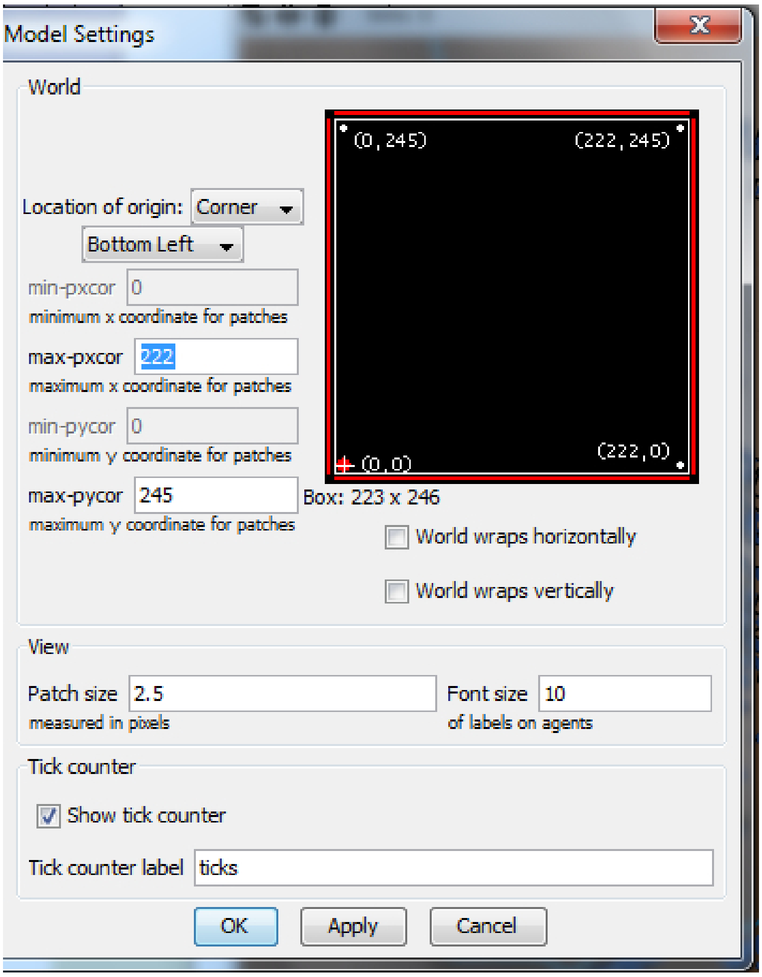

projection: Rivers
InitialPopulation: FromShapefile
RiskOfChildrenGettingInfected 7 %
RiskOfFliesGettingInfected 35 %
RiskOfHH 15 %
day-length 10
EH-infected 0
HEH-infected 0
VT-infected 0
HH-infected 0
ticks: 0

Figure 5.1: User Interface

5.3 LOADING OF THE SPATIAL ENVIRONMENT

NetLogo has a GIS extension which provides the ability to load vector GIS data (points, lines and polygons) which are in ESRI shapefile (.shp) format. The extension supports raster data in the form of ESRI ascii Grid files (.asc or .grd). There are two ways of loading raster layers (as the environment) into NetLogo. The first one is loading them using the GIS extension and the raster layers should be ASCII files with an .asc extension. The second way is of loading the raster layers (ASCII files) with a .txt extension. Both ways, raster data is loaded into the patches of NetLogo world. For this prototype, the second method was adopted because after conversion of the raster layers to ASCII files in ArcGIS, they came out with a .txt extension.

(a) Defining the environment (Model settings)
First and foremost, it is crucial to define the transformation between NetLogo space and GIS data. The world is defined with the properties of the Elevation data in terms of its numbers of columns and rows. The ASCII file will be read from the bottom-left corner, the number of rows and columns are defined excluding the point of origin, and therefore both numbers of rows and columns are lowered by one. The point of origin is (0;0). Fig 4.2 shows the Model settings screen shot.



Model Settings

World

Location of origin: Corner

Bottom Left

min-pxcor 0
minimum x coordinate for patches

max-pxcor 222
maximum x coordinate for patches

min-pycor 0
minimum y coordinate for patches

max-pycor 245
maximum y coordinate for patches

Box: 223 x 246

☐ World wraps horizontally

☐ World wraps vertically

View

Patch size 2.5
measured in pixels

Font size 10
of labels on agents

Tick counter

☑ Show tick counter

Tick counter label ticks

Figure 5.2: Model Settings

(b) Load raster layers (ASCII files) The following code was used in NetLogo to read the text file.

```
file-open "elevationn.txt"
```

```
foreach sort patches [ask ? [ set terrain file-read
set pcolor scale-color brown terrain 230 280 ]]
file-close
```

The first line is an instruction to open the ASCII file called “elevationn.txt”. The second line is reads the elevation values of each patch. Third line is displaying the elevation using a brown color ramp, where by the higher elevated areas will be light brown and the lower elevated areas will have a dark brown color. Last line closes the text file.

The same was done for the houseflies’ distribution raster layer but the houseflies’ raster layer is not exhibited, it is not going to be visualized in this prototype. Therefore the same code was used but excluding line 3 of displaying.

Note: All ASCII (raster) files should have the same pixel size, same number of rows and same number of columns.

5.4 LOADING OBJECTS

The objects’ layers used in the prototype are houses, refuse dumpsites and rivers/streams and they were prepared in ArcMap as shapefiles. The shapefiles are loaded into NetLogo using the GIS extension and they are defined using the WGS84 coordinate system. But the files containing the projection and coordinate system have names similar to the names of their respective shapefiles, therefore on the projection chooser, I have chosen the Rivers. Choosing either refuse dumpsites or houses will not make any difference.

5.5 LOADING THE SYNTHETIC POPULATION AND CREATING INDIVIDUALS TURTLES’ BREED

NetLogo provides a functionality of creating turtles in two ways, by instructing either the observer or by patches. If it is the observer making turtles it is called creating turtles but if it is a patch making the turtle, the patch sprouts turtles. Turtle can be randomly created or can be created at specific coordinates. For this prototype, a point shapefile was loaded and the observer was instructed to create turtles called individuals by reading the shapefile.

A synthetic population that was created was distributed into houses (over geographic area) in ArcMap and was stored as a point shapefile. The initial population for this prototype is loaded into NetLogo as a shapefile using the GIS extension. The observer in NetLogo reads the shapefile and instead of displaying the individuals as point graphics, it is creating the individuals with their respective positions at a house and being part of household as exactly read from the point shapefile. Their attributes are read from the dbf file of the shapefile. This was done for easy and efficient manipulation of the attributes, the individuals were (assigned) transferred attributes of their respective households and the house where they stay. The following code was used.

```
set Individuals-dataset gis:load-dataset "data/Individuals.shp"
gis:set-world-envelope-ds (gis:envelope-union-of (gis:envelope-of Individuals-dataset)
                                                 (gis:envelope-of Houses-dataset)
                                                 (gis:envelope-of Rivers-dataset))
```

```
;create individual agents from point shapefile
foreach gis:feature-list-of Individuals-dataset
[
  let location gis:location-of gis:centroid-of ?
  if not empty? location
  [
    create-individuals 1
    [
      set homeLocation location
      set size 3
      set Ages gis:property-value ? "AGES"
      set personID gis:property-value ? "PersonID"
      set Gender gis:property-value ? "Gender"
      set HouseholdI gis:property-value ? "HouseholdI"
      set BloodTyp_1 gis:property-value ? "BloodTyp_1"
      set incomeLeve gis:property-value ? "incomeLeve"
      set TapeWater gis:property-value ? "Tapewater"
      set hygieneLev gis:property-value ? "hygieneLev"
      set infected? false
      set recovered? false
      set immune? false
      set carrying-contaminated-water? false
      set recovery-period 0
      set HH-infected? false
      set VT-infected? false
      set HEH-infected? false
      set EH-infected? false
      set xcor item 0 homeLocation
      set ycor item 1 homeLocation
      set color blue
```

5.6 ENABLING TURTLES (POPULATION) WITH BEHAVIOUR

The individuals have a sequence of activities that they perform on daily basis. Due to lack of data about daily activities of the people in Kumasi, this prototype will simulate a few daily activities that were considered to be relevant for the diffusion process. The activities have been assigned to individuals according to their ages and income levels. Each activity is perform at a specific tick of a day and is independent of the length of a day. In real life, adults are responsible for collecting water for household use and most children at times play at refuse dumpsites. This mostly influence the decision I made of making people aged fifteen and above have the responsibility of fetching water while children are assigned the task of dumping household wastes.

As the model is initialised, shortest distances to the nearest open space refuse dumpsite and place of fetching water at the river (nodes) are calculated on the fly (in the setting up procedure). Each individual is made to remember his/her home (where s/he stays), where to fetch water and dump waste. Therefore, every agent has a home location, water location and refuse dumpsite location respectively.

All individuals aged fifteen years and above, fetch water daily at time-step one of each day unless they are ill. Children aged between five and twelve years go to the dumpsite to dump household wastes daily (during the second time-step of each day) unless they are ill. Occasionally, some of the children play at the dumpsites especially children from families of low income. At the same second time-step of the day, food is prepared (created) by a household and is provided (served) a plate per individual. At third time-step of each day, everyone eats and drinks despite his health status.

5.6.1 Fetching water

Agents who reside at houses where there is no tape water and belonging to households of low income level collect water for daily household use (cooking, bathing and washing utensils) from the river (nearest node/vertex). For agents with access to tape water at their houses, there are times when water stops coming (running) from the tape due to rampant water shortages, the individuals belonging to middle income leveled families are fetching water from the river once in 2 days. For agents belonging to families with high income level are fetching water from the river once in 3 days. Assumption made is that all people will fetch water from the river.

Depending on the hygiene standard of their respective households, the assumption made is that families with high hygiene standards will treat the water before drinking and those with low hygiene standards may not treat the water before use.

Collection of water is done in time step 1 of each day, those collecting the water will appear and disappear at the water collecting points/nodes. When they appear at the water location, the water they will be carrying home will have contamination status of the water at that specific water location.

Figure 5.3: People fetching water in a cycle of 3 days

5.6.2 Dumping waste

Solid, liquid and at times human faecal wastes are dumped at an open space refuse dumpsite (nearest to a household). Children aged between five and twelve years are responsible for the dumping of household waste at time step 2 of each day unless if he/she is ill. Usually children from low income families play at these dumpsites, they pick up waste materials such as plastics to make plastics balls. These plastics balls are used to play street football. Most girls will pick up containers or tin to use when playing childrens' out-door games. Upon picking dirty materials, their hands can get soiled (contaminated) by vibro cholerae if that dumpsite is infected.

At every second time-step of the day, children appear at the dumpsite and will disappear from there after. For children who would have played at the dumpsite, if the dumpsite is contaminated, they have a certain chance whereby their hands can get soiled. This value is set/controlled by the slider titled "RiskOfChildrenGettingInfected".

5.6.3 Cooking

At time step 2 of each day, food is prepared by a household, no specific person is responsible for preparing the food. The reason for providing food at time step 2 of each day is to give time for the food to get contaminated by houseflies before being consumed.

5.6.4 Eating and drinking

Every individual has to eat and/or drink at time step 3 of each day, and it is independent of the health status of the person. Everyone is provided by their own plate of food whereas, for the water drinking, people drink water that is available at the home. Therefore, agents cannot see whether the food or water is contaminated.

Table 5.1 gives a summary of the daily activities and their corresponding time (tick) at which they are performed.

Table 5.1: Daily ticks and activities

TICK	ACTIVITY
1	Fetching water
2	Dumping waste and Cooking
3	Eating and Drinking

5.7 DISEASE MODEL

At the beginning of the simulation, every individual is susceptible to cholera infection, people with blood group 'O' being more susceptible than others. If one gets the infection, and once they recover, they then become immune to cholera till end of the simulation run. People can get the cholera infection through any of the four different ways which are EH, HEH, VT or HH unless they are infected manually by the button named "Infect". The use of this button is to randomly introduce the infection to the population, in the event that the primary transmission mechanism (EH) is off which introduces infection to the community.

If the user is considering external factors which may introduce illness to the population such as migration, the the button can be used. It infects 20 people, the assumption is that the infection will be spread evenly hence at least one of the dumpsites will become infected in order for VT and HEH transmission mechanisms to start working.

The sick individuals fully recover from cholera after at least two weeks (14 days) to 18 days. Each agent belongs to either "Susceptible" or "Infected" or "Recovered" category.

At the beginning of the simulation, everyone is susceptible to cholera infection and the susceptibility level is dependent on the blood group of each individual. Hygiene levels and income levels make people get exposed to the infection. For an individual who gets cholera infection, s/he goes through three transitional phases as presented by Fig 5.5

Figure 5.4: Susceptibility-Infectious-Recovered

Assumptions

(a) Population is fixed (no entries/births or departures/deaths)
(b) Latent period is zero
(c) Infectious period = disease duration
(d) After recovery, individuals are immune

The total numbers of each category are plotted on the graph titled "Overall Health Status". Susceptible and recovered people do their daily chores whereas the sick people stay at home. All infections through either EH or HEH or VT or HH occur only at time step 3 when people will be eating and/or drinking. To eliminate re-infection of the already infected and the recovered, it has been made a condition to all the transmission mechanisms that no one can be re-infected.

As the infection occurs independent of the mechanism of transmission, **dumpsite infection level** also goes up. Dumpsite infection level indicates the number of sick individuals who use that open space refuse dumpsite. Therefore, for each infection of an individual, either of the dumpsites infection level increases by one but not for both.

5.8 SUB-MODELS

There are four sub-models that have been integrated together with the daily activity model and disease model which are the four mechanisms of transmission of cholera and they are operated using switches which can be turned on or off. A transmission mechanism can only work when its switch is on.

EH Transmission

EH transmission occurs when naturally existing vibrio cholerae in water bodies are consumed by a person through either drinking contaminated water or eating fish that carries the infection. During summer time when temperatures are high more V. cholerae exist and the bacteria multiply.

Households who do not have access to tape water (no tape at their houses) mostly staying at residential areas for people with low income (mostly slums) fetch water for household use from the river. These people are fetching water at nodes (vertices) created during digitising of the river. These vertices have 50% chance of being infected or not. Since these turtles do not see the cholera bacteria by naked eyes, they just get their water and go back home. If the node is infected, the individual carrying the water will have his/her attribute “carrying-contaminated-water”, true. The following piece of code was used.

```
ask individuals
[ if Ages >= 15 and TapeWater = 0  and incomeLeve = "low"
  [
    ifelse ticks mod day-length = 0 and not infected?
    [
      let locationX item 0 waterpointLocation
      let locationY item 1 waterpointLocation
      setxy locationX locationY

      let contam? false
      ask patch-here
      [
        set contam? contaminated-water?
      ]
      set carrying-contaminated-water? contam?
    ]
    [
      let locationX item 0 homeLocation
      let locationY item 1 homeLocation
      setxy locationX locationY
    ]
  ]
]
```

Line 1 and 2, the observer is asking individuals who are 15 years and above, who do not have tape at their houses and who belong to households with low income level. These two lines, the observer puts these individuals in an agentset (set of agents). Line 4 is giving the condition, which is whenever a day starts (first tick of each day) and if that individual from the agent set is not infected (lines 6,7,8) should appear at the water location at the river (the nearest node). During initialisation, the individuals were made to remember their water locations and home locations. When at the water location, each node is either contaminated or not. The contamination status of a node can be known by asking a patch directly below it. Therefore the patch’s contamination status is made to be the contamination status of water carried by the individuals “carrying-contaminated-water?”. When it is not the first time-step of a day, the people should be home. All the people who fetch water at the same node will carry water with contamination status of the node.

For individuals who are carrying contaminated water, when it comes time to drink the water, depending on the hygiene level of the household, blood type of the individual, they have different chances of contracting cholera. People from families with low hygiene level and who are of blood group 'O' have 30% chance of getting infected, whereas people from the same families with other blood type, have 15% chance of getting infected. People from households with average hygiene standards and with blood group 'O' have 10% chance of getting infected.

Any person who gets the infection through drinking water from the river that was contaminated through this way, is recorded as EH-infected. If an individual is sick, s/he stays at home until recovered then they can resume their daily chores of fetching water.

EH submodel is an extension of the fetching water procedure, then the result is carried forward to the drinking water procedure. Disease model will give the results.

HH Transmission

If one of the household member is infected by cholera, the other member will assist (help) the sick by cleaning them since s/he will be weak. As they will be cleaning them, depending on their income level, those with high income level they use gloves and disinfectants which include soap. Depending also on the hygiene standard of a household, the care givers have to wash their hands thoroughly after helping. Care givers with less hygiene at times might not wash their hands carefully hence their hands can keep the bacteria. When they touch food during eating, they can get the infection.

Susceptible turtles with low hygiene that are within a radius of one pixel from an infected agent that are at home (assuming they are helping the infected) are exposed to HH transmittance. The risk of getting infected through HH is set using the slider named "RiskOfHH", on the user interface and this parameter applies to individuals with low income, low hygiene and have 'O' blood group. People with low hygiene and other blood groups have half chance of the value set "RiskOfHH" of contracting cholera through HH infection. Individuals with average income level, low hygiene standard and blood type 'O' have one third of the risk of getting cholera through HH set on the slider.

People who contract cholera through this way are recorded as HH-infected. The income level of a household is very crucial to this transmission mechanism since it determines whether they can afford to buy disinfectant (e.g. soap) to use when washing hands.

Vector Transmission (VT)

At every second tick of each day, solid, liquid and human faecal wastes are dumped at the dumpsites. Each dumpsite has an infection-level which increases when a susceptible person who uses that dumpsite gets infected. When a dumpsite has a dumpsite-infection-level of at least 10 implying that there are at least 10 patients that are suffering from cholera in the community who use that specific open space refuse dumpsite. The dumpsite becomes infected and patches around it become red in color.

When a dumpsite becomes infected the assumption is that a substantial amount of contaminated faecal waste has reached that dumpsite. An infected dumpsite infects the filth breeding houseflies, the risk of houseflies getting infected is set using the slider named "RiskOfFliesGettingInfected", which is on the user interface. "RiskOfFliesGettingInfected" percent of patches (pixels) containing a threshold of 850 houseflies has a that chance of transmitting the infection to (contaminating) the food prepared and served at that patch.

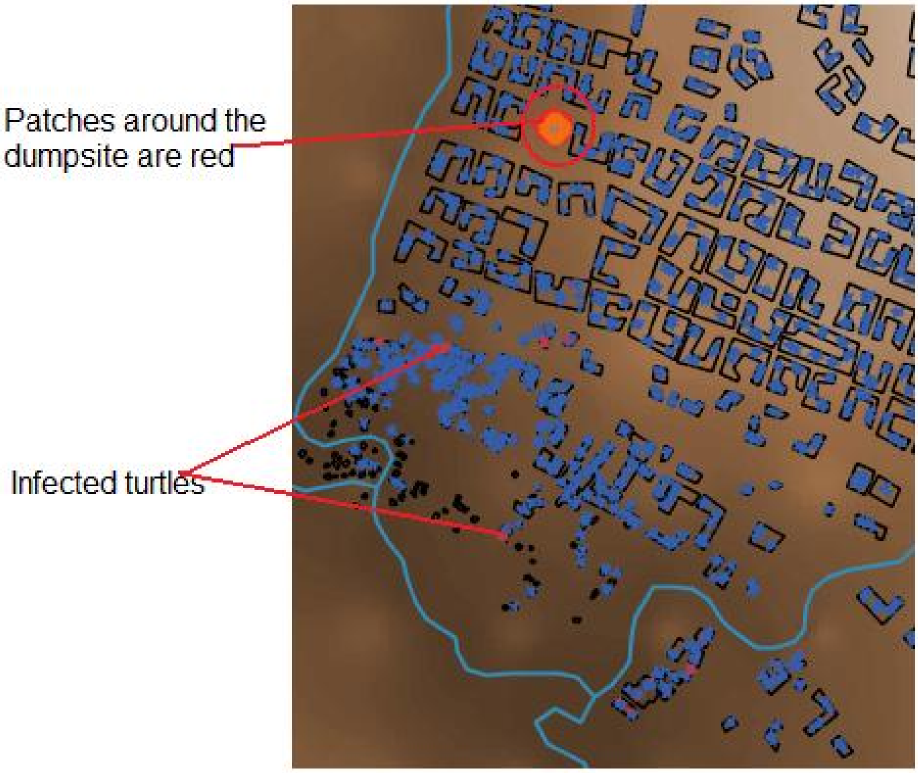


Figure 5.5: Infected dumpsite

```
if VT?
[
  let nrflies 0
  ask patch-here
  [
    set nrflies houseflies
    if (nrflies > 850.000) and (random 100) < RiskOfFliesGettingInfected
    [
      let prey one-of foods-here
      ask prey
      [ if random 100 < 30
        [ set contaminated? true
```

The piece of code above is asking if the switch VT? for vector transmission is on, then the observer checks the pixel value from a layer of houseflies distribution, if the value is above 850, the patch has “RiskOfFliesGettingInfected” percent chance of being infected. **A group of houseflies contained in a patch is modelled as a single agent**.

Upto 30% of the foods in an infected pixel can get contaminated but the contamination is also dependend on the standard of hygiene of a household. If a household has high hygiene level, then no food will be contaminated. People get infected when they are eating contaminated food. Contracting the disease depends on the blood type of an individual, leaving those with ‘O’ blood type more susceptible.

Turtles with low hygiene level, blood type ‘O’ and low income have 20% chance of getting the infection after eating food that has been contaminated by houseflies. The same percentage applies to individuals from high income level families and poor hygiene, this is like this because they have a wide variety of food stuffs they eat which include fruits and vegetables which are eaten raw or half cooked. Agents from middle income level families but with low hygiene have 10% chance of getting infected after consuming contaminated food.

VT submodel is an extension of the cooking procedure, then the result is carried forward to the eating water procedure. Disease model will give the results of the infected.

HEH Transmission

(a) *Hydrological model*

This sub-model (transmission mechanism) contains a hydrological model which is a modification of an existing hydrological model called “Grand Canyon” found in the NetLogo Models Library.

An average of 80 mm of rainfall literally poured once in two days and is being represented by 200 raindrops which are falling randomly over the whole area at time-step 3 of a day but (raining) once in two days. The observer creates the water particles and each drop checks from its 8 neighbouring pixels for a patch with the lowest elevation value and moves to that patch until it reaches the river (with lowest elevation).

The pixel (patches) size of the elevation data is 3.5 metres and the whole area covered is approximately one square kilometre. For this hydrological model a water particle is moving a distance of one pixel at each time step (1 pixel per tick). Taking into consideration the speed of run-off flow in reality, there are three options to increase the speed of raindrops which are:

- Making the particles skip a number of pixels per tick.
  This makes the particles to move a certain distance (patches) within the required time.
  Disadvantage - one cannot visually track the flow of the particle.
- Increase the number of ticks per day.
  A particle will flow for a longer in a day.
  Disadvantage - the model takes longer time to run and becomes slower since it will be having a large number of turtles.
- Increase the pixel size.
  Meaning will have to resize the pixels for the three raster layers loaded into the prototype.

The piece of code below was used to define and instruct water particles how they should flow:

```
to flow
 let target min-one-of neighbors [waterflow + (count turtles-here * water-height)]
  ifelse [waterflow + (count turtles-here * water-height)] of target
     < (waterflow + (count turtles-here * water-height))
    [ move-to target ]
    [ set breed waters ]
```

A particle checks its 8 neighbouring pixels, for the value of the patch and and volume of water in each of the pixels (Line 2). If the sum of value of pixel and volume of water of the neighbouring pixel is less than the sum of its pixel value and volume of the particle, then it can move to that neighbor otherwise if equal or greater, the drop should remain in that pixel and change its breed to water.

If a particle falls onto or passes through dumpsite pixels which are red (infected), it changes its contamination status to "contaminated? true" and its color becomes red.

When a raindrop gets to (reaches) the pixel of lowest elevation, it changes its breed to water breed. A breed is a type/group of turtles that belong together.

As a particle is infected, it will carry vibrio cholerae down to water bodies (river). A limitation of using such small pixels, it takes a long time for a particle flowing from the furthest pixel (from the river) to reach the river. Therefore, for all particles that are not contaminated may not reach the river in a day ("day-length"), they disappear by either evaporation or infiltration. Those which reach the river will also disappear to avoid flooding. Only the infected raindrops reach the river no matter how long they take and remain in the system. If uninfected particles collide with an infected particle, it becomes infected as well.

If a raindrop falls onto a patch without any neighbouring pixels with lower elevation (due to natural dips), it stays on that patch until it disappears (evaporates). This effect of natural dips has been minimised by "Sink and Fill" effect performed on the elevation (data) raster layer. Some raindrops which flow towards or fall at the edges automatically get out of the system.

People who fetch water at a node that has a contaminated (red) rain particle within 1 patch radius (3.5 m), carry home water that is contaminated. Drinking the contaminated water caused HEH - infection.

(b) *Infection of children playing at dumpsites*

This transmission occurs only when a dumpsite is infected, meaning the waste is contaminated. Children aged between five and twelve years and belong to households of low income level play at dumpsites occasionally (that is once in five days), and have a chance of getting soiled hands which is set on the slider named "RiskOfChildrenGettingInfected", on the interface. When they have low hygiene level, they have 10% chance of contracting cholera when they use their dirty hands to eat (without washing their hands thoroughly with soap and clean water). Those who get infected are recorded as HEH-infected but their source of infection will be a dumpsite.

Chapter 6

Verification, Results and Discussion

6.1 VERIFICATION

The purpose of the verification stage (phase) is to run and verify the simulation. Therefore, the modeler runs the simulation and inspects whether the simulation program has been coded correctly from the conceptual design [44]. The coding of this simulation prototype was done in seven major stages which are:

(a) Loading the environment and agents
(b) Enabling agents with daily activities
(c) Implementing disease model
(d) Coding the EH transmission mechanism
(e) Coding the HH transmission mechanism
(f) Coding the VT mechanism
(g) Coding the HEH transmission mechanism

NetLogo provides a function of checking the syntax after writing a procedure (before running it). This check helps to identify syntax errors from the code.

6.1.1 Expectations of the prototype

The expectations of the simulation prototype are to visualise individuals performing their daily activities - fetching water, dumping waste, eating food and drinking water. Cholera infection has to spread after ingestion of contaminated food or water and transmitted through any of the transmission mechanisms. Individuals who get infected should fully recover after at least two weeks. A change in the values of the parameters (“RiskOfChildrenGettingInfected”, “RiskOfFliesGettingInfected” and “RiskOfHH”) should have an influence on the total number infection cases.

Enabling agents with daily activities

The daily activities implemented are fetching water, dumping waste, playing at dumpsite, cooking, eating and drinking.

Fetching water

It is important to inspect the agents fetching water, their ages and time step at which they go to the river. I extracted an output file of the individual’s IDs, their ages and the time step

Table 6.1: Agents fetching water

PersonID	Age	Tick
460	33	80
846	37	80
137	34	80
975	37	80
1387	25	80
3334	25	80
409	18	80
49	41	80
23	30	80

at which they go to get water (as shown in Table 6.1) below. It shows a small portion of the output showing the personID, age and timestep.

Such an output helps to identify whether the right people (age) are going to fetch water at the right time (time step). Fig 6.1 is a screen shot from NetLogo showing people with fetching water in the least cycle of three days.

Figure 6.1: People fetching water in a cycle of 3 days

On day 1, a fewer number of people are fetching water as compared to day 2 and on the third day the almost all the river nodes are occupied by people fetching water. Only people with low income fetch water on day 1, on day 2 people with both low income and average income fetch water whereas on the third day people aged 15 and above collect water at the river.

Dumping Waste

Only children between 5 and 12 years go to the nearest dumpsite to dump waste. Occasionally, children from low income families play at the dumpsites where they pick up plastics and empty cans to use when playing children games.

Table 6.2: Agents dumping waste

PersonID	Age	Tick
2699	8	31
1683	10	31
41	10	31
3101	9	31
2098	8	31
434	9	31
1000	6	31
81	8	31
1265	9	31

An output presented in Table 6.2 was generated to identify whether the right children (age) are going to fetch water at the right time (time step).

Table 6.3 shows a list of children who played at a dumpsite, the implementation is correct since only children from low income level families play at dumpsites.

Table 6.3: Children playing at dumpsites

PersonID	Tick	Age	Income level
1212	51	10	low
3405	51	8	low
528	51	8	low
3396	51	7	low
631	51	9	low
1299	51	9	low
257	51	11	low
1485	51	7	low
760	51	11	low

Cooking food

Food is prepared by a household at the every second time step of each day and is served (a plate for each individual). Fig 6.2 shows food that is prepared which was assigned green color and it is at the locations of the individuals (agents).

At time step 3 of each day, the food disappears only if the individual at that location eats it.

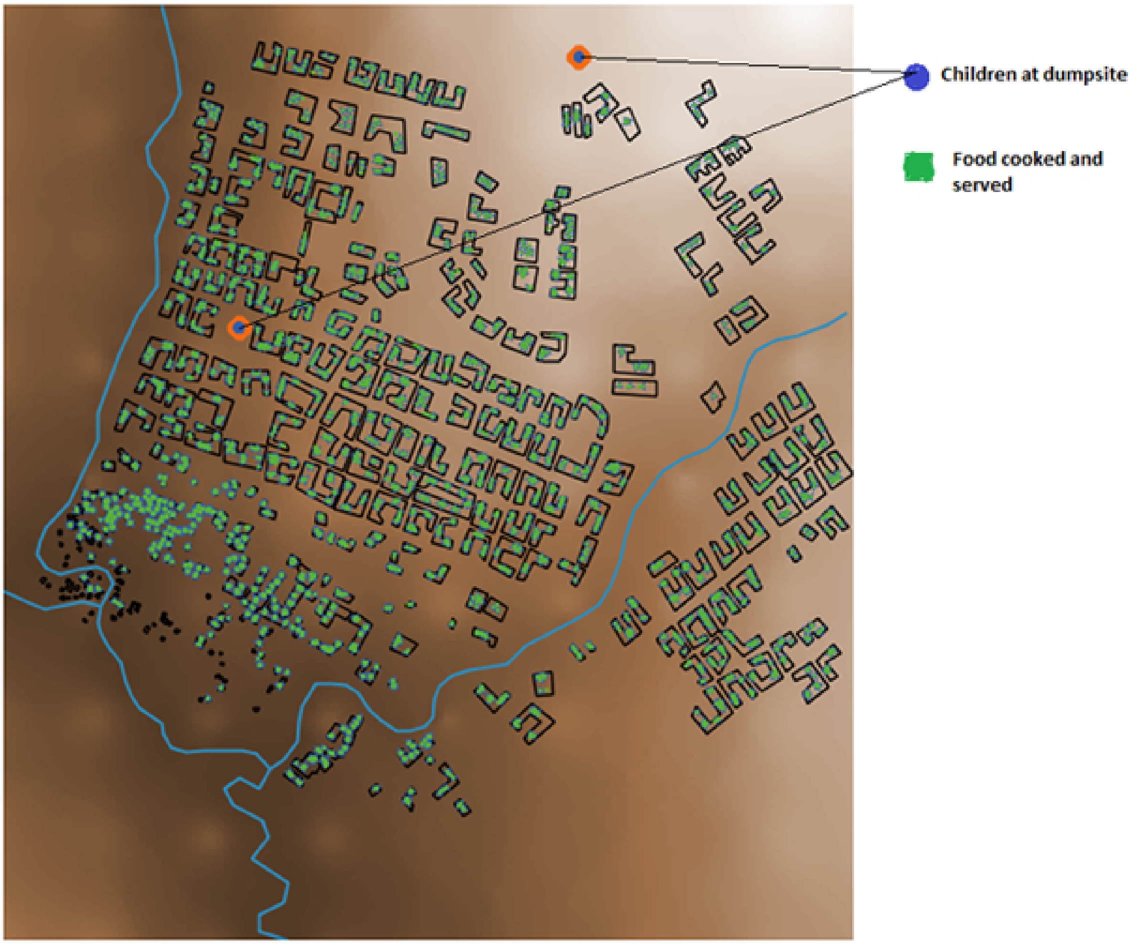


Figure 6.2: Children at dumpsite and food that is served

Transmission mechanisms

EH is the primary mechanism of cholera transmission, which will initiate the cholera outbreak. Transmission starts from the low income families, as they dump their (infected) waste at a dumpsite, it becomes infected, the houseflies which breed at dumpsites become infected hence VT and HEH transmission start. As the rains pours, some cholera bacteria is carried by runoff to the water bodies. HEH transmission from drinking water contaminated from infection from human faecal wastes will take place. HH occurs whenever there is/are agent(s) suffering from the infection.

Fig 6.3 shows the line graphs of the number of individuals infected throughout the simulation period. EH transmission (primary transmission) first causing the secondary transmission mechanisms, and the delay period of the VT and HEH is coming out as presented.

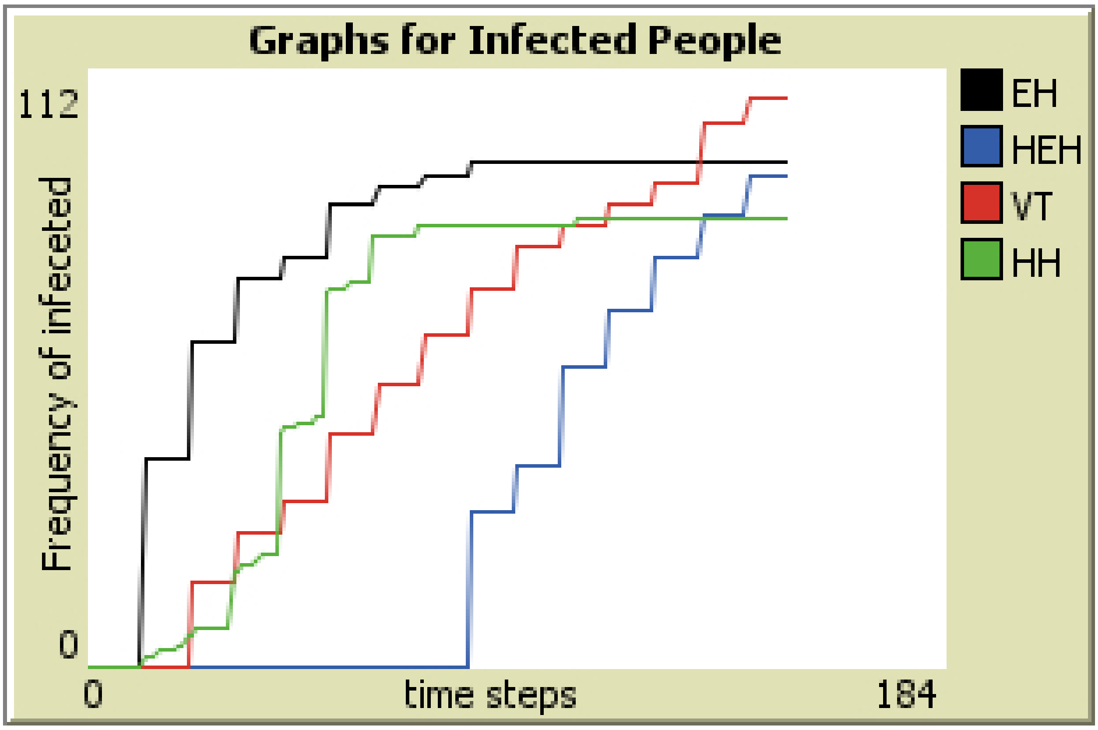


Figure 6.3: Line graphs for each transmission mechanism

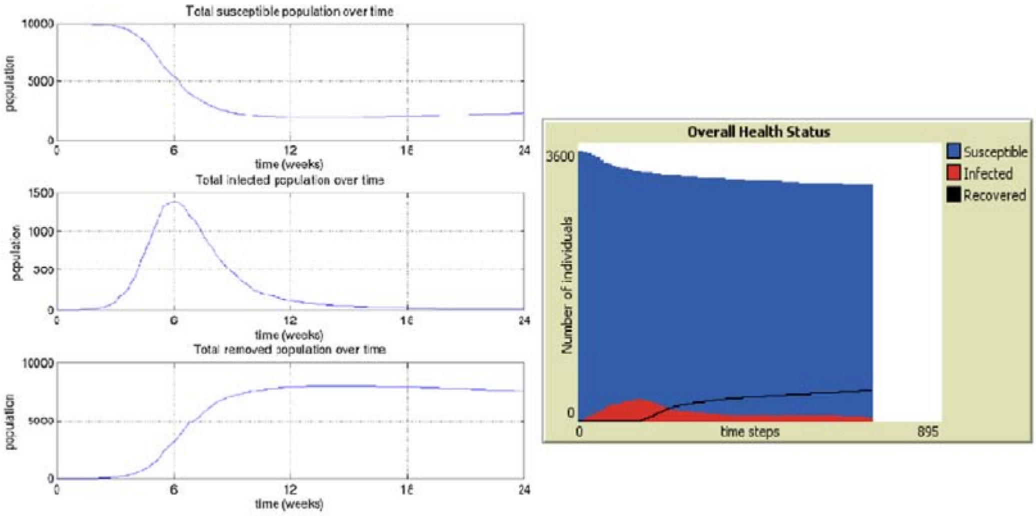
Comparison of the "Overall Health Status" graph with SIR Epidemic Model

The Susceptible-Infected-Recovered (SIR) model (from literature) is a simplified (general) dynamic model for many infectious diseases and has been used by epidemiologists to describe how the numbers of people in different compartments vary with time when an outbreak occurs. The total number of people who are susceptible decreases as people get infected. The total number of individuals infected reduces as soon as some of the infected people start to recover.

As the simulation runs, the population in the three "SIR" compartments are plotted at each time step on a graph titled "Overall Health Status". In Fig 6.4 the line graphs on the left, shows how total numbers of people belonging to three SIR compartments vary with time and they are a result (output) of an age structured SIR model developed by Agheksanterian Gobbert [3]. For a mild cholera outbreak, their simulation resulted in approximately 75% of the initial susceptible population contracting cholera. In the same figure, the graph on the right side shows how the total numbers of people vary with time as the simulation runs.

The numbers of infected cases increase and reach a peak value, the number of susceptibles start decreasing as the infection starts and people start recovering fully after 14 days (2 weeks).

This helps to verify whether the system works as expected. For both cases, the recovered become immune and they are not susceptible to the infection till the end of the simulation. The population goes through the transition stages of the SIR model. The total number of population does not change and not all people get the infection by the end of the simulation.



Source: agheksanterian2007modeling

Figure 6.4: Health Status

6.1.2 Determine the effect of varying the parameters and testing whether the sliders work

Expectation of the prototype was that the lower the parameter, the lower the number of infections for its respective transmission mechanism.

6 sets of test runs (each set is an average of 10 simulation runs) with randomly chosen (but different) parameters were performed in order to observe if there is an effect on the total number of infection cases from varying parameters.

Results

Table 6.4: Effect of changing parameters

SET	Risk of children getting infected	Risk of flies getting infected	Risk of HH	EH	HEH	VT	HH	TOTAL
1	50%	50%	10%	100	189	83	16	387
2	35%	35%	15%	96	188	77	16	376
3	25%	25%	10%	88	175	67	15	343
4	15%	15%	5%	89	170	56	11	325
5	10%	10%	5%	95	164	48	6	312
6	0%	0%	0%	85	100	0	0	185

The average values were plotted in a graph for easy comparison.

HEH
As the value of Risk of children getting infection from dumpsite parameter decreases, average number of HEH-infection cases also decrease. The lower the parameter, the lower the infection cases.

VT
As the value of Risk of flies getting infection from an infected dumpsite decreases, the average number of VT-infections decrease as well. The lower the parameter, the lower the

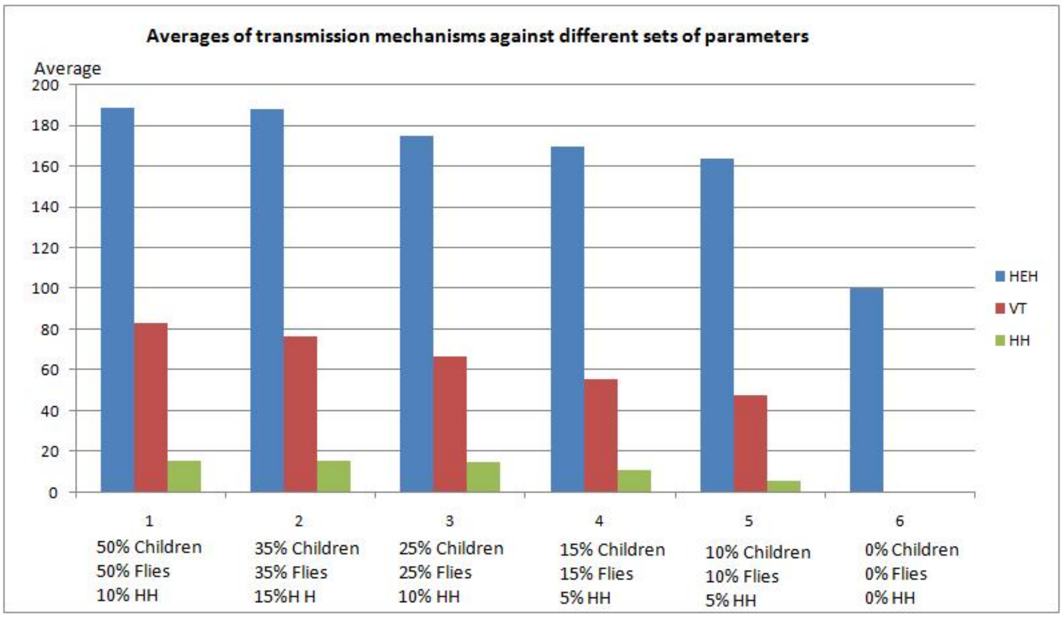


Figure 6.5: Averages against different sets of parameters

infection cases.

HH

As the value of Risk of HH infection from infected people decreases, the average number of HH-infections decrease but not in all cases.

Zero values on sliders

As the sliders were set at zero, there were no (zero) HH infection cases and VT infection cases recorded, but for HEH, the was no infection of children recorded but drinking water that was contaminated by run-off recorded on the HEH monitor.

Conclusion

In conclusion, the sliders are working as expected and effect of varying the values of parameters on respective transmission mechanism is positive, the higher the value for parameter the more infection cases are recorded from the respective transmission mechanisms.

6.2 RESULTS AND DISCUSSION

6.2.1 Determining number of simulation runs

5 simulation runs were performed to determine whether the same output is found and the results are presented in Table 6.4

The results in Table 6.4 were plotted into a line graph for easy interpretation.

Considering the output of the 5 independent runs (Table 6.5 and Fig 6.5), there is a lot of variation in the totals and EH infections. HEH infection cases vary but less than the EH cases. The variation in VT and HH can be considered since it is not much as compared to

Table 6.5: Individual runs

RUN	EH	HEH	VT	HH	TOTAL
1	192	217	74	19	502
2	82	160	70	16	328
3	54	183	83	16	336
4	113	216	73	15	417
5	29	172	84	14	299

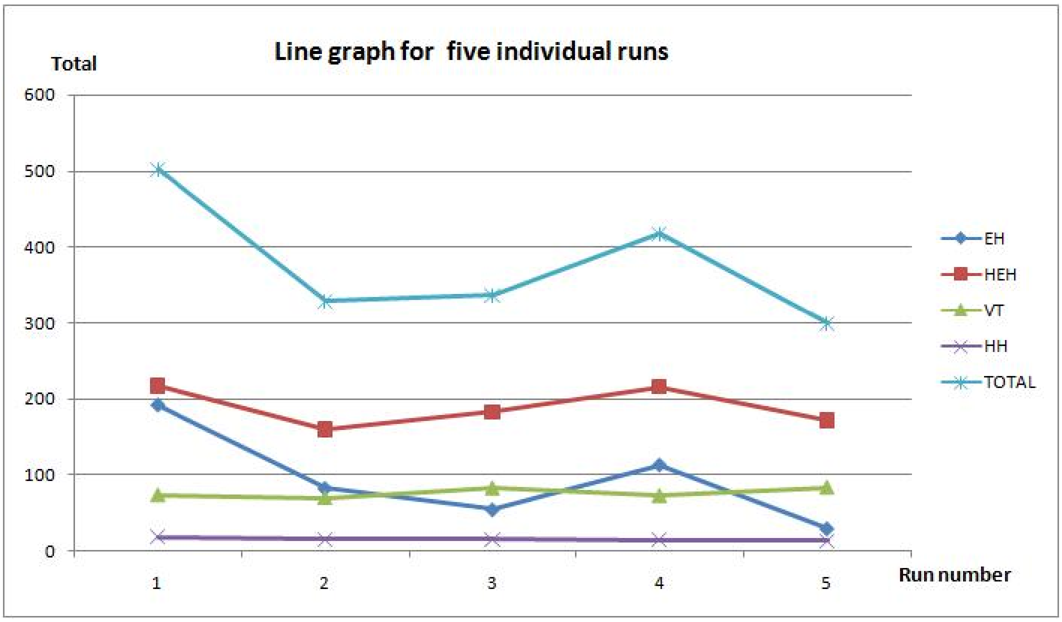


Figure 6.6: Plot of individual runs

the other 2 mechanisms. However, individual runs are not the best to consider therefore more runs have to be made.

How many times should the simulation be run to generate stable and accurate results?

This cholera simulation is non-deterministic (stochastic) model due to randomness effect mainly from EH transmission mechanism in the model, hence it is crucial to perform repeat runs to determine the number of runs that are necessary to extract stable and valid results for analysis. The set of runs performed for this section are 720 ticks long each (72 days in total) and are using constant parameters of 10 ticks representing a day, 35% “RiskOfChildrenGettingInfected”, 35% “RiskOfFliesGettingInfected” and 15% “RiskOfHH”. In total, 50 runs were done and an average of first ten runs was calculated, then an average of the first 20 runs, then 30 runs, 40 runs and for all the 50 runs.

Results

The average values for the different number of runs per transmission mechanism and their totals are presented in Table 6.6.

The results in Table 6.6 were plotted into a line graph for easy interpretation.

Average values for EH, HEH, VT, HH and total for first 10 runs, 20 runs, 30 runs, 40 runs

Table 6.6: Average values for different number of runs

AVERAGE OF	EH	HEH	VT	HH	TOTAL
First 10	96	188	77	16	376
First 20	95	186	77	16	373
First 30	100	186	77	17	376
First 40	96	183	79	17	373
All 50	92	182	79	16	368

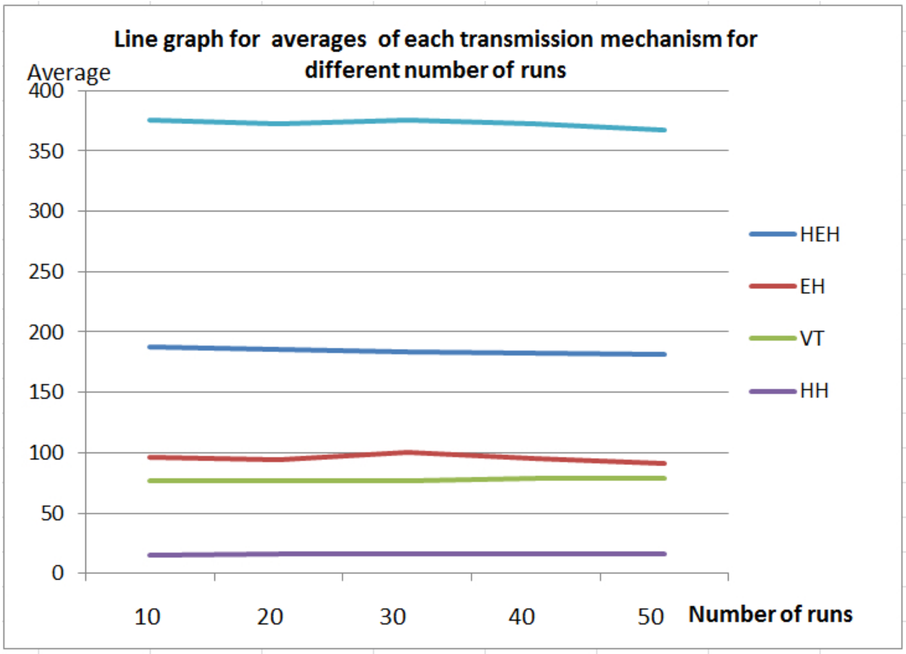


Figure 6.7: Averages of different number of runs

and all 50 runs were plotted into line graphs (Fig. 6.7). Averages of HEH, VT and HH are almost constant. The line series for EH has slight fluctuations which are accountable for the fluctuation in the line series of the averages of the totals.

Discussion and conclusion

The slight instability observed for the EH is caused by the randomness effect which comes during initialisation of the model, randomness of upto 50% of node/vertices can be contaminated. This stochastic effect can not be avoided, but in general the averages are relatively stable. The randomness in HEH main results from the randomness in the raindrops as they will be falling randomly over the whole area. This rainfall stochastic behaviour cannot be controlled.

Therefore 10 simulation runs shall be used to perform further results since the averages of the number of 10 or more runs are almost the same (stable).

6.2.2 Determination of parameters

Sensitivity analysis

Sensitivity analysis is a technique used to determine/investigate how a model responds to (changing) varying values of an independent parameter provided as input. The main purpose of sensitivity analysis is to identify critical/key parameters, which are the key drivers of a model's results. This analysis is also important to help understanding of the relationship between input and output variables and testing the robustness of the results.

Procedure

Parameter Risk of Children Getting Infected from the dumpsite is varied seven times (60%, 50%, 35%, 25%, 15%, 10% and 5%) whilst parameter Risk of Flies Getting Infected from the dumpsite was held constant at 35% and Risk of HH remains constant at 15%. Each (set) simulation run is an average of 10 runs.

Parameter Risk of Flies Getting Infected from the dumpsite is then varied seven times (60%, 50%, 35%, 25%, 15%, 10% and 5%) whilst parameter Risk of Children Getting Infected from the dumpsite was held constant at 35% and Risk of HH remains constant at 15%. Each simulation run is an average of 10 individual runs due to the stochastic nature of the prototype.

Parameter Risk of HH is varied as well while parameters Risk of Flies Getting Infected from the dumpsite and Risk of Children Getting Infected from the dumpsite were held constant at 35%.

Results

(a) **Varying parameter Risk of Children Getting Infected from the dumpsite**

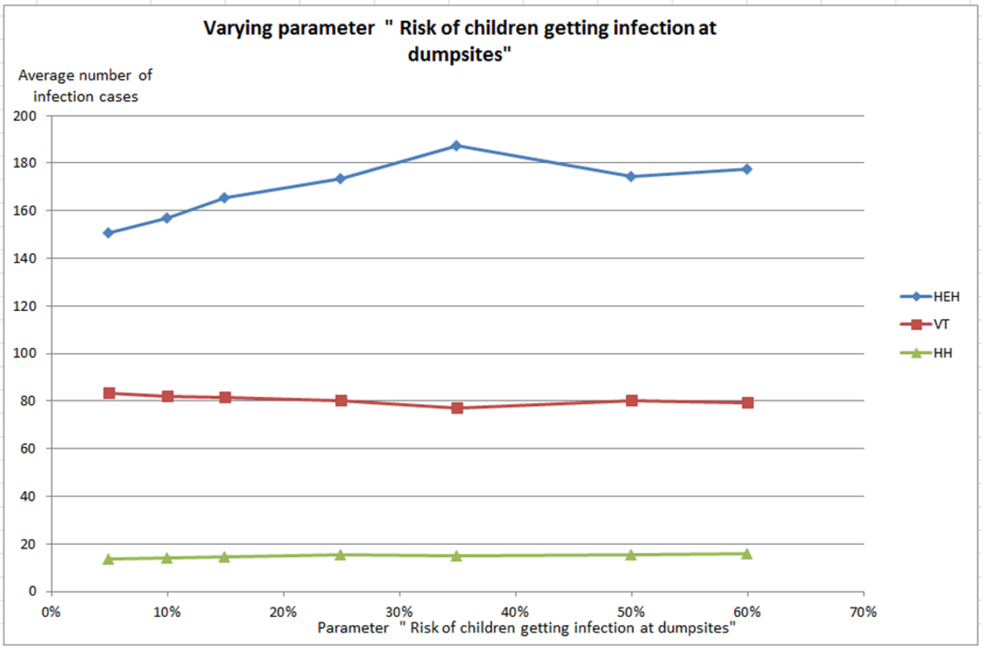


Figure 6.8: Effects of varying Risk of children getting infection parameter

As the parameter Risk of Children Getting Infected from the dumpsite is varied (from 5% to 60%), very small variations are observed on the averages of infection cases of VT and HH (Fig. 6.8). There is no direct relationship between HEH infection cases with either VT infection cases or HH infection cases. EH infections were not considered since EH is the primary transmission mechanism hence it has no effect from its secondary transmission mechanism. Changes in the parameter value only affects its respective transmission mechanism HEH.

As parameter Risk of Children Getting Infected from the dumpsite is varied from 5% to 60%, generally there is an increase in the number of HEH infections but there is a peak at 35%. Considering the averages of HEH for many runs (Fig 6.7), the averages are all above 180. Hence cannot justify that it is an outlier.

Further analysis was done to determine the values for the parameter that can be chosen for the simulation model. The starting point is 25% since it is at the middle of the set of runs performed already. With reference to the last 2 columns of Table 6.7:

Table 6.7: Sensitivity of HEH

Parameter	EH	HEH	VT	HH	TOT	TOT. Diff	% Diff.	HEH-Diff	%HEH Diff
60%	95.4	177.6	79.4	15.8	368.2	-10.8	-2.9	4	2.3
50%	91	174.4	80.1	15.5	361	-18	-4.7	0.8	0.5
35%	95.9	187.2	77	15.2	375	-4	-1.1	13.6	7.8
25%	109.6	173.6	80.1	15.6	379	0	0	0	0
15%	86.2	165.6	81.5	14.6	348	-31	-8.2	-8	-4.6
10%	74.8	157	81.9	14	328	-51	-13.5	-16.6	-9.6
5%	66.8	150.6	83.4	13.7	315	-64	-16.9	-23	-13.2

- At 35%, marks the peak of effect of increasing the values of the parameter, and from 25% there is an increase which contributes to the HEH infection.
- Going down from 25% the model is sensitive to the decrease in the value of the parameter.

Conclusion

The parameter values ranging from 5% (as the least value) to 35% (as the highest value) shall be considered. The lower values at 50% and 60% can be caused by natural variation.

(b) **Varying parameter Risk of Flies Getting Infected from the dumpsite**

There is no effect of changing the parameter Risk of Flies getting infection at the dumpsite on HEH and HH infection numbers. It only affects its respective transmission mechanism VT.

There is a sharp increase of VT infections as the parameter is varied from 5% to 25%. From 25% to 60%, there is a slow increase in the total numbers of VT infections.

Further analysis was done to determine the values for the parameter that can be chosen for the simulation model. The starting point is 25% since it is at the middle of the set of runs performed already.

With reference to the last 2 columns of the Table 6.8:

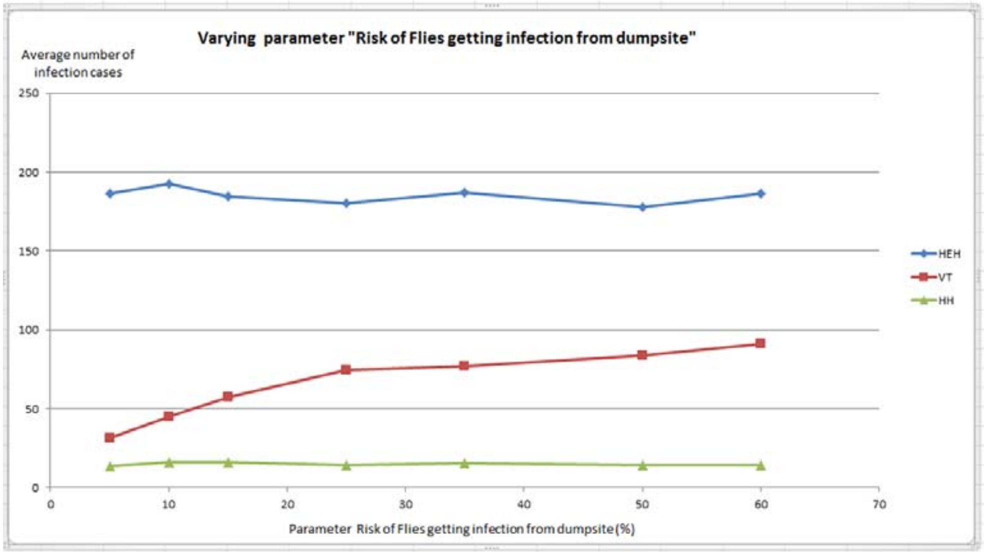


Figure 6.9: Effects of varying Risk of Flies infection parameter

Table 6.8: Sensitivity of VT

Parameter	EH	HEH	VT	HH	TOT	TOT. Diff	% Diff.	VT-Diff	%VT Diff
60%	98.1	186.2	91.3	14.4	390	63.2	19.4	16.9	22.7
50%	101.1	177.9	84	14.5	377.5	50.7	15.5	9.6	12.9
35%	95.9	18.2	77	15.2	375.3	48.5	14.9	2.6	3.5
25%	57.8	180.4	74.4	14.2	326.8	0	0	0	0
15%	12.2	184.5	57.2	15.9	369.8	43	13.2	-17.2	-23.1
10%	125.9	192.3	44.8	16.2	379.2	52.4	16	-29.6	-39.8
5%	70.7	186.2	31.2	13.5	301.6	-25.2	-7.7	-43.2	-58.1

- Starting from 25% moving upwards, there is a gain in the values.
- Moving downwards from 25%, there is a loss in the total values.

Conclusion

The parameter can have any value ranging from 5% to 60%.

(c) **Varying parameter Risk of human-to-human transmission**

As presented in the graph (Fig 6.10), varying parameter for Risk of HH transmission does not have any effect on both VT and HEH, since HH is secondary transmission to both VT and HEH. It only affects its respective transmission mechanism HH.

The main reason for the decrease in the average of HH infection is due to the EH. Despite a high value of parameter Risk of HH, if EH is having lower infection numbers, then HH will also have lower infection cases.

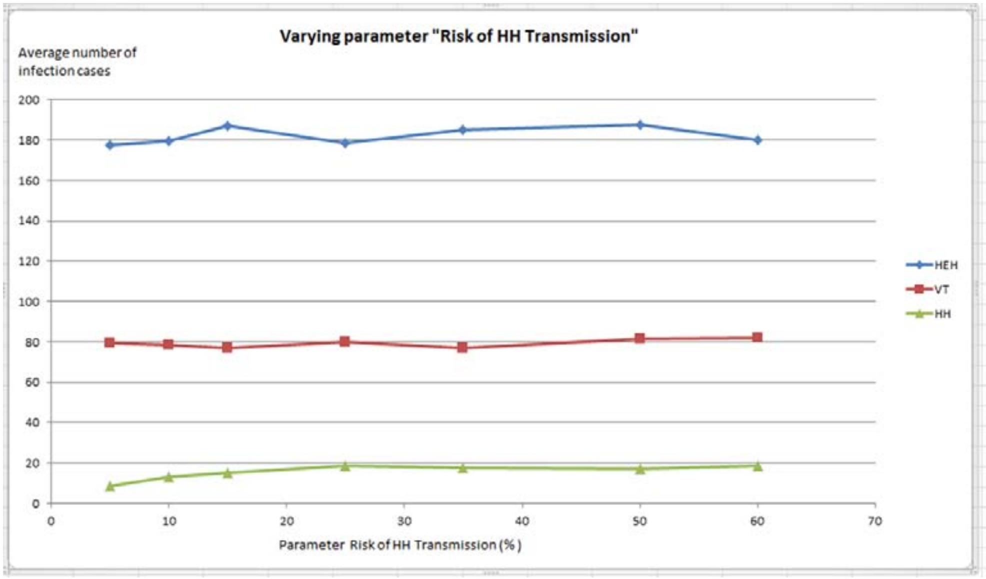


Figure 6.10: Effects of varying Risk of HH parameter

2 scatter plots were were plotted for the corresponding (values) EH infections and HH infections Fig 6.11.

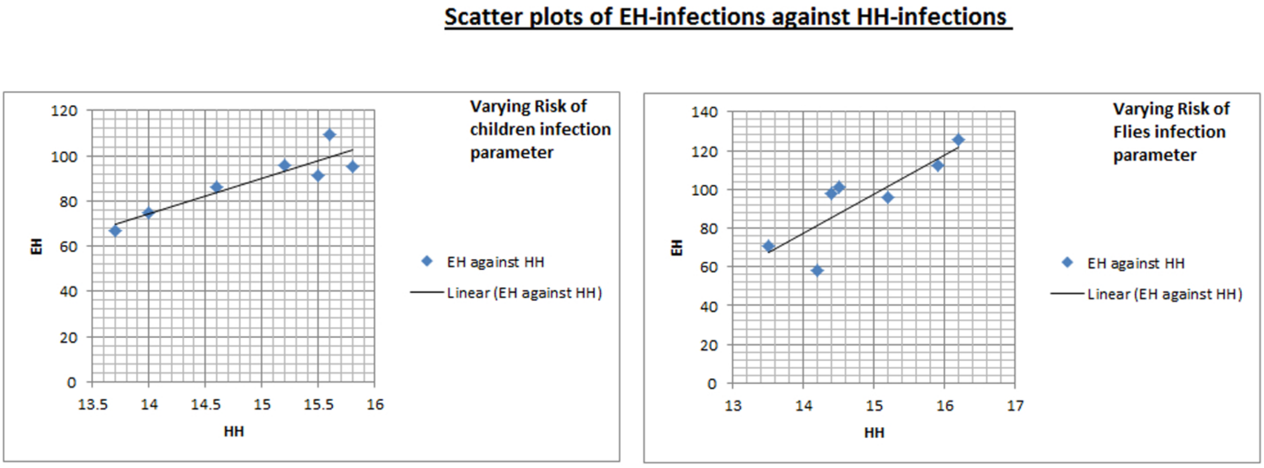


Figure 6.11: Scatter Plots for EH against HH

Most EH infections occur to people who have low income level, and HH transmission is the most immediate secondary transmission that takes place due to lack of money to buy disinfectants such as soap. As they will be assisting the ill, they do not put on gloves since they can not afford.

There is a direct relationship between primary cases due to EH and the secondary cases for HH. Therefore, the more the EH infections, the more the HH infection cases.

As the parameter Risk of HH was was being varied, and making 25% as the starting value

(Table 6.9, the last 2 columns). As the parameter value decrease there is greater loss in the number of infections especially at 5% almost 56% cases will not be recorded. Going up from 25% to 35% there actually a loss of which a gain was expected. This shows that 25% is the maximum value for the parameter that can be considered, used in the simulation prototype.

Table 6.9: Sensitivity of HH

Parameter	EH	HEH	VT	HH	TOT	TOT. Diff	% Diff.	HH-Diff	%HH Diff
60%	81.9	180.3	82	18.8	363	-17.5	-4.6	0	0
50%	60.4	187.5	81.5	17	346.4	-34.1	-9	-1.8	-9.6
35%	90.6	185	76.7	17.6	369.9	-10.6	-2.8	-1.2	-6.4
25%	103	178.6	80.1	18.8	380.5	0	0	0	0
15%	95.9	187.2	77	15.2	375.3	-5.2	-1.4	-3.6	-19.2
10%	116.3	179.4	78.6	12.9	387.2	6.7	1.8	-5.9	-31.4
5%	94.3	177.3	79.3	8.3	359.2	-21.3	-5.6	-10.5	-55.9

Conclusion

25% is the optimal value for parameter Risk of HH, that shall be used.

Determining the effect of combining the maximum and minimum values of parameters (“Risk of children getting infected” 5% and 35% and “Risk of flies being carrying infection” 5% and 60%. “Risk of HH is constant at 25%

A combination of the parameters were used for further simulation runs and the total number of cases are presented in Table 6.10.

Table 6.10: Matrix

	5%Chidren	35%Children
5%Flies	300	323
60%Flies	318	380

The results of the set combinations were used to plot epidemic curves Fig. 6.11 in order to choose compare the structure of curves to epidemic curves from two cholera outbreaks in African countries.

Epidemic curve

Epidemic curve is a graphic depiction of the number of cases outbreak cases by date of illness onset. New infection cases per day are plotted against an interval of time to describe a specific epidemic or outbreak. The graph can either be a histogram or a line curve. Epidemic curves assist to determine the transmission trends/patterns and identifying the source of the outbreak and/or transmission method.

Four epidemic graphs were plotted corresponding to the following set of combination of parameters presented in Table 6.10:

The epidemic curves (Fig 6.11) show continuous transmission over the simulated period of 72 days. Epidemic curve 1, most people got infected on the 10th day after the simulation started with 23 infection cases. Epidemic curve 2, a peak was reached on the 12th day of the onset of the simulation having 34 cases recorded. Epidemic curve 3, a peak was reached on the 27th day with a recording of 17 cases. Epidemic curve 4, 28 infection cases occurred on the 12th day of the simulation.

These curves (Fig 6.12) were compared to epidemic curves (Fig 6.13) for 2005 Kakuma cholera outbreak at a refugee camp in Kenya and cholera outbreak cases reported at Taung clinic in South

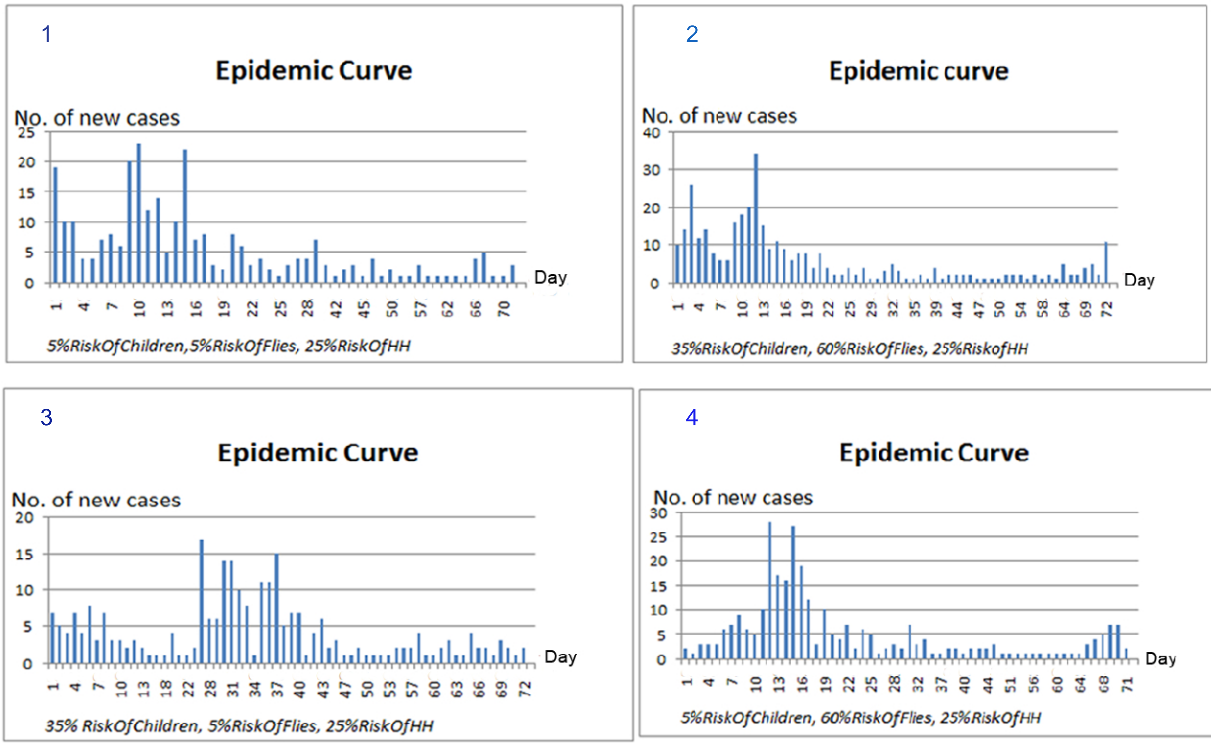


Figure 6.12: Epidemic Curves

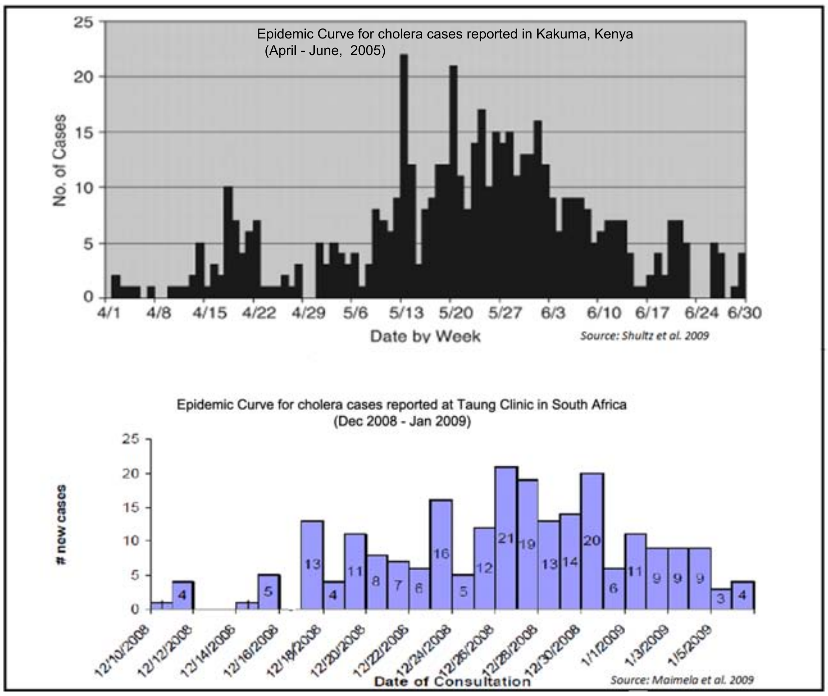


Figure 6.13: Epidemic curves for cholera outbreaks in African countries

Africa (December 2008 to January 2009). For Kakuma epidemic curve, a peak was reached around the 43th day after the start of the outbreak and almost 22 cases were reported. For the South African epidemic, 21 infection cases were recorded on the 18th day after the start of the outbreak. The peak of the simulated outbreak cases are coming sooner (earlier) as compared to the epidemic curves for Kakuma outbreak in Kenya and for Taung clinic in Sekhukhune District in South Africa, epidemic curves 1, 2 and 4, their peaks are coming sooner, and epidemic curve 3, the peak comes later.

Therefore, the simulation is producing earlier peaks but correctly presenting continuous transmission of cholera for a prolonged period of time. One possible reason for the early peak in the simulated results is caused by the threshold value of 10 implemented for the dumpsite infection level which marks the start of infectious status of a dumpsite. A recommendation is to increase the value of dumpsite infection level. A reason to explain the prolonged transmission is that the agents (individuals) are not changing their behaviour and this is a recommendation proposed. For all the epidemic curves (simulated data and real data), they are not smooth and each day comes with not a predictable outcome.

Comparison of model's simulated results to literature simulated results

Agheksanterian and Gobbert [3] developed an SIR age-structured model to model the spread of cholera epidemic which can be applied to environments such as refugee camps and total of 10 000 susceptible population, almost 75% in the simulation got infected after 24 weeks from the onset of the outbreak.

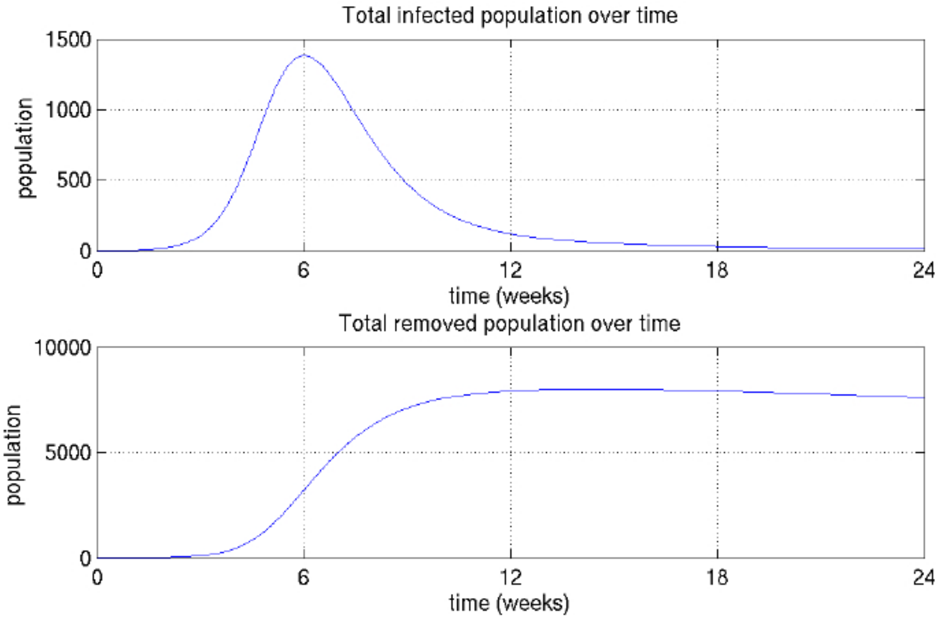


Figure 6.14: Simulated results of SIR age structured model

Fig 6.14 shows how the total population for infected and recovered varies with time from Agheksanterian and Gobbert [3]. A peak of infection cases occurred at 6 six weeks after the onset of the epidemic in a low rate shedding of cholera. The recovery

Graph showing how the total population for infected and recovered varies with time from the simulated results

Both plots Fig 6.14 and Fig 6.15 show a recovery period of almost 2 weeks (14 days) after the first record of infection case. Both graphs have long tails showing that cholera epidemics have a

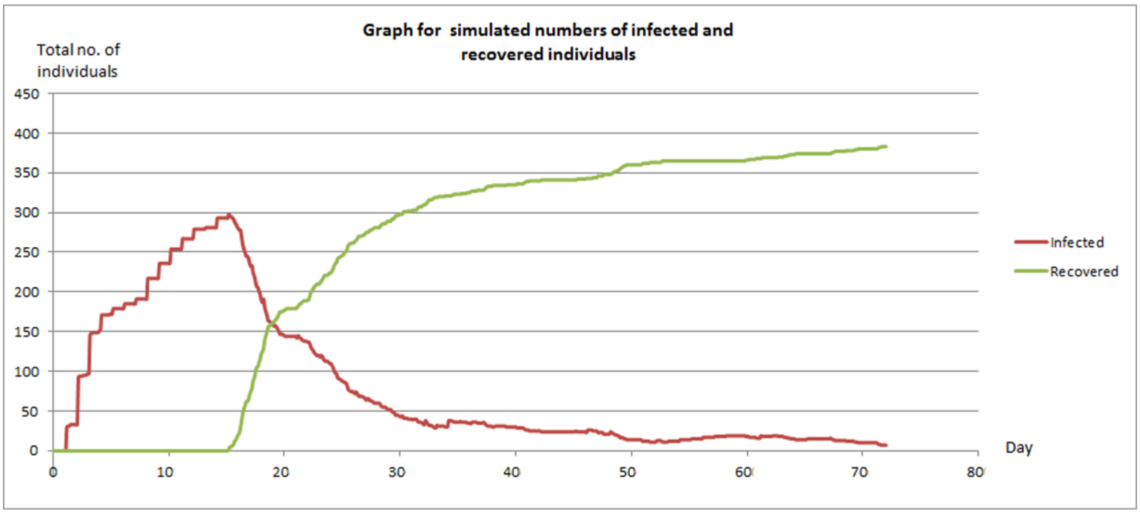


Figure 6.15: Simulated results of Agent-based simulation prototype

long propagation period. For the total numbers of infection cases, simulated results from the SIR age structured model is over estimating since it is only considering the varying ages where the simulation prototype is considering heterogeneity in the initial dataset of individuals, that is the age differences, income levels and hygiene standards.

Comparison of total number of cases reported

From the 2005 cholera report for Kumasi cholera outbreak, a total of 949 cases were reported in a total population of 1 000 345. From the simulation, an average of 330 infection cases were simulated in total population of 3 501. Therefore, the model is over estimating hence adjustments need to be made to the model. I recommend that individuals should have their own (personal) hygiene levels and more personal characteristics instead of generalising as a household attribute. Heterogeneity within households need to be considered. To the modelling part, the chances of individuals contracting cholera need to be lowered, they are most probably too high

6.2.3 Spatial distribution of infection cases

Results from a single simulation run were used to show the spatial distribution of the infection cases and the map produced is presented as Fig 6.13.

Only individuals who got infected are presented in the map and are distinguished from each other by the source of infection that is the transmission mechanism from which they got the infection from. Almost all EH infections occurred in the area with households without access to tape water. Majority of the households living in that area have low income, and this justifies why they got infected through EH transmission.

A greater number of the infected people have HEH transmission as their source. Most of the HEH infections occurred in the area where households with average incomes reside. The HEH infections that occurred in areas where low income families stay, are for children who play at dumpsites that are infected.

For the VT infections, the spatial locations where they took place correspond correctly to areas where high numbers of houseflies were distributed in the density distribution map (Fig 4.9).

HH infections occurred in area where households with low incomes reside and this corresponds to what literature states that poverty is a contributing factor to cholera transmission.

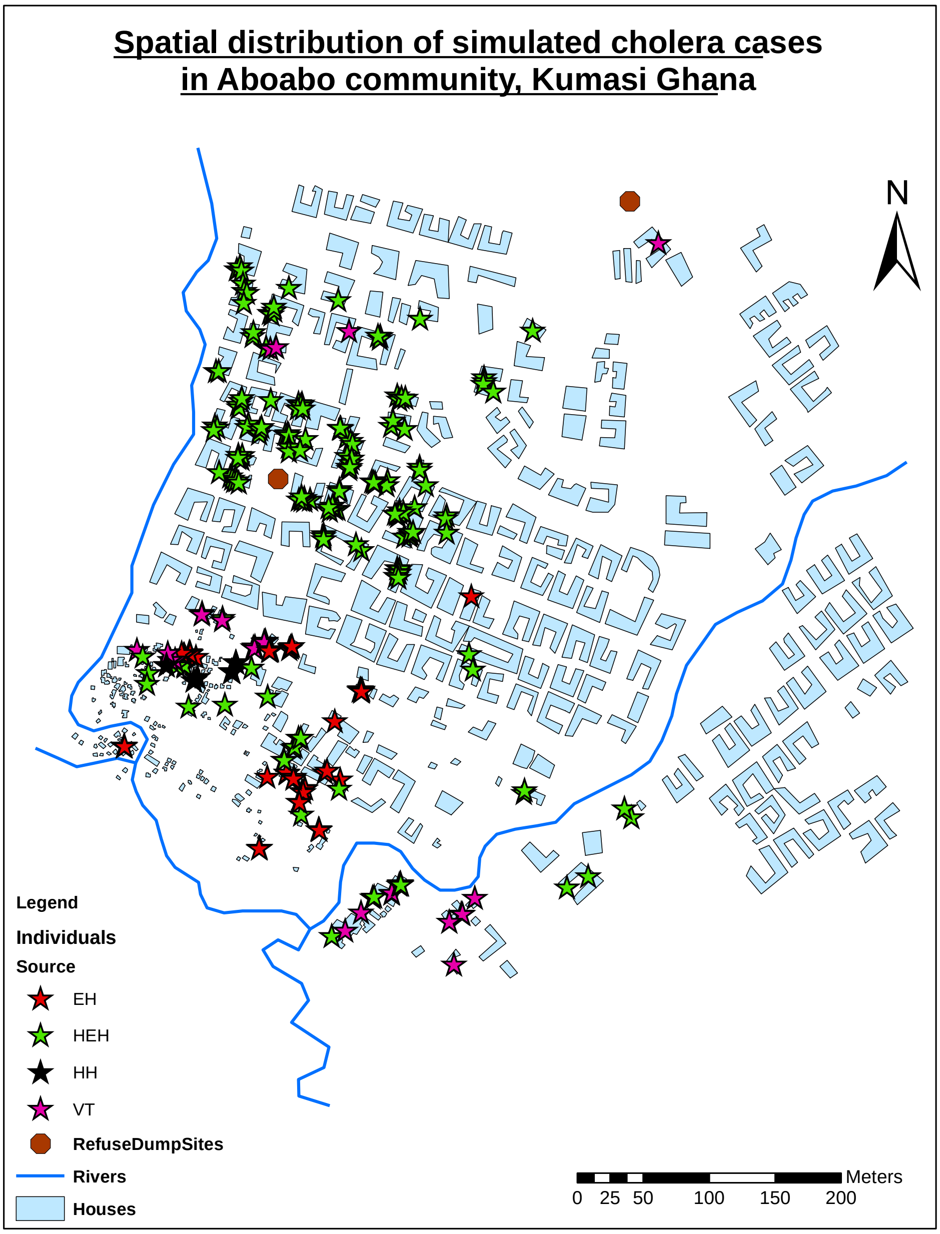


Figure 6.16: Spatial distribution of simulated cholera cases

There are no infection cases that occurred around the area surrounding the right branch of the river, this is because run-off was flowing towards the left branch of the river only hence no contamination from run-off from dumpsite reach there.

6.2.4 Analysing cholera infections by hygiene level, blood type and income level

1. The total number of simulated infection cases by household hygiene level were plotted into a bar graph for analysis Fig 6.17.

   232 infected cases occurred to individuals with low hygiene, whereas 29 were average hygiene and only 2 with high level of hygiene got infected. This proves that hygiene standard

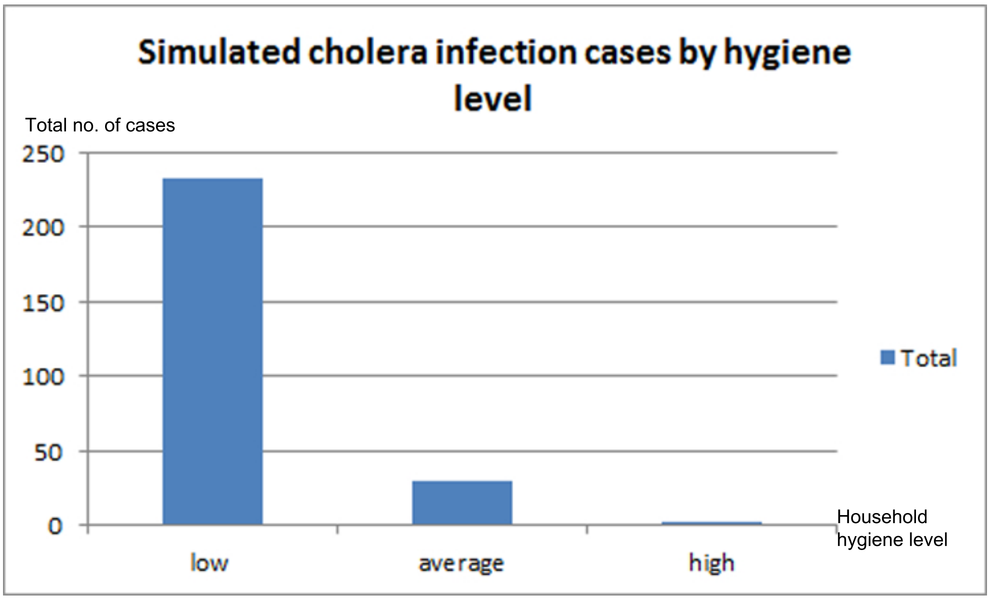


Figure 6.17: Hygiene levels of simulated cholera cases

is an important factor that influences the transmission of cholera as stated in literature. High standard hygiene mean that the people boil or treat water before drinking and cover their food after preparation. Therefore, the higher the level of hygiene, the lower the risk of contracting cholera.

2. The total number of simulated infected cases by individual blood type (group) were plotted into a bar graph for analysis Fig 6.18.

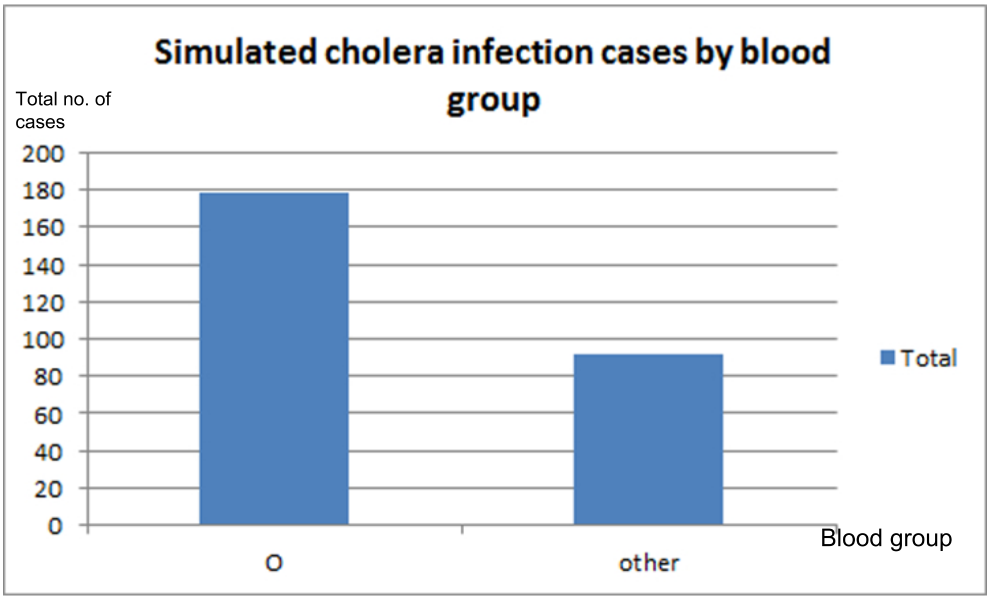


Figure 6.18: Blood groups of simulated cholera cases

179 infected cases occurred to agents with blood group ‘O’, and 92 individuals with other blood groups. This proves that blood group is a factor that influence the transmission of

cholera. According to Swerdlow et al. [76], people in with blood group 'O' in South America were estimated to be more prevalent to cholera than other blood groups.

3. The total number of simulated infected cases by household income levels were plotted into a bar graph for analysis Fig 6.19.

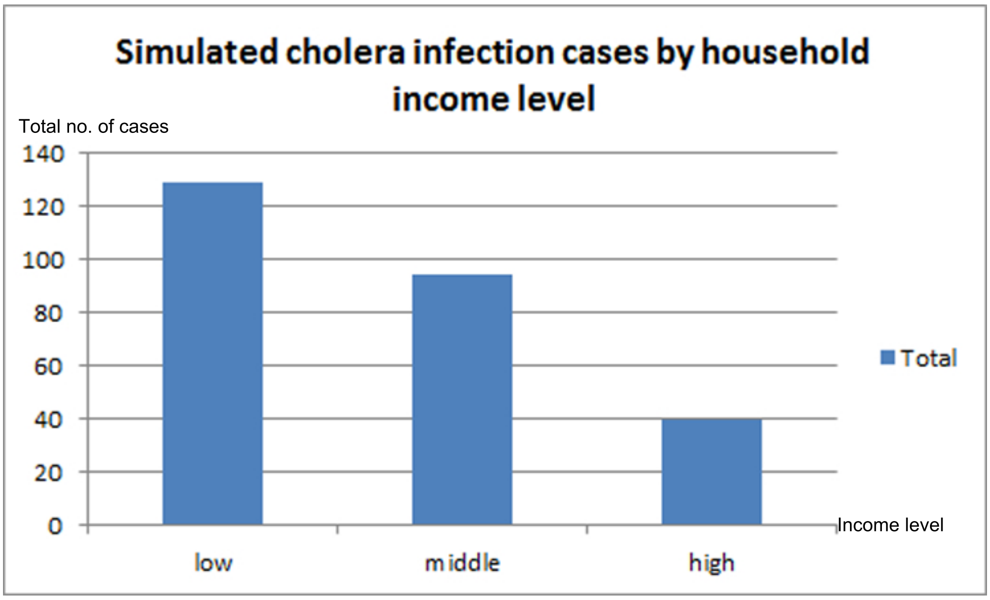


Figure 6.19: Income levels of simulated cholera cases

129 individuals who belong to households with low income level were infected by cholera, whereas 94 belonging to households with middle income and 40 from high income families were infected during the same simulation run. Income level relates to poverty, cholera infection is influenced by poverty, the lower the income the higher the chance of contracting cholera due to lack of money to buy disinfectants such as soap.

6.2.5 Reproductive rate (R_o)

Basic reproduction rate (ratio/number) is the number of secondary infection cases expected to arise from the introduction of a primary infection case into a completely susceptible population [82]. R_o is defined as the actual average number of secondary cases per primary case [82]. R_o for cholera in a community is defined by the product of social and environmental factors. The importance of the aquatic reservoir depends on the sanitary conditions of the community [21].

Estimating a numerical value for any disease is a difficult task, for cholera in endemic areas, Bangladesh had a estimation of 12 for 1997 to 2001 [41]. For this simulation model, all initial cases from EH are leading to the secondary cases HEH, VT and HH infection cases. Therefore, the assumption is R_o for the simulation is greater than 1, hence the outbreak occurs until there are no more people who are susceptible.

Chapter 7

Conclusion, limitations and recommendations

7.1 CONCLUSION

The objective of this research was to develop an Agent-based modeling prototype that simulates the diffusion of cholera in Kumasi, Ghana and evaluate the contributions made from each of the four transmission mechanism responsible for spread of cholera. This research resulted in the development of a spatially explicit prototype simulation model in which four different types of cholera transmission mechanisms were integrated for the study area of Kumasi in Ghana. The prototype contained two branches from a river, two dumpsites and a settlement containing different types of income groups. The contributions made by each transmission mechanism were evaluated and found out that human-to-environment-to-human transmittance plays a major role in the diffusion of cholera during an epidemic. Human-to-human transmission has the lowest contribution and these findings confirm what is found in literature. People with blood type 'O' are at high risk of contracting cholera as compared to others and low hygiene standards, poor access to safe drinking water, heavy rainfall and low income levels are factors which influence the spread of cholera.

The output (results) of the model have been compared to epidemic data from two outbreaks in African countries, to the aggregated statistics of 2005 Kumasi cholera outbreak and results of a technical report by Agheksanterian and Gobbert [3] on SIR age structured model that simulates the spread of cholera. The epidemic curves produced from the simulation results and the other sources, show similarities that cholera epidemics generally have long and continuous propagation period. The epidemic curves from simulated results have multiple peaks similar to the peaks from data of outbreaks in African countries. The infected-recovered models/plots showed almost the same recovery period of 2 weeks (14 days).

During the development process it became evident that there are still many unknown elements in all of the transmission mechanisms that make it difficult to validate/develop this model.

Flies

There was no information about the numbers of flies and their distribution (no measurements from the field), theoretically little known. This led to a lot of assumptions in the generation of the housefly layer (distribution of the flies) and in the chance of the flies infecting the food during/after preparation, and therefore the risk of people contracting this disease.

Hydrological system

Although there is far more information on hydrological systems in literature, compared to the flies, the hydrological system is very complex and yet at the same time very important to the spread of cholera as it is mainly a water born disease. Even in the small area, there are factors that need to be considered which include a river system, run-off from the dumpsites, infiltration rate, evaporation, the soil type and a link to rainfall data that make the hydrological model very complex.

Survival of the V. cholerae bacteria in the environment

Although the bacteria are known to survive in water, food and soil, it is unclear where in the river water and run-off it will actually survive and for how long.

7.2 LIMITATIONS

Data

Main limitation of this research was the unavailability of good datasets. The shapefiles of the river, open space refuse dumpsites and communities had no metadata, hence were assigned a datum that might not be the one used when they were created. There was no layer for buildings and it was created from a google earth image which was very haze hence difficult to see to delineations of the buildings. And also unavailability of knowledge on the purpose of buildings, whether they are houses or for other uses.

A very good dataset of an epidemic outbreak of Cholera would be very helpful in further developing this prototype. The DEM that was collected had too big pixel sizes for such a small area.

The synthetic population was created only from age structures provided in a census report and there was no information about blood groups of the people in Ghana

NetLogo

The turtle breed of individuals was created from a point shapefile that was loaded, it is a limitation of NetLogo that it is very difficult to create individual with many attributes/characteristics with varying age, income level, and hygiene level yet at the same time having them subscribed to families. It would be useful to vary the population size and their characteristics.

Modelling

Due to lack of information and time constraints, people with high income were generalized to be fetching water at the river, the access to other water sources such as buying bottled water was not implemented (but has been included in the conceptual design), and hence the heterogeneity amongst households with high income was violated.

There are other activities that expose people to cholera infection such as eating fish carrying the bacteria, washing utensils with contaminated water, using contaminated water to bath and buying food from market places that were not modeled. Children do not go to school and adults do not go to work, but they could drink water at school/work that is infected. Including movement in space of people would assist in improving to the current model to better suit the reality.

Flooding of rivers to the shallow wells was not modelled as modelled in the conceptual design phase. This was mainly because of lack of information about the locations of the wells and the hydrological model need to be improved.

Important shortcoming of the current model is that people do not change behavior during a Cholera outbreak, they indeed are supposed to improve their awareness. I recommend that change of behavior be included in the model.

7.3 RECOMMENDATIONS

I think it is very important to develop an agent-based Cholera model because potentially this type of model is very useful for revealing additional risk factors.

After validation, a model can also be used in education, making people aware of the fact that changes in behavior will help to reduce the risk of contracting this disease. Interactions by the government by improving sanitation, providing safe drinking water, clean rivers, control housefly population and breeding areas, and moving dumpsites to suitable places (perhaps to locations where run-off will not reach) and contaminate the water sources.

Most crucial is that the hydrological model has to be improved (correct) since it plays a major role in the cholera diffusion. A more advanced model that includes the soil type, infiltration, evaporation and steepness of the area. As contaminated run-off gets to the river, the bacteria concentration in water should be considered since the bacteria do not stay on one location.

The houseflies distribution layer need to be improved by introducing heterogeneity to their behavior since in the current model, houseflies contained in a pixel are modeled as a single agent and the food that is present is the pixel will be contaminated. Of which in reality, not all flies will be carrying the bacteria and no infection may result

Experts in cholera epidemiology should be involved in the development process of the model but due to time limitation it was not done.

Individuals should have independent hygiene standards from the household hygiene. This helps to bring heterogeneity within families. The blood groups of individuals should be improved by taking into consideration that if parents are both of blood type ‘O’, it is not realistic to have child(ren) with any other blood group which is not ‘O’.

External factors such as migration into an area which may introduce the cholera into a new area. Therefore, information about the migration is necessary.

Appendix A

APPENDIX

```
;;procedures saved as .nls files
__includes [ "FetchingWater.nls" "DumpingWaste.nls" "PlayingAtDumpSite.nls"
"Cooking.nls" "Eating.nls" "HigherPopDensity.nls" "DrinkingWater.nls"
"Recover.nls" "HHTransmission.nls" "Dumping.nls" "InfectDumpsites.nls"
      "Reports.nls" "HEHTransmission.nls" "ChildrenEat.nls" ]
extensions [ gis ]
globals [ elevation-dataset
          Houses-dataset
          Rivers-dataset
          RefuseDumpSites-dataset
          Individuals-dataset
          border
          water-height
           ]

;;different turtles
breed [ individuals individual ]
breed [ households household ]
breed [ houses house ]
breed [ waters water ]
breed [ raindrops raindrop ]
breed [ wastes waste ]
breed [ foods food ]

;;declaring attributes
patches-own [ terrain houseflies waterflow contaminated-water? fliesInfected?
dumpsite-infection-level ]
houses-own [ TapeWater ]
individuals-own [ personID Ages BloodTyp_1 Gender HouseholdI infected?
homeLocation waterpointLocation dumpsiteLocation incomeLeve
TapeWater hygieneLev recovery-period recovered? carrying-contaminated-water?
 HH-infected? VT-infected? HEH-infected? EH-infected? immune?
 dumpsiteNumber source soiledHands? ]
households-own [ householdID incomeLevel hygieneLev ]
foods-own [ contaminated? homeLocation ]
waters-own [ contaminated? stay-period ]
raindrops-own [ contaminated? stay-period ]
wastes-own [ contaminated? dumpsiteLocation ]
```

```
;;Setting up procedures loading all raster layers
to Setup
  clear-all
  clear-turtles
  clear-all-plots

; read elevation file
 file-open "elevationn.txt"
 foreach sort patches [ask ? [ set terrain file-read
 set pcolor scale-color brown terrain 230 280 ]]
 file-close

  ;set day-length day-length

  set water-height 3
  set border patches with [ count neighbors != 8 ]


  ;;read houseflies file
  file-open "flieshouse.txt"
  foreach sort patches [ ask ? [ set houseflies file-read ]]
     ;set pcolor scale-color yellow houseflies 0 1912 ]]
  file-close

  set Rivers-dataset gis:load-dataset "data/Rivers.shp"
  set Houses-dataset gis:load-dataset "data/Houses.shp"
  set RefuseDumpSites-dataset gis:load-dataset "data/RefuseDumpSites.shp"
  gis:set-world-envelope-ds (gis:envelope-union-of (gis:envelope-of Rivers-dataset)
                                    (gis:envelope-of Houses-dataset)
                                    (gis:envelope-of RefuseDumpSites-dataset))
  gis:set-drawing-color sky
  gis:draw rivers-dataset 2
  gis:set-drawing-color black
  gis:draw Houses-dataset 1.5
  gis:set-drawing-color orange
  gis:draw refuseDumpSites-dataset 6

   ;;read waterflow file
  file-open "experiement.txt"
  foreach sort patches [ ask ? [ set waterflow file-read ]]
  file-close

  run InitialPopulation


end

;;creating individuals from a point shapefile
```

```
to FromShapefile
  set Individuals-dataset gis:load-dataset "data/Individuals.shp"
 gis:set-world-envelope-ds (gis:envelope-union-of (gis:envelope-of Individuals-dataset)
                                                (gis:envelope-of Houses-dataset)
                                                (gis:envelope-of Rivers-dataset))

  ;create individual agents from point shapefile
  foreach gis:feature-list-of Individuals-dataset
  [
    let location gis:location-of gis:centroid-of ?
    if not empty? location
    [
      create-individuals 1
      [
        set homeLocation location
        set size 3
        set Ages gis:property-value ? "AGES"
        set personID gis:property-value ? "PersonID"
        set Gender gis:property-value ? "Gender"
        set HouseholdI gis:property-value ? "HouseholdI"
        set BloodTyp_1 gis:property-value ? "BloodTyp_1"
        set incomeLeve gis:property-value ? "incomeLeve"
        set TapeWater gis:property-value ? "Tapewater"
        set hygieneLev gis:property-value ? "hygieneLev"
        set infected? false
        set recovered? false
        set immune? false
        set carrying-contaminated-water? false
        set recovery-period 0
        set HH-infected? false
        set VT-infected? false
        set HEH-infected? false
        set EH-infected? false
        set soiledHands? false
        set xcor item 0 homeLocation
        set ycor item 1 homeLocation
        set color blue

        ;; choose where to get water
        let smallestDistance 1000
        foreach gis:feature-list-of Rivers-dataset
        [
          foreach gis:vertex-lists-of ?
          [
            foreach ?
            [
              set location gis:location-of ?
              let locationX item 0 location
```

```
      let locationY item 1 location

      let dist distancexy locationX locationY
      if dist < smallestDistance
      [
        set smallestDistance dist
        set waterpointLocation list locationX locationY
      ]
    ]
  ]
]

let locationX item 0 waterpointLocation
let locationY item 1 waterpointLocation
let waterPatch patch locationX locationY
ask waterPatch
[
  ifelse EH?
  [
    set contaminated-water? (random 100) < 30
  ]
  [
    set contaminated-water? false
  ]
]

; choose the nearest dumpsite
let leastdistance 1000
foreach gis:feature-list-of RefuseDumpSites-dataset
[ foreach gis:vertex-lists-of ?
  [ foreach ?
    [ set location gis:location-of ?
      let locationX1 item 0 location
      let locationY1 item 1 location
      let dumpsitePatch patch locationX1 locationY1
      ask dumpsitePatch
      [
        set dumpsite-infection-level 0
      ]

      let dist distancexy locationX1 locationY1
      if dist < leastDistance
      [
        set leastDistance dist
        set dumpsiteLocation list locationX1 locationY1
        ]
        ]
      ]
```

```
            ]
        ]
    ]
  ]
end
;;go procedure
to Go
  tick
  update-plots
  fetch-water
  dump-waste
  play-at-dumpsite
  cook-food
  eat-food
 ; eating-food
  drink-water
  recover-from-illness
  transmission
   ask border
  [
    ;; when raindrops reach the edge of the world
    ;; kill them so they exit the system and we
    ;; don’t get pooling at the edges
    ask turtles-here [ die ]
  ]
  faecal-oral-route
  dumpsite-infection
  water-die
  raindrops-die

end
;; all plots
to update-plots
  set-current-plot "Overall Health Status"
  let susceptible count individuals with [ not infected? and not recovered?]
  set-current-plot-pen "Susceptible"
  plot susceptible
  set-plot-y-range 0 3600
  set-current-plot "Overall Health Status"
  let ill count individuals with [ infected? ]
  set-current-plot-pen "Infected"
  plot ill
  set-plot-y-range 0 3600
  set-current-plot "Overall Health Status"
  set-current-plot-pen "Recovered"
  plot count individuals with [ recovered? ]
  set-current-plot "Graphs for Infected People"
  set-current-plot-pen "HH"
```

```
  plot count individuals with [ HH-infected? ]
  set-current-plot "Graphs for Infected People"
  set-current-plot-pen "HEH"
  plot count individuals with [ HEH-infected? ]
   set-current-plot "Graphs for Infected People"
  set-current-plot-pen "EH"
  plot count individuals with [ EH-infected? ]
   set-current-plot "Graphs for Infected People"
  set-current-plot-pen "VT"
  plot count individuals with [ VT-infected? ]
end
 ;;get sick from infection button
to get-sick
   set infected? true
   file-open "Results.txt"
   file-print (word PersonID"" ";" ticks"" ";" HouseholdI"" ";" BloodTyp_1"")
   file-close

   set color red
   set recovery-period ((day-length * 14) + random (day-length * 2))

   let locationX1 item 0 dumpsiteLocation
   let locationY1 item 1 dumpsiteLocation
   let dumpsitePatch patch locationX1 locationY1
   ask dumpsitePatch
   [
      set dumpsite-infection-level (dumpsite-infection-level + 1)
   ]
end
 ;;cooking
 to cook-food

  if ticks mod day-length = 1
  [  foreach gis:feature-list-of Individuals-dataset
     [
       let location gis:location-of gis:centroid-of ?
       if not empty? location
       [ set-default-shape foods "fish"
         create-foods 1
         [ set color green
           set size 1
           set contaminated? false
           set homeLocation location
           set xcor item 0 homeLocation
           set ycor item 1 homeLocation
           setxy xcor ycor

           if VT?
```

```
           [
             let nrflies 0
             ask patch-here
              [
                set nrflies houseflies
       if (nrflies > 850.000) and (random 100) < RiskOfFliesGettingInfected
                [
                  let prey one-of foods-here

                      if random 100 < 30
                       [
                          ask wastes
                       [
                          if contaminated?
                          [
                            ask prey
                            [
                              set contaminated? true
                              set color brown
                            ] ] ]
                        ] ]  ]
             ] ]  ]
       ]
     ]

end
;;drinking water
to drink-water
  if ticks mod day-length = 2
  [
    if EH?
    [ ask individuals with [carrying-contaminated-water?
    and not infected? and not recovered? ]
      [ if hygieneLev = "low" and incomeLeve = "low"
      and (random 100 < 30) and BloodTyp_1 = "O"
        [ get-sick
          set EH-infected? true
            set source "EH"
            ;file-open "EH.txt"
            ;file-print (word PersonID"" ";" ticks"" ";"
            ;source"" ";" BloodTyp_1"" ";" HouseholdI"")
            ;file-close
        ]
      ]
      ask individuals with [ carrying-contaminated-water?
      and not infected? and not recovered? ]
      [ if hygieneLev = "low" and incomeLeve = "low"
      and (random 100 < 15) and BloodTyp_1 = "other"
```

```
    [ get-sick
      set EH-infected? true
        set source "EH"
        file-open "EH.txt"
        file-print (word PersonID"" ";" ticks"" ";"
        source"" ";" BloodTyp_1"" ";" HouseholdI"")
        file-close
    ]
  ]
   ask individuals with [ carrying-contaminated-water?
   and not infected? and not recovered? ]
  [ if hygieneLev = "aver" and incomeLeve = "low"
  and (random 100 < 10) and BloodTyp_1 = "O"
    [ get-sick
      set EH-infected? true
        set source "EH"
        file-open "EH.txt"
        file-print (word PersonID"" ";" ticks"" ";"
        source"" ";" BloodTyp_1"" ";" HouseholdI"")
        file-close
    ]
  ]
]

if HEH?
[
  ask individuals with [ carrying-contaminated-water?
  and not recovered? and not infected? ]
  [
    if hygieneLev = "low" and incomeLeve = "middle"
    and (random 100 < 30) and BloodTyp_1 = "O"
    [ get-sick
      set HEH-infected? true
      ;set source "HEH"
      ;file-open "HEH.txt"
      ;file-print (word PersonID"" ";" ticks"" ";"
      ;source"" ";" BloodTyp_1"" ";" HouseholdI"")
      ;file-close
    ]
  ]
   ask individuals with [ carrying-contaminated-water?
   and not recovered? and not infected? ]
  [
   if hygieneLev = "low" and incomeLeve = "middle"
   and (random 100 < 15) and BloodTyp_1 = "other"
    [ get-sick
      set HEH-infected? true
      ;set source "HEH"
```

```
          ;file-open "HEH.txt"
          ;file-print (word PersonID"" ";" ticks"" ";"
          ;source"" ";" BloodTyp_1"" ";" HouseholdI"")
          ;file-close
        ]
      ]
      ask individuals with [ carrying-contaminated-water?
      and not infected? and not recovered? ]
      [
        if hygieneLev = "low" and incomeLeve = "high"
        and (random 100 < 10) and BloodTyp_1 = "O"
        [ get-sick
          set HEH-infected? true
          set source "HEH"
          file-open "HEH.txt"
          file-print (word PersonID"" ";" ticks"" ";"
          source"" ";" BloodTyp_1"" ";" HouseholdI"")
          file-close
        ]
      ]
       ask individuals with [ carrying-contaminated-water?
       and not infected? and not recovered? ]
      [
        if hygieneLev = "low" and incomeLeve = "high"
        and (random 100 < 5) and BloodTyp_1 = "other"
        [ get-sick
          set HEH-infected? true
          set source "HEH"
          file-open "HEH.txt"
          file-print (word PersonID"" ";" ticks"" ";"
          source"" ";" BloodTyp_1"" ";" HouseholdI"")
          file-close
        ]
      ]
    ]
  ]
end
;;dumping waste at a dumpsite
to dump-waste
   ask individuals
  [ if Ages < 12 and Ages > 5
     [ ifelse ticks  mod day-length = 1 and not infected?
      [
        let locationX1 item 0 dumpsiteLocation
        let locationY1 item 1 dumpsiteLocation
        setxy locationX1 locationY1
      ]
      [
```

```
        let locationX item 0 homeLocation
        let locationY item 1 homeLocation
        setxy locationX locationY
      ]
    ]
  ]
end
;;eating
to eat-food
   if ticks mod day-length = 2
   [
     ask individuals
     [
      let prey one-of foods-here
      if prey != nobody
      [
        ask prey
        [
         if VT?
          [
            if contaminated?
           [
             ask individuals-here
            [
              if not recovered? and not infected? and hygieneLev = "low"
              and (random 100 < 20) and BloodTyp_1 = "O" and incomeLeve = "low"
             [
               get-sick
               set VT-infected? true
               set source "VT"
                file-open "VT.txt"
                file-print (word PersonID"" ";" ticks"" ";" source"" ";" HouseholdI"")
                file-close
             ]
               if not recovered? and not infected? and hygieneLev = "aver"
               and (random 100 < 3) and BloodTyp_1 = "O" and incomeLeve = "low"
              [
                get-sick
                set VT-infected? true
                set source "VT"
                file-open "VT.txt"
                file-print (word PersonID"" ";" ticks"" ";" source"" HouseholdI"")
                file-close
             ]
              if not recovered? and not infected? and hygieneLev = "low"
              and (random 100 < 10) and BloodTyp_1 = "O" and incomeLeve = "aver"
              [
                get-sick
```

```
                set VT-infected? true
                set source "VT"
                file-open "VT.txt"
                file-print (word PersonID"" ";" ticks"" ";" source"" HouseholdI"")
                file-close
             ]
              if not recovered? and not infected? and hygieneLev = "low"
              and (random 100 < 20) and BloodTyp_1 = "O" and incomeLeve = "high"
              [
                get-sick
                set VT-infected? true
                set source "VT"
                file-open "VT.txt"
                file-print (word PersonID"" ";" ticks"" ";" source"" ";" HouseholdI"")
                file-close
             ]
           ]
          ]
         ]
          if HEH?
          [
           if not contaminated?
           [
             ask individuals-here
             [
                if not recovered? and hygieneLev = "low" and (random 100 < 10)
                and BloodTyp_1 = "O"
                and incomeLeve = "low" and soiledHands? and not infected?
              [
                get-sick
                set HEH-infected? true
                ;set source "HEH - Dump"
                ;file-open "Dump.txt"
                ;file-print (word PersonID"" ";" ticks"" ";" source"" ";" HouseholdI"")
                ;file-close
             ]
             ]
           ]
          ]
        die
      ]
     ]
    ]
   ]
end
;;fetching water
 to fetch-water
```

```
ask individuals
[ if Ages >= 15 and TapeWater = 0  and incomeLeve = "low"
  [
    ifelse ticks mod day-length = 0 and not infected?
    [
      let locationX item 0 waterpointLocation
      let locationY item 1 waterpointLocation
      setxy locationX locationY

      let contam? false
      ask patch-here
      [
        set contam? contaminated-water?
      ]
      set carrying-contaminated-water? contam?
    ]
    [
      let locationX item 0 homeLocation
      let locationY item 1 homeLocation
      setxy locationX locationY
    ]
  ]
  ]

   ask individuals
  [
   if Ages >= 15 and incomeLeve = "high"
   [
    ifelse ticks mod (day-length * 3) = 0 and not infected?
    and (random 100 < RiskOfChildrenGettingInfected)

    [
      let locationX item 0 waterpointLocation
      let locationY item 1 waterpointLocation
      setxy locationX locationY

      let contam? false

      ask raindrops in-radius 1
      [
        set contam? contaminated?
      ]
      set carrying-contaminated-water? contam?
    ]
    [
      let locationX item 0 homeLocation
      let locationY item 1 homeLocation
      setxy locationX locationY
```

```
      ]

    ]
   ]

   ask individuals
   [
   if Ages >= 15 and incomeLeve = "middle"
    [
      ifelse ticks mod (day-length * 2) = 0 and not infected?
      and (random 100 < RiskOfChildrenGettingInfected)
      [
        let locationX item 0 waterpointLocation
        let locationY item 1 waterpointLocation
        setxy locationX locationY

        let contam? false
        ask raindrops in-radius 1
        [
          set contam? contaminated?
        ]
        set carrying-contaminated-water? contam?
      ]
      [
        let locationX item 0 homeLocation
        let locationY item 1 homeLocation
        setxy locationX locationY
      ]

    ]
   ]

end
 ;;creating rainfall ready for HEH transmission
to faecal-oral-route
  if HEH?
  [ if ticks mod (day-length * 2) = 2

     [ create-raindrops 200
       [
        move-to one-of patches
        set size 2
        set color sky
        set contaminated? false
        set stay-period 0
       ]
     ]
      ask raindrops
```

```
    [
      flow

      ask patch-here
      [
        if pcolor = red
        [
          ask raindrops in-radius 5
          [
            set color red
            set contaminated? true
          ]
        ]
      ]
      if contaminated?
      [
        ask raindrops in-radius 2
        [
          set color red
          set contaminated? true
        ]
      ]
    ]
   ask individuals
  [ if Ages >= 15 and incomeLeve = "middle"
   [
    ifelse ticks mod (day-length * 3) = 0 and not infected?
    [
      let locationX item 0 waterpointLocation
      let locationY item 1 waterpointLocation
      setxy locationX locationY

      let contam? false
      ask raindrops in-radius 3
      [
        set contam? contaminated?
      ]
      set carrying-contaminated-water? contam?
    ]
    [
      let locationX item 0 homeLocation
      let locationY item 1 homeLocation
      setxy locationX locationY
    ]

   ]
  ]
```

```
  ]


end
;;raindrops flow
to flow ;; turtle procedure
  ;; get the lowest neighboring patch taking into account
  ;; how much water is on each patch.
  let target min-one-of neighbors [waterflow + (count turtles-here * water-height)]
  ;; if the elevation + water on the neighboring patch is
  ;; lower than here move to that patch.
  ifelse [waterflow + (count turtles-here * water-height)] of target
     < (waterflow + (count turtles-here * water-height))
    [ move-to target ]
    [ set breed waters ]
    ask waters
    [ set contaminated? false
      set stay-period 0
      if color = red
      [ ask waters in-radius 3
        [ set color red
          set contaminated? true ]
        ]
      ]

end
;;eliminating unwanted water
to water-die
  ask waters
  [ set stay-period (stay-period + 1)
    if not contaminated?
    [
      if stay-period > day-length
      [ die ]
    ]
  ]

end
;;eliminating unwanted raindrops
to raindrops-die
  ask raindrops
  [ set stay-period (stay-period + 1)
    if not contaminated?
    [
      if stay-period > day-length
      [ die ]
    ]
  ]
```

```
end

to transmission

  ;;Human-to-Human transmission
  if ticks mod day-length = 2
  [
  if HH?
  [
    ask individuals with [ infected? ]
    [
      let my-friends individuals in-radius 0.1 with [ self != myself ]
       ;people around me but not myself
      ask my-friends with [ not infected?  and not recovered? ]
      [
        if hygieneLev = "low" and (random 100) < RiskOfHH and BloodTyp_1 = "O"
        and incomeLeve = "low"
        [
          get-sick
          set HH-infected? true
          set source "HH"
          file-open "HH.txt"
          file-print (word PersonID"" ";" ticks"" ";" HouseholdI"" ";"
          BloodTyp_1"" ";" source"")
          file-close
        ]
        if hygieneLev = "low" and (random 100) < (RiskOfHH / 2)
        and BloodTyp_1 = "other" and incomeLeve = "low"
        [
          get-sick
          set HH-infected? true
          ;set source "HH"
          ;file-open "HH.txt"
          ;file-print (word PersonID"" ";" ticks"" ";" HouseholdI""
          ;";" BloodTyp_1"" ";" source"")
          ;file-close
        ]
        if hygieneLev = "low" and (random 100) < (RiskOfHH / 3)
        and BloodTyp_1 = "O" and incomeLeve = "aver"
        [
          get-sick
          set HH-infected? true
          set source "HH"
          file-open "HH.txt"
          file-print (word PersonID"" ";" ticks"" ";" HouseholdI""
           ";" BloodTyp_1"" ";" source"")
          file-close
```

```
        ]
      ]
      ]
    ]
  ]
end
;;infecting the dumpsite
to dumpsite-infection
  foreach gis:feature-list-of RefuseDumpSites-dataset
    [
      let location gis:location-of gis:centroid-of ?
      if not empty? location
       [
          if ticks mod day-length = 2
           [ create-wastes 1
        [ set contaminated? false
          set dumpsiteLocation location
          set xcor item 0 dumpsiteLocation
          set ycor item 1 dumpsiteLocation
          setxy xcor ycor

         let dumpsitePatch patch xcor ycor
         ask dumpsitePatch
           [ ifelse dumpsite-infection-level >= 10
             [
                ask patches in-radius 4
                [
                   set pcolor red
                   ask wastes
                   [ set contaminated? true
                   ]
                ]
              ]

           [
             ask patches in-radius 4
             [
                set pcolor orange
             ]
          ]
        ]
       ]
     ]
     ]
   ]
end

;;cildren playing at dumpsite
```

```
to play-at-dumpsite
  ask individuals
   [ if Ages < 12 and Ages > 5 and incomeLeve = "low" and
   (random 100 < 10) and not infected? and not recovered?
    [
      ifelse ticks  mod (day-length * 5) = 1
       [
         let locationX1 item 0 dumpsiteLocation
         let locationY1 item 1 dumpsiteLocation
         setxy locationX1 locationY1
         if dumpsite-infection-level >= 100 and
         (random 100 < RiskOfChildrenGettingInfected)
          [
            ask individuals-here
            [
              set soiledHands? true
            ]
          ]
         ]
       [
        let locationX item 0 homeLocation
        let locationY item 1 homeLocation
        setxy locationX locationY
      ]
     ]
    ]

end
;;recovering from illness
to recover-from-illness
  ask individuals with [ infected? ]
  [
     ifelse recovery-period > 0
     [
       set recovery-period recovery-period - 1
     ]
     [
       set infected? false
       set recovered? true
       set immune? true
       set color black
     ]
  ]
end

;;report to  monitors
to-report HH-infected
  report count individuals with [ HH-infected? ]
```

```
  report 0
end

to-report HEH-infected
  report count individuals with [ HEH-infected? ]
  report 0
end

to-report EH-infected
  report count individuals with [ EH-infected? ]
  report 0
end

to-report VT-infected
  report count individuals with [ VT-infected? ]
  report 0
end
```

LIST OF REFERENCES

[1] *Tutorial on agent-based modeling and simulation*. Winter Simulation Conference, 2005.

[2] Shaheen A. Abdulkareem. Simulating the spread of pertussis in enschede region using agent-based modelling. Master's thesis, ITC, 2010.

[3] A. Agheksanterian and M.K. Gobbert. Modeling the Spread of Epidemic Cholera: an Age-Structured Model. Technical report, Citeseer, 2007.

[4] M. Ali, M. Emch, J.P. Donnay, M. Yunus, and RB Sack. Identifying environmental risk factors for endemic cholera: a raster GIS approach. *Health & Place*, 8(3):201–210, 2002.

[5] M. Ali, M. Emch, J.P. Donnay, M. Yunus, and RB Sack. The spatial epidemiology of cholera in an endemic area of Bangladesh. *Social Science & Medicine*, 55(6):1015–1024, 2002.

[6] Kumasi Metropolitan Assembly. About this metropolis: Housing. `http://www.ghanadistricts.com/districts1on1/kma/?arrow=atd&_=6&sa=5490`, 2006.

[7] Ellen-Wien Augustijn-Beckers and Rolf A. de By. Towards routine large-scale, discrete spatial event simulations. `http://www.geocomputation.org/2007/Posters/P06.pdf`, 2007.

[8] R. Axelrod and L. Tesfatsion. A guide for newcomers to agent-based modeling in the social sciences, inHandbook of Computational Economics', Vol. 2, 2006.

[9] Stefania Bandini, Sara Manzoni, and Giuseppe Vizzari. Agent based modeling and simulation: An informatics perspective. *Journal of Artificial Societies and Social Simulation*, 12(4):4, 2009.

[10] Victoria Beddow. Kumasi: Sanitation status. `http://iwawaterwiki.org/xwiki/bin/view/Articles/24)+KUMASI+(Ghana)+3`, 2010.

[11] E. Bertuzzo, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe, and A. Rinaldo. On spatially explicit models of cholera epidemics. *Journal of the Royal Society Interface*, 7(43):321, 2010.

[12] E. Bonabeau. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the United States of America*, 99(Suppl 3):7280, 2002.

[13] M. Bouden, B. Moulin, and P. Gosselin. The geosimulation of West Nile virus propagation: a multi-agent and climate sensitive tool for risk management in public health. *International Journal of Health Geographics*, 7(1):35, 2008.

[14] D. Buor. Water needs and women's health in the Kumasi metropolitan area, Ghana. *Health & Place*, 10(1):85–103, 2004.

[15] CCJ Carpenter. Principles and Practice of Cholera Control. *Publ Hlth Papers WHO No*, 40, 1970.

[16] F. Castiglione. Agent based modeling. *Scholarpedia*, 1(10):1562, 2006.

[17] C.J.E. Castle and A.T. Crooks. Principles and concepts of agent-based modelling for developing geospatial simulations. *Centre for Advanced Spatial Analysis (University College London): Working Paper*, 110, 2006.

[18] Elizabeth M. Chingayipe. Factors influencing cholera case detection by the communities in chiradzulu district. Master's thesis, Malawi University, 2008.

[19] The Chronicle. Kma boss tours flooded areas in kumasi. `http://allafrica.com/stories/200907131709.html`, 2009.

[20] CNN. Cholera. `http://edition.cnn.com/HEALTH/library/cholera/DS00579.html`, 2009.

[21] C.T. Codeço. Endemic and epidemic dynamics of cholera: the role of the aquatic reservoir. *BMC infectious diseases*, 1(1):1, 2001.

[22] A.E. Collins, ME Lucas, MS Islam, and LE Williams. Socio-economic and environmental origins of cholera epidemics in Mozambique: guidelines for tackling uncertainty in infectious disease prevention and control. *International journal of environmental studies*, 63(5):537–549, 2006.

[23] R. Connell, P. Dawson, and A. Skvortsov. Comparison of an Agent-based Model of Disease Propagation with the Generalised SIR Epidemic Model. 2009.

[24] G. Constantin de Magny, B. Cazelles, and J.F. Guégan. Cholera threat to humans in Ghana is influenced by both global and regional climatic variability. *EcoHealth*, 3(4):223–231, 2006.

[25] G. Constantin de Magny, R. Murtugudde, M.R.P. Sapiano, A. Nizam, C.W. Brown, A.J. Busalacchi, M. Yunus, G.B. Nair, A.I. Gil, C.F. Lanata, et al. Environmental signatures associated with cholera epidemics. *Proceedings of the National Academy of Sciences*, 105(46):17676, 2008.

[26] NS Crowcroft. Cholera: current epidemiology. *Communicable disease report. CDR review*, 4(13):R157, 1994.

[27] Minnesota department of Health. Causes and symptoms of cholera. `http://www.health.state.mn.us/divs/idepc/diseases/cholera/basics.html`, 2009.

[28] O. Diekmann and J. A. P. Heesterbeek. *Mathematical epidemiology of infectious diseases: model building, analysis*. Harvard University, 2000.

[29] DNA. Low power voltage affecting water supply in kumasi. `http://www.ghanaweb.com/GhanaHomePage/NewsArchive/artikel.php?ID=177728`, 2010.

[30] J.B. Dunham. An agent-based spatially explicit epidemiological model in MASON. *Journal of Artificial Societies and Social Simulation*, 9(1):3, 2005.

[31] J. Edwards. Agent-Based Epidemiology Modelling.

[32] eMedTV. Cholera symptoms. `http://diarrhea.emedtv.com/cholera/cholera-symptoms.html`, 2010.

[33] O. Felsenfeld. Notes on food, beverages and fomites contaminated with Vibrio cholerae. *Bulletin of the World Health Organization*, 33(5):725, 1965.

[34] Christian Fertner. Complexity and agent-based modelling in urban research. `http://www.byforskning.dk/publikationer/Siden%20publikationer/Working%20papers/WP17.pdf`, 2008.

[35] R. Fotedar. Vector potential of houseflies (Musca domestica) in the transmission of Vibrio cholerae in India. *Acta tropica*, 78(1):31–34, 2001.

[36] F. Gargiulo, S. Ternes, S. Huet, and G. Deffuant. An iterative approach for generating statistically realistic populations of households. *PloS one*, 5(1):95–184, 2010.

[37] A. Getchell. Agent-based Modeling. 2008.

[38] Ellen Geurts. Working on cities: an experiement from kumasi, ghana. `http://web.up.ac.za/sitefiles/file/44/1068/3229/9086/African%20Perspectives/PDF/Papers/GEURTS.pdf`, 2009.

[39] V. Grimm, U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, J. Goss-Custard, T. Grand, S.K. Heinz, G. Huse, et al. A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, 198(1-2):115–126, 2006.

[40] V. Grimm, E. Revilla, U. Berger, F. Jeltsch, W.M. Mooij, S.F. Railsback, H.H. Thulke, J. Weiner, T. Wiegand, and D.L. DeAngelis. Pattern-oriented modeling of agent-based complex systems: lessons from ecology. *Science*, 310(5750):987, 2005.

[41] D.M. Hartley, J.G. Morris, and D.L. Smith. Hyperinfectivity: a critical element in the ability of V. cholerae to cause epidemics? *PLoS Medicine*, 3(1):63, 2006.

[42] H. Hu, P. Gong, and B. Xu. Spatially explicit agent-based modelling for schistosomiasis transmission: Human-environment interaction simulation and control strategy assessment. *Epidemics*, 2010.

[43] A. Huq, R.B. Sack, A. Nizam, I.M. Longini, G.B. Nair, A. Ali, J.G. Morris Jr, MN Khan, A. Siddique, M. Yunus, et al. Critical factors influencing the occurrence of Vibrio cholerae in the environment of Bangladesh. *Applied and environmental microbiology*, 71(8):4645, 2005.

[44] T. Iba, Y. Matsuzawa, and N. Aoyama. From conceptual models to simulation models: Model driven development of agent-based simulations. In *9th Workshop on Economics and Heterogeneous Interacting Agents*. Citeseer, 2004.

[45] M.S. Islam, B.S. Drasar, and R.B. Sack. The aquatic environment as a reservoir of Vibrio cholerae: a review. *Diarrhoeal Diseases Research*, page 197.

[46] N.R. Jennings and M.J. Wooldridge. *Agent technology: foundations, applications, and markets*. Springer Verlag, 1998.

[47] J.S. Kuma, R.O. Owusu, and S.K.Y. Gawu. Evaluating the Water Supply System in Kumasi, Ghana. *European Journal of Scientific Research*, 40(4):506–514, 2010.

[48] K.M. Kwofie. A spatio-temporal analysis of cholera diffusion in Western Africa. *Economic Geography*, 52(2):127–135, 1976.

[49] C. Linard, N. Ponçon, D. Fontenille, and E.F. Lambin. A multi-agent simulation to assess the risk of malaria re-emergence in southern France. *Ecological Modelling*, 220(2):160–174, 2009.

[50] E.K. Lipp, A. Huq, and R.R. Colwell. Effects of global climate on infectious disease: the cholera model. *Clinical microbiology reviews*, 15(4):757, 2002.

[51] B. Logan and G. Theodoropoulos. The distributed simulation of multi-agent systems.

[52] Mikola Lysenko and Roshan M. D'Souza. A framework for megascale agent based model simulations on graphics processing units. *Journal of Artificial Societies and Social Simulation*, 11(4):10, 2008.

[53] C.M. Macal and M.J. North. Agent-based modeling and simulation: Desktop ABMS. In *Proceedings of the 39th conference on Winter simulation: 40 years! The best is yet to come*, pages 95–106. IEEE Press, 2007.

[54] A. Mawudeku and M. Blench. Global public health intelligence network (GPHIN). In *7th Conference of the Association for Machine Translation in the Americas*, pages 8–12, 2006.

[55] E.D. Mintz, T. Popovic, and P.A. Blake. Transmission of Vibrio cholerae Ol. *Vibrio cholerae and cholera: molecular to global perspectives*, page 345, 1994.

[56] R. Moeckel, K. Spiekermann, and M. Wegener. Creating a synthetic Population. In *Proceedings of the 8th International Conference on Computers in Urban Planning and Urban Management (CUPUM)*, 2003.

[57] F.B. Osei. SPATIAL STATISTICS OF EPIDEMIC DATA: THE CASE OF CHOLERA EPIDEMIOLOGY IN GHANA. 2010.

[58] F.B. Osei and A.A. Duker. Spatial and demographic patterns of Cholera in Ashanti region-Ghana. *International Journal of Health Geographics*, 7(1):44, 2008.

[59] Frank B. Osei and Alfreid A. Duker. Spatial and demographic patterns of cholera in ashanti region - ghana. *International Journal of Health Geographics*, 2008a.

[60] Frank B. Osei, Alfreid A. Duker, Ellen-Wien Augustijn, and Afred Stein. Spatial dependency of cholera prevalence on potential cholera reservoirs in an urban area, kumasi, ghana. *International Journal of Applied Earth Observation and Geoinformation*, 2010.

[61] S.E. Page. Agent based models. *The New Palgrave Dictionary of Economics*, 2006.

[62] P. Patlolla, V. Gunupudi, A. Mikler, and R. Jacob. Agent-Based Simulation Tools in Computational Epidemiology. *Innovative Internet Community Systems*, pages 212–223, 2006.

[63] M. Rabbani, G. Hassan, and WB Greenough. Food as a vehicle of transmission of cholera.

[64] S.F. Railsback, S.L. Lytinen, and S.K. Jackson. Agent-based simulation platforms: Review and development recommendations. *Simulation*, 82(9):609–623, 2006.

[65] Home Remedies. Cholera. `http://www.home-remedies-for-you.com/remedy/Cholera.html`, 2004.

[66] Duncan A. Robertson. Agent-based modeling toolkits netlogo, repast and swarm (book & resorce reviews). *Academy of Management*, pages 525–527, 2005.

[67] J. Sakshaug and T. Raghunathan. Synthetic Data for Small Area Estimation. In *Privacy in Statistical Databases*, pages 162–173. Springer.

[68] S. Sasaki, H. Suzuki, K. Igarashi, B. Tambatamba, and P. Mulenga. Spatial analysis of risk factor of cholera outbreak for 2003-2004 in a peri-urban area of Lusaka, Zambia. *The American journal of tropical medicine and hygiene*, 79(3):414, 2008.

[69] P. Scheelbeek, S. Treglown, T. Reid, and P. Maes. Household fish preparation hygiene and cholera transmission in Monrovia, Liberia. *The Journal of Infection in Developing Countries*, 3(9):727, 2009.

[70] M.C.S.L.I. Seminar. Introduction to Agent-based Modeling and Simulation. 2006.

[71] Joana Simoes, Chris Skelly, and Joao Simoes. Iesim: simulating and understanding world events in a game-like environment. *Worshop proceedings*, 2010.

[72] I. Sinai. Using the home for income-generation:: The case of Kumasi, Ghana. *Cities*, 15(6):417–427, 1998.

[73] Springhouse. Cholera: Excerpt from professional guide to diseases (eighth edition). `http://www.wrongdiagnosis.com/c/cholera/book-diseases-7a.htm`, 2005.

[74] A. Srbljinovic and O. Skunca. An introduction to agent based modelling and simulation of social processes. *Arxiv preprint cond-mat/0409312*, 2004.

[75] A.B. Sutiono, H. Suwa, and T. Ohta. MULTI AGENT BASED SIMULATION FOR TYPHOID FEVER WITH COMPLICATIONS: AN EPIDEMIC ANALYSIS. In *Proceedings of the 51st Annual Meeting of the ISSS*, volume 51, 2007.

[76] D.L. Swerdlow, E.D. Mintz, M. Rodriguez, E. Tejada, C. Ocampo, L. Espejo, T.J. Barrett, J. Petzelt, N.H. Bean, L. Seminario, et al. Severe life-threatening cholera associated with blood group O in Peru: implications for the Latin American epidemic. *The Journal of infectious diseases*, 170(2):468–472, 1994.

[77] W. Tang. Development of a Spatially-explicit Agent-based Simulation Package for Modeling Complex Adaptive Geographic Systems. *Proceedings of the UCGIS Summer Assembly*, 2007.

[78] E. Teweldemedhin, T. Marwala, and C. Mueller. Agent-based modelling: a case study in HIV epidemic. In *Fourth International Conference on Hybrid Intelligent Systems, 2004. HIS'04*, pages 154–159, 2004.

[79] A.G. Tipple and JA Hellen. Priorities for public utilities and housing improvements in Kumasi, Ghana: an empirical assessment based on six variables. *Géographie des systèmes de santé. Urbanisation et santé dans le Tiers Monde. Transition épidémiologique changement social et soins de santé primaire. Tours: ORSTOM collection colloques et séminaires*, pages 199–226, 1989.

[80] S. Vattam, A.K. Goel, S. Rugaber, C. Hmelo-Silver, and R. Jordan. From Conceptual Models to Agent-based Simulations: Why and How. In *Fourteenth International Conference on AI in Education, Brighton UK*, 2009.

[81] K. Wachsmuth, I.K. Wachsmuth, P.A. Blake, and Ø. Olsvik. *Vibrio cholerae and cholera: molecular to global perspectives.* Amer Society for Microbiology, 1994.

[82] J. Wallinga and P. Teunis. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *American Journal of Epidemiology*, 160(6):509, 2004.

[83] D. Whittington, D.T. Lauria, K. Choe, J.A. Hughes, M. Swarna Albert, et al. Household sanitation in Kumasi, Ghana: a description of current practices, attitudes, and perceptions. *World Development*, 21(5):733–748, 1993.

[84] WHO. Cholera. `http://www.emro.who.int/sudan/pdf/factsheet_cholera_07.pdf`, 2000.

[85] WHO. Weekly epidemiological record. `http://www.who.int/wer/2010/wer8513.pdf`, 2010.

[86] Wikipedia. Cholera. `http://en.wikipedia.org/wiki/Cholera`, 2010.

[87] Wikipedia. Conceptual model (computer science). `http://en.wikipedia.org/wiki/Conceptual_model_(computer_science)`, 2010.

[88] Wikipedia. Kumasi. `http://en.wikipedia.org/wiki/Kumasi`, 2010.

[89] Emma Wikner. Modeling waste to energy systems in kumasi, ghana. Master's thesis, Swedish University of Agricultural Sciences (SLU) in Uppsala, 2009.

[90] U. Wilensky et al. NetLogo 4.1.2 User Manual, 2010.

[91] Yuhong Zhou. Explorative research on methods for discrete space time simulation integrated with the event based approach and agent concept. Master's thesis, ITC, 2006.