Agent-based modeling for Humanitarian Issues: Disease and Refugee Camps

Atesmachew Hailegiorgis[1] and Andrew T. Crooks[1]

[1] Center for Social Complexity, Krasnow Institute for Advanced Study, George Mason University
379 Research Hall, MS 6B2, Fairfax, VA 22030. USA
{ahailegi, acrooks2}@gmu.edu

Abstract. The displacement of people in times of crises represents a challenge for humanitarian agencies. This challenge is especially acute within developing countries, which home the majority of displaced people. Within this paper, we will demonstrate a spatially explicit agent based model that explores the spread of cholera in the Dadaab refugee camps. Poor sanitation and housing conditions contribute to frequent incidents of cholera outbreaks. We model the spread of cholera by explicitly representing the interaction between humans (host) and their environment, and the spread of the epidemic using Susceptible-Exposed-Infected-Recovered (SEIR) model. We utilize elevation surface data combined with rainfall to carry pollutants (disposal and feces). We model agents as mobile and purposeful oriented individuals engaging in daily activities all of whom are susceptible to diseases. Infected agents spread cholera bacteria through excretion of feces to the environment and this can then be spread throughout the system. Results from the model show that the spread of cholera grows radially from contaminated water sources. Agents' social behavior and movements contribute to the spread of cholera to other camps where water sources were relatively safe. This modeling effort highlights the potential of agent based modeling to explore the spread of cholera in a humanitarian context and its impact on service provision.

Keywords: Agent-based modeling, Disease propagation, Cholera, GIS, Refugee camps.

1 Introduction

People either become displaced by man-made and natural disasters such as drought, flooding, ethnic conflict, and political violence. The United Nations High Commissioner for Refugees (UNHCR, 2011a) Global Trend report indicated that about 44 million people were displaced worldwide in 2010 with the majority being in developing countries. With 27.5 million people being displaced within their own country (internally displaced persons), while the rest were became refugees in other countries (international refugees). The displacement of people in times of crises represents a challenge for humanitarian agencies who attempt to provide basic services (such as food, water and healthcare) at camps or other locations to ease the suffering of these displaced people. Displaced people is especially acute in Africa, where the most common causes of death have been diarrheal diseases, measles, acute respiratory infections and malaria (Toole & Waldman, 1997). Many of which are preventative in the developed world.

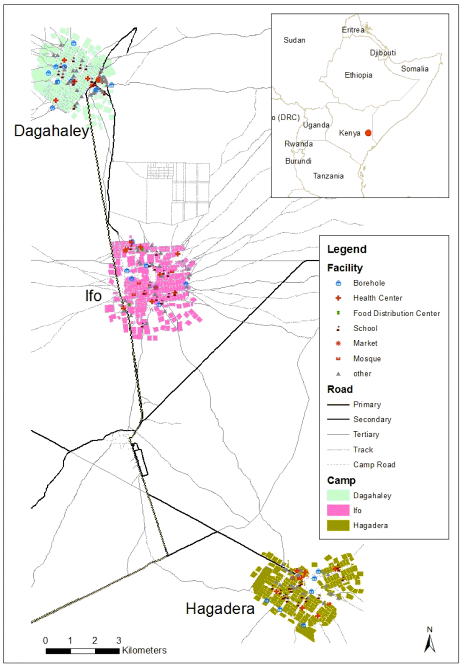

Fig. 1. Location of the study area.

Cholera is an acute problem in Africa especially when related to refugee camps (Siddique, 1994; Swerdlow *et al.,* 1997) as they often suffer from poor sanitation and low per capita availability of water (Cronin *et al.,* 2008). Within this paper we focus one such camp complex. That of the Dadaab refugee camps which are located near the Kenya-Somalia border as shown in Fig. 1. This camp complex hosts nearly half a million refugees (UNHCR, 2012) many of which come from Somalia due to drought, famine, and violence. The camps have poor sanitation and housing conditions with frequent incidences of cholera outbreaks (see UNHCR, 2011a).

This model is part of the larger RiftLand modeling effort (Cioffi-Revilla *et al,* 2011). The RiftLand agent-based model seeks to model human subsistence, conflict and displacement in East Africa at multiple spatial scales. The RiftLand model covers and area approximately 2.5 million km^2 including all of Kenya, Uganda, Rwanda and Burundi; and parts of Sudan, Democratic Republic of Congo and Tanzania. While other submodels have focused on vegetation growth (Gulden *et al.,* 2011a) or the displacement and migration of households (Gulden *et al.,* 2011b), non of them have explored the fine scale dynamics of refugee camps and the transmission of diseases.

The remainder of the paper is as follows, within Section 2 we discuss prior work with respect to refugee camps and disease modeling more generally before introducing the model in Section 3 (while Appendix 1 outlines the model in more detail). Section 4 presents some preliminary results from our work while in Section 5 we provide a summary of our paper and identify future avenues of research.

2 Background

Cholera is an intestinal disease caused by the bacterium *Vibrio cholerae*, which colonizes the human intestine (Bertuzzo *et al.,* 2010) and is characterized by diarrhea and severe dehydration. The main transmission mechanism for cholera is by drinking water or eating food contaminated by *V. cholerae* which enters the system via feces (stools) from infected people. Pandemics of cholera have been seen throughout the world from the Indian sub-continent, Africa, Europe and the Americas (Codeço, 2001). Even though cholera itself is both preventable and treatable via the treatment of raw sewage or by providing clean drinking water, using oral cholera vaccines[1] or once infected using rehydration therapy. It remains a health hazard in many developing countries where such care or prevention is not possible. It is difficult to obtain the actual numbers of deaths per year due to under or no reporting, but it is estimated to kill between 100,000–150,000 people per year (Longini *et al.,* 2007).

Our understanding of how cholera spreads and infects people relates to some of earliest work with respect to the spatial analysis and disease outbreaks. Snow (1855) used mapping to explore the association of between water contamination and the risk of cholera in the 1854 cholera outbreak in London. This analysis not only lead to the foundation of modern epidemiology (Longley *et al.,* 2010) but also showed the importance of space when exploring the spread of the diseases. Snows work has proved to be an inspiration of other studies of cholera, all of which extends his basic research. This includes exploring the spatial variations of cholera (see Oseia *et al.,* 2010) to that of the how seasonality and inter-annual variability of cholera outbreaks can be triggered by climatic events (Lobitz *et al.,* 2000). Others have used climate models to forecast the outbreak of cholera events (e.g. Reiner *et al.,* 2012).

However, while cholera has a long tradition of being mapped and modeled using spatial analysis techniques, modeling the propagation and the spread of the disease on the human population has had a much briefer history. But such models are needed as they play a crucial role in the study of infectious diseases and as a tool for exploring what might happen and comparing different control strategies (Fisman *et al.,* 2009). This relates to the notion that even spatial analysis techniques are limited with respect to capturing spatial interaction, and one might argue that even spatial autoregressive measures used as surrogates for such interaction as in the work of Oseia *et al.* (2010) are no substitute for actual agent mixings. With respect to cholera models, there are several existing models, many of which are aggregate mathematical models using differential equations (e.g. Capasso & Paveri-Fontana, 1979; Codeço, 2001). More recently a number of spatially explicit mathematical models have been developed

[1] However, these are not 100% successful; see Longini *et al.* (2007) for a discussion.

(e.g. Longini *et al.,* 2007; Bertuzzo *et al.,* 2010; Tuite *et al.,* 2011). While such styles of models have proved useful they have also been criticized by some. For example, Epstein (2009) notes that such models are ill-suited to model complex human systems as such models don't incorporate direct contact between individuals, they treat populations as aggregates and assume uniform mixing assumptions (Eubank *et al.,* 2004) and homogeneity of behaviors. This is clearly not the case as people interact with each other in many different ways (Crooks & Heppenstall, 2012).

Some have suggested (e.g. Eubank *et al.,* 2004; Epstein, 2009) that to capture the rich dynamics of human populations and to move away from the restrictive assumptions of past generations of epidemiological models we need to embrace agent-based models which focus on the individual and how through individual interaction more aggregate structures emerge, such as the spread of diseases. Agent-based models have been applied to a number of epidemiological problems ranging from the spread of the swine flu (H1N1, Epstein, 2009) or influenza (Mniszewski *et al.,* 2008), or mumps (Simoes, 2012), to the spread of airborne contaminates and the evacuations of cities (Epstein *et al.,* 2011). With respect to agent-based models and cholera, to the authors' knowledge, this is one of the first applications. The other application the authors are aware of is Augustijn-Beckers *et al.,* (2011) who extends the empirical spatial analysis work of Oseia *et al.,* 2010 to model the spread of cholera in the city Kumasi in south central Ghana. However, our model differs in two aspects to this model. First we do not use a SIR model but rather a SEIR model which we go into details about in Section 3. Secondly we are focusing on refugee camps rather than a small city where water becomes contaminated via a refuse site. Moreover, with respect to refugee camps there has been little work carried out with using agent-based models. For example Johnson *et al.,* 2009 explored violence within camps and how military personnel might respond to it. While Anderson *et al.* (2006, 2007) developed an abstract model exploring how humanitarian assistance policies implemented by governments and non-governmental organizations impact on camp refugees. Our model on the other hand is spatially explicit and models peoples daily activities from which we can than add a disease into the system and see what happens.

3 Methodology

The Appendix provides details about the inner workings of the model however to give a sense of what the model look likes, Fig. 2 shows its graphical user interface. From top left, we have the spatial environment which the agents interact within; while charts monitor the dynamics of important system statistics over time such as the age distribution, household size, health statistics and refugees daily activities. We also provide a clock for time of day and all model parameters can be set by the user. Such an interface allows for ease of use in understanding and debugging the model (Grimm, 2002).

The purpose of the model is to explore the spread of cholera in the Dadaab refugee camps by modeling the interaction of humans (hosts) with their environment. We have tried to stylize agents' behavior by incorporating their daily routine that may happen within the camps so that we can have better understanding of the dynamics of cholera transmission. We represent the environment using geo-referenced spatial data (see Appendix 1, Section 3.2).

Fig. 2. Example or the graphical user interface of the model.

The main agent in the model is the refugee agent who represents an individual who lives in the Dadaab refugee camps. A refugee agent has family and fixed home location. Agents of the same family cooperate and share resources. Agents determine their current activity based on their personal attributes (e.g. age, sex), and their current water need. They also take into consideration the time and distance when they make their goal choices (e.g. go and get water). Agents who are less than 5 years old are always stay at home unless they are sick and need to visit the health facilities. The other age groups can engage in any of the activities depending on their needs and abilities. Such as going to school, going to a mosque, getting food from a distribution center or using the latrines (see Appendix 1, Section 3.3.1).

We use a Susceptible – Exposed – Infected – Recovered (SEIR) model to capture the dynamics of the spread of cholera as unlike SIR models it is able to capture the time between ingestion of contaminated water and showing the symptom. All the refugee agents are considered as susceptible. We assume that a susceptible agent who ingests contaminated water will be become exposed to the cholera disease. We distinguished two types of infectious agents: symptomatic and asymptomatic. A symptomatic agent shows the symptom between 8 to 17 hours (Nelson *et al.,* 2009). Infants show the symptom more quickly than adults. However, if the agent is asymptomatic, the agent will pass the infection phase to the recovery stage without showing any sign of symptom. In reality as most infected people have no or mild symptoms but still are carriers of the disease for up to two weeks (Longini *et al.,* 2007) others who show symptoms can die from dehydration within hours.

An infected agent spread *V. cholerae* through excretion of feces to the environment until he is treated and recovered. If the agent uses an open field latrine, the bacteria are accumulated in the field and can be transported by rain (See Appendix 1, Section 3.3.2). If the agent uses ventilated improved pit latrine (VIPL), the bacteria will be contained within it even during the rain. Unless there is seepage, the potential of VIPL to contaminate a water source is minimal. For simplicity, we did not consider contamination of water sources through seepage in this model unless it is exogenously introduced. A recovered agent will stay as recovered for some time before the agent becomes susceptible again. A susceptible and recovered agent may also spread *V. cholerae* through excretion of feces to the environment with a much lesser amount as set in the parameter.

4 Scenario Testing

Based on the current status of the model, we present results from two experiments that are focusing on the spread of cholera caused by two potential sources of water contamination in the refugee camps. We initialize the model with 50,000 refugees and run the simulation for one month. We assume that all the facilities (22 water points) provide sufficient water for the refugees. As the daily consumption of a person is 15 liters/day, we limit the capacity to each water point proportional to the agent numbers. Each water point holds a maximum of 15,000 liters of water and will refill at a rate of 15 liters/minute until it reach to the maximum level. The other parameters are set based on the default value as shown in Table 1. A refugee agent will fetch water from their nearest water point when there is a need. However, agents may be forced to go further if the nearest water point is very crowded or does not have sufficient water.

4.1 Contamination of Water Through Seepage

In this scenario, we consider the possibility of contamination of water sources through seepage. We explore the spatial distribution of cholera by infecting one of water facility (borehole) in the center of the Ifo camp exogenously. We assume that there is no rainfall in this experiment and run the model for one week.

Fig. 3 shows the summary from this scenario. In the early stages of the model run, about 1500 refugee agents who had ingested contaminated water were rapidly passed from susceptible to exposed stage. This exposure resulted in a spike in cholera epidemic as shown in Fig. 3B. However, as more agents prioritized their activity and received treatment, the cholera outbreak significantly decreased. As the model run progressed (between 2,500 and 12,500 minutes), cyclic incidents of cholera outbreak were observed but with very fewer number of cases as compared to the initial period. The number of agents who passed to the recovered stage progressively increased. This is mainly because the probability getting infected again for agents after receiving treatment was minimal due to their acquired immunity.

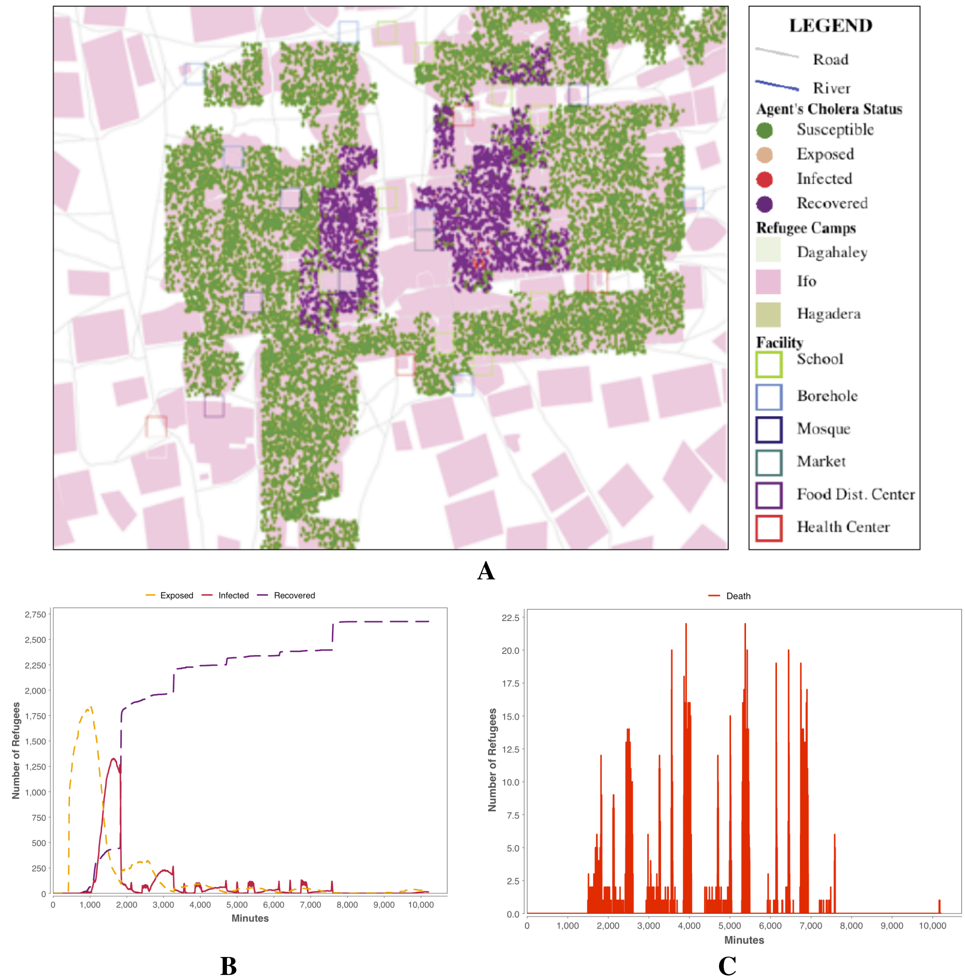

Fig. 3. The spread of cholera when there is contamination through seepage. **A:** The spatial distribution of agents. **B:** The number of exposed, infected and recovered individuals over time. **C:** The number of deaths over time.

Although agents were able to get treatment easily, there were cases of fatalities in periods where infection level is high as shown in Fig. 3C. However, the number of deaths is not linear when compared to the number of infected agents. Two factors may contribute for this variation. First, health facilities have limited capacity. Secondly, agents might of delayed their visit health facilities due to their priority of activities (i.e. queuing at the health facilities) and eventually loses their body resistance before receiving treatment.

The spatial pattern of cholera was localized around the contaminated water source as shown in Fig 3A. The spread of the disease grows radially as the simulation progresses from the contaminated water source. However, the spread of cholera fails to affect agents who are living far from the contaminated water source, except for few agents who visited their relatives or friends in the contaminated area.

4.2 Contamination of Water Through Runoff

In this scenario, we incorporated rainfall in the model. Hence, agents utilize water from two sources, either from facilities (boreholes or tanks) or rain, which is mainly accumulated in reservoirs (ditches or wells). We also alter the following parameter setting before the model is instantiated. That of setting equal value for the parameter for preference of water from borehole and surface runoff. In addition we introduce rain to occur in every three days for 20 successive minutes at a rate of 1 mm/minute. As in the previous scenario 1 we ran the simulation for one week and infected the same borehole.

As Fig. 4 shows the number of infected agents is significantly different compared to the first scenario. The pattern of the incidence of cholera outbreak is also different. Although the scenario showed a decreasing pattern over time, pronounced cyclic events of cholera outbreaks occurred due to the rainfall. For instance, the highest cholera outbreak occurred around 3200 minutes into the simulation as shown in Fig. 4, which is after the second day of rain. The spike can be explained due to more *V. cholerae* bacteria being in the system due to the first outbreak the day before. As more infected agents use open field latrines, the environment accumulates a larger number of *V. cholerae* bacteria. The bacterium is then being transported via surface runoff along the elevation gradient. As the bacterium is diluted in the water, it increases the level of contamination. Hence agents who have access to contaminated surface water are then being exposed to cholera.

In this scenario, the spatial distribution of cholera is spread out throughout the camps. Although the highest cluster of cases is found around the contaminated borehole, there are other clusters of cholera in other places as the bacteria spread through runoff as shown in Fig. 4A.

5 Summary and Outlook

Although the model is still under development, it yields striking results on the spread of cholera in refugee camps. The model shows that agents who are exposed to contaminated water are very susceptible to cholera infection. The agents are also some of the causes of the epidemic when they use open field latrines. Moreover, the model also showed that cholera cases were identified in a place where water sources were relatively safe. This was due to runoff caused by rain and as agents were travelling throughout the camps to visit their friends and relatives. Such individuals visiting areas where the cholera epidemic had occurred would then become infected and then carry it back to their homes. This signifies that although a humanitarian agency might able to contain the spread of cholera at the source by treating the contaminated water, it could be challenging to eradicate the spread of cholera as the bacteria could traveled far either by an infected person or through runoff.

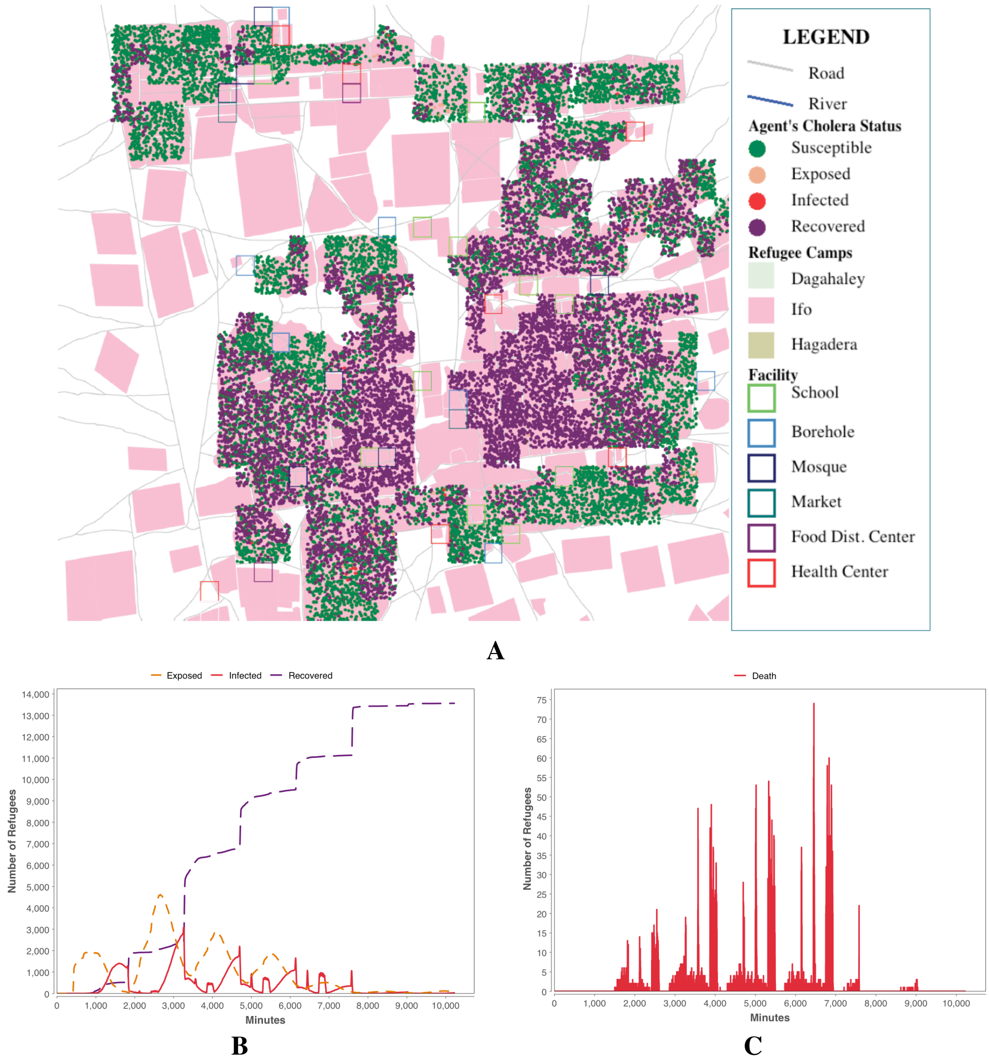

Fig. 4. The spread of cholera when there is contamination through runoff. **A:** The spatial distribution of agents. **B:** The number of exposed, infected and recovered individuals over time. **C:** The number of deaths over time.

This modeling effort also highlights the potential of agent-based modeling to explore the spread of cholera in humanitarian contexts. Agent-based modeling is helpful in exploring the intensity and spatial extent of the spread of cholera. It also showed the dynamics of the spread of cholera that emerges from the complex interaction of humans and their environment. As individuals make their own decision in their daily routine and interact with other individuals and/or the environment, they act as a victim or a cause for cholera epidemic.

Although this paper shows the potential of agent-based modeling in capturing complex humanitarian issues, it also worth mentioning to some of the challenges that we have faced in the modeling process. The main challenge was getting data of high enough resolution to depict the study area. In our case, although we were able to convert the pdf maps into shapefiles and conduct editing and digitizing to capture all the necessary attribute information, we could not get detailed socio-economic data at this resolution. We were relying on aggregate statistical data and relevant literatures to fill this gap. Another challenge on doing model at this resolution and time scale we have chosen which impacts the speed of the model. With respect to ongoing work, we are attempting to validate our model with actual events in the Dadaab refugee camps and improve the dilution of *V. cholerae* bacteria in surface runoff along with optimizing the codebase to improve the runtime of model. Once these have been done the model can then be potentially used to aid understanding of the spread of cholera and to help develop control strategies for when cholera breaks out.

Acknowledgments. This work was supported by the Center for Social Complexity at George Mason University and by the Office of Naval Research (ONR) under a Multi-disciplinary University Research Initiative (MURI) grant no. N00014-08-1-0921. The authors would like to acknowledge input from the Mason-HRAF Joint Project on Eastern Africa (MURI Team). The opinions, findings, and conclusions or recommendations expressed in this work are those of the authors and do not necessarily reflect the views of the sponsors.

References

1. **Anderson, J., Chaturvedi, A. and Cibulskis, M. (2006),** 'Modeling the Health of Refugee Camps: An Agent-Based Computational Approach', in Lee, D. and Nutter, B. (eds.), *Proceedings of the 19th IEEE Symposium on Computer-Based Medical Systems*, Salt Lake City, UT, pp. 641 - 645.
2. **Anderson, J., Chaturvedi, A. and Cibulskis, M. (2007),** 'Simulation Tools for Developing Policies for Complex Systems: Modeling the Health and Safety of Refugee Communities', *Health Care Management Science*, 10(4): 331-339.
3. **Augustijn-Beckers, E., Useya, J., Zurita-Milla, R. and Osei, F. (2011),** 'Simulation of Cholera Diffusion to Compare Transmission Mechanisms', *Proceedings of the 11th International Conference on Geocomputation*, University College London, London, England, pp. 39-42.
4. **Batty, M. and Torrens, P.M. (2005),** 'Modelling and Prediction in a Complex World', *Futures*, 37(7): 745-766.
5. **Bertuzzo, E., Casagrandi, R., Gatto, M., I., R.-I. and Rinaldo, A. (2010),** 'On Spatially Explicit Models of Cholera Epidemics', *Journal of the Royal Society Interface*, 7(43): 321-333.
6. **Capasso, V. and Paveri-Fontana, S.L. (1979),** 'A Mathematical Model for the 1973 Cholera Epidemic in the European Mediterranean Region', *Revue d'Epidemiologie et de Sante Publique*, 27(2): 121-132.
7. **CARE (2012),** Horn of Africa Food Crisis, Available at http://www.care.org/emergency/Horn-of-Africa-food-poverty-crisis-Dadaab-2011/index.asp [Accessed on July 1st, 2012].
8. **CGIAR Consortium for Spatial Information (2012),** CGIAR Consortium for Spatial Information, Available at http://www.cgiar-csi.org/ [Accessed on July 1st, 2012].

9. **Cioffi-Revilla, C., Gulden, T., Kennedy, W. and Coletti, M. (2011),** Agent Based Model for Analyzing Conflict, Disasters, and Humanitarian Crises in East Africa, *Working Paper, Mason-Yale Joint Project on Eastern Africa, Center for Social Complexity, Krasnow Institute for Advanced Study, George Mason University*. Available from the first author.
10. **Codeço, C.T. (2001),** 'Endemic and Epidemic Dynamics of Cholera: The Role of the Aquatic Reservoir', *BMC Infectious Diseases*, 1(1), Available at http://www.biomedcentral.com/1471-2334/1/1.
11. **Cronin, A.A., Shrestha, D., Cornier, N., Abdalla, F., Ezard, N. and Aramburu, C. (2008),** 'A Review of Water and Sanitation Provision in Refugee Camps in Association with Selected Health and Nutrition Indicators – The Need for Integrated Service Provision', *Journal of Water and Health*, 6(1): 1-13.
12. **Crooks, A.T. and Heppenstall, A. (2012),** 'Introduction to Agent-Based Modelling', in Heppenstall, A., Crooks, A.T., See, L.M. and Batty, M. (eds.), *Agent-based Models of Geographical Systems*, Springer, New York, NY, pp. 85-108.
13. **Epstein, J.M. (2009),** 'Modelling to Contain Pandemics', Nature, 460: 687.
14. **Epstein, J.M., Pankajakshan, R. and Hammond, R.A. (2011),** 'Combining Computational Fluid Dynamics and Agent-Based Modeling: A New Approach to Evacuation Planning', *PLoS ONE*, 6(5): e20139. doi:10.1371/journal.pone.0020139.
15. **Eubank, S., Guclu, H., Kumar, A.V.S., Marathe, M.V., Srinivasan, A., Toroczkai, Z. and Wang, N. (2004),** 'Modelling Disease Outbreaks in Realistic Urban Social Networks', *Nature*, 429: 180-184.
16. **Fisman, D. and Pandemic Influenza Outbreak Research Modelling Team (Pan-InfORM) (2009),** 'Modelling an influenza pandemic: a guide for the perplexed', *Canadian Medical Association Journal*, 181(3-4): 171-173.
17. **Franco, A.A., Fix, A.D., Prada, A., Paredes, E., Palomino, J.C., Wright, A.C., Johnson, J.A., McCarter, R., Guerra, H. and Morris, J.G. (1997),** 'Cholera in Lima, Peru, Correlates with Prior Isolation of Vibrio Cholerae from the Environment', *American Journal of Epidemiology*, 146(12): 1067-1075.
18. **Grimm, V. (2002),** 'Visual Debugging: A Way of Analyzing, Understanding, and Communicating Bottom-up Simulation Models in Ecology', *Natural Resource Modeling*, 15(1): 23 - 38.
19. **Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S., Huse, G., Huth, A., Jepsen, J., Jorgensen, C., Mooij, W., Muller, B., Pe'er, G., Piou, C., Railsback, S., Robbins, A., Robbins, M., Rossmanith, E., Ruger, N., Strand, E., Souissi, S., Stillman, R., Vabo, R., Visser, U. and Deangelis, D. (2006),** 'A Standard Protocol for Describing Individual-Based and Agent-Based Models', *Ecological Modelling*, 198(1-2): 115–126.
20. **Gulden, T., Cotla, C.R., Hailegiorgis, A. and Crooks, A.T. (2011a),** 'A Generic Vegetation Growth Sub-model in a Large Human/Environment Interaction Model of East Africa', *The Computational Social Science Society of America Conference* (2011), Santa Fe, NM.
21. **Gulden, T., Harrison, J.F. and Crooks, A.T. (2011b),** 'Modeling Cities and Displacement through an Agent-based Spatial Interaction Model', *The Computational Social Science Society of America Conference* (2011), Santa Fe, NM.
22. **Hartley, D.M., Morris, J.G. and Smtih, D.L. (2006),** 'Hyperinfectivity: A Critical Element in the Ability of V. Cholerae to Cause Epidemics?' *PLoS Medicine*, 3(1): 63-69, Available at doi:10.1371/journal.pmed.0030007.
23. **Johnson, R.T., Lampe, T.A. and Seichter, S. (2009),** 'Calibration of an Agent-based Simulation Model Depicting a Refugee Camp Scenario', in Rossetti, M.D., Hill, R.R., Johansson, B., Dunkin, A. and Ingalls, R.G. (eds.), *Proceedings of the 2009 Winter Simulation Conference*, Austin, TX, pp. 1778-1786.

24. **Lobitz, B., Beck, L., Huq, B., Wood, B., Fuchs, G., Faruque, S.M. and Colwell, R. (2000),** 'Climate and Infectious Diseases: Use of Remote Sensing for Detection of Vibrio Cholerae by Indirect Measurement', *Proceedings of the National Academic Sciences*, 97(4): 1438-1443.
25. **Longini, I.M., Nizam, A., Ali, M., Yunus, M. and Shenvi, N. (2007),** 'Controlling Endemic Cholera with Oral Vaccines', *PLoS Med*, 4(11): e336.
26. **Longley, P.A., Goodchild, M.F., Maguire, D.J. and Rhind, D.W. (2010),** *Geographical Information Systems and Science* (3rd Edition), John Wiley & Sons, New York, NY.
27. **Mniszewski, S.M., Del Valle, S.Y., Stroud, P.D., Riese, J.M. and Sydoriak, S.J. (2008),** 'EpiSimS Simulation of a Multi-Component Strategy for Pandemic Influenza', *Proceedings of the 2008 Spring Simulation Multiconference*, Ottawa, Canada, pp. 556-563
28. **Nelson, E.J., Harris, J.B., Morris, J.G., Calderwood, S.B. and Camilli, A. (2009),** 'Cholera Transmission: The Host, Pathogen and Bacteriophage Dynamic', *Nature Reviews Microbiology*, 7(10): 693-702.
29. **Oseia, F.B., Dukera, A.A., Augustijnb, E.W. and Steinb, A. (2010),** 'Spatial Dependency of Cholera Prevalence on Potential Cholera Reservoirs in an Urban Area, Kumasi, Ghana', *International Journal of Applied Earth Observation and Geoinformation*, 12(5): 331-339.
30. **Reiner, R.C., Kinga, A.A., Emch, M., Yunusd, M., Faruque, A.S.G. and Pascuala, M. (2012),** 'Highly Localized Sensitivity to Climate Forcing Drives Endemic Cholera in a Megacity', *Proceedings of the National Academy of Sciences*, 109(6): 2033-2036.
31. **Siddique, A.K. (1994),** 'Cholera Epidemic Among Rwandan Refugees: Experience of ICDDR,B in Goma, Zaire', *Glimpse*, 16(5): 3-4.
32. **Simoes, J. (2012),** 'An Agent-Based/Network Approach to Spatial Epidemics', in Heppenstall, A.J., Crooks, A.T., See, L.M. and Batty, M. (eds.), *Agent-based Models of Geographical Systems*, Springer, New York, NY, pp. 591-610.
33. **Snow, J. (1855),** On Mode of Communication of Cholera (2nd Edition), Churchill, London, England.
34. **Sphere (2004),** *The Sphere Project: Humanitarian Charter and Minimum Standards in Disaster Response*, Oxfam Publishing, Oxford, England.
35. **Swerdlow, D.L., Malenga, G., Begkoyian, G., Nyangulu, D., Toole, M.J., Waldman, R.J., Puhr, D.N. and Tauxe, R.V. (1997),** 'Epidemic Cholera Among Refugees in Malawi, Africa: Treatment and Transmission', *Epidemiology and Infection*, 118(3): 207-214.
36. **Toole, M.J. and Waldman, R.J. (1997),** 'The Public Health Aspects of Complex Emergencies and Refugee Situations', *Annual Review of Public Health*, 18(1): 283-312.
37. **Tuite, A.R., Tien, J., Eisenberg, M., Earn, D.J., Ma, J. and Fisman, D.N. (2011),** 'Cholera Epidemic in Haiti, 2010: Using a Transmission Model to Explain Spatial Spread of Disease and Identify Optimal Control Interventions', *Annals of Internal Medicine*, 154(9): 593-601.
38. **UNHCR (2006),** *Practical Guide to the Systematic Use of Standards and Indicators in UNHCR Operations* (2nd Edition), UNHCR, Geneva, Switzerland.
39. **UNHCR (2011a),** *Global Trends 2010*, United Nations High Commissioner for Refugees, Geneva, Switzerland, Available at http://www.unhcr.org/4dfa11499.html.
40. **UNHCR (2011b),** Health And Nutrition Sector Update Dadaab Refugee Camps 15th to 21st October 2011, Epidemiological Week 42 Update No. 14, Nairobi, Kenya.
41. **UNHCR (2011c),** Kenya - Lagdera District Refugee Camp Overview, Available at http://data.unhcr.org/horn-of-africa/regional.php [Accessed on December 21st, 2011].
42. **UNHCR (2012),** Dadaab - World's Biggest Refugee Camp 20 Years Old, Available at http://www.unhcr.org/4f439dbb9.html [Accessed on July 1st, 2012].
43. **UNITAR (2012),** Horn of Africa Data, Available at http://www.unitar.org/unosat/horn-africa-data [Accessed on July 1st, 2012].
44. **Werker, E. (2007),** 'Refugee Camp Economies', *Journal of Refugee Studies*, 20(3): 461-480.

Appendix 1: Overview, Design concepts, and Details Protocol

In the following appendix, the description of the model will be given based on the Overview, Design concepts, and Details (ODD) ODD protocol by Grimm, *et al.* (2006).

1 Overview

1.1 Purpose

The purpose of the model is to explore the spread of cholera, which is caused by the interaction of human (host) with their environment. As with all models a number of simplifying assumptions have been made to covert the complexities of reality into something which can be modeled (Batty and Torrens, 2005) which we detail below.

1.2 State Variable and Scales

The model focuses mainly on the spread of cholera in the Dadaab refugee camps. We have tried to stylize agents' behavior by incorporating their daily routine that may happen within the Dadaab refugee camps so that we can have better understanding of the dynamics of cholera transmission. We represent the environment using geo-referenced spatial data. Fig. 5 shows the UML diagram of the model.

The main agent in the model is the refugee agent who represents an individual refugee who lives in the Dadaab refugee camps. A refugee agent has family and fixed home location. Agents of the same family cooperate and share resources.

Agents are instantiated with different attributes that contribute to their heterogeneity. Agents differ in their personal characteristics (e.g. age, sex and education level), social ties (e.g. number of family members and friends), their body immunity type (symptomatic and asymptomatic), and goals and priorities. Behaviorally, agents are mobile and purpose-oriented. They determine a specific activity (goal) at a given time, depending on their priorities, and move towards it to fulfill their satisfaction.

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

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

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

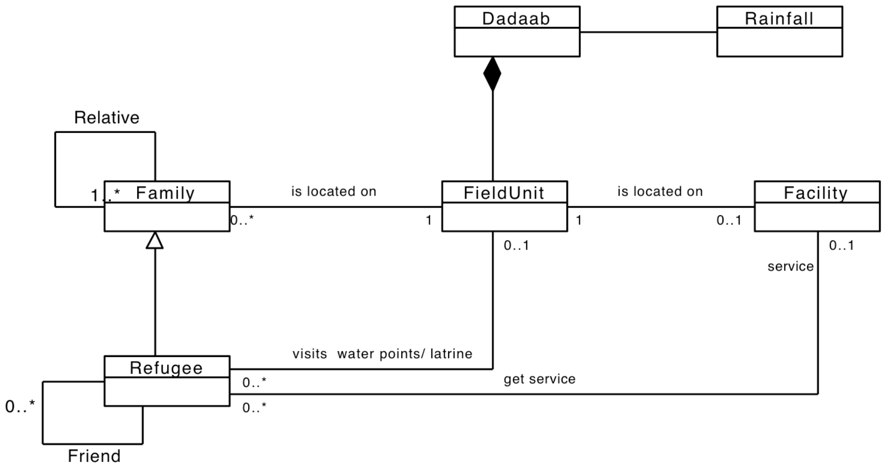

Fig. 5. High level representation of the model using UML diagram.

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

The infrastructure represents the road networks of the Dadaab refugee camps. All types of roads (primary roads, secondary roads, feeder roads, and trails) are represented as the same type. There is no cost or preference on the types of the road the agents chose; however, roads have a capacity that constrains the flow of traffic (agents) in a given time. The level of road capacity is represented as ‘crowd parameter’. A road can only be occupied by a given number of agents as it is set in the crowd parameter at a given time. If the road ahead is crowded, agents should stay where they are and wait until the road is cleared[2].

The elevation dataset represents the topography of the Dadaab refugee camps. It has the same spatial resolution and extent with the other datasets and mainly used as an input for modeling the runoff of water.

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

1.3 Process Overview and Scheduling

In each time step, each agent makes decision to stay where they are or move to their goal based on their priorities (e.g. needing food or water, to attend school). Agent movement from home to goal determined by the time and their success at the goal as shown in Fig. 6.

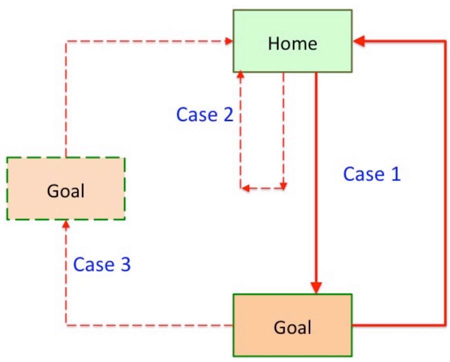

Fig. 6. Agent's movement options between his home and his goal.

Each agent moves towards his goal by selecting the nearest road path. Since an agent can be constrained by time, the success of the agent to reach its goal is not guaranteed. If an agent is successful and reaches its goal, the agent might return to home if he accomplishes his activity or stays for a while until he is able to accomplish his activity before he returns to his home (case 1). If the agent is too late to accomplish his goal, the agent will return home before reaching his goal (case 2). In some instants, an agent might consider choosing another goal before going back to their home (case 3). In the current version of the model, this will only happen when

[2] In future versions of the model, agents will be given more dynamic route planning capabilities.

the agent fails to collect enough water from their nearest water point (e.g. the water point has run out of water) and are therefore forced to search for other water points to achieve its goal.

2 Design Concepts

2.1 Observation

The visualization window of the model as shown in Fig. 2 portrays the extent and dynamics of cholera. We monitor the progress of cholera as agents portray different colors depending on their health status. At the global level, we monitor the following statistics: total number each activities, total number of susceptible agents, total number of exposed agents, total number of infected agents, total number of recovered agents, and total number of deaths. We also monitor the rainfall and surface runoff on the environment.

2.2 Interactions

The dynamics of the model are driven by the interaction of agents with their environment. Agents of the same family share resources (e.g. water). Agents interact with other agents who are their relatives or friends. Agents have a fixed number of relatives, which is given to them at the initialization of the model; however, they select friends randomly at runtime.

2.3 Sensing

Agents are assumed to know the nearest facilities, their relatives and friends locations. They also know how to navigate along the road network towards their goal. The environment reacts to rainfall and creates surface runoff via the elevation gradient.

2.4 Emergence

We anticipate the spatial distribution of cholera to be emergent phenomena. Although users can infect a specific number of boreholes to observe the spread of cholera, the spatial distribution and spread of cholera from one camp to another emerges due to the complex interaction of agents with their environment.

2.5 Stochasticity

Some processes in the models are stochastic. Agent selections of some activities or location of friends are drawn from normal distribution. Assignment of some values to agents (e.g. to be a symptomatic or asymptomatic agent) at the initialization of the model is also carried out stochastically based on a parameter set by the user.

3 Details

3.1 Initialization

The initialization of the model relies on socioeconomic and spatial data of the study area. Most of the parameter values were calibrated based on relevant literatures. The model gives flexibility to users to run different experiments by changing some of the default values. For instance, users can experiment by increasing or decreasing the number of agent in the simulation. All the default values and the parameters are summarized in Table 1.

3.2 Inputs

The input spatial dataset were generated from publicly available data sources. The camps information (camp boundaries, houses, facilities, and infrastructure) were processed from UNHCR (2011b) PDF maps and Google Earth (KLM format) from UNITAR, (2012). The PDF maps of each camp were converted into a ESRI vector shapefile using PDF Converter. The converted data was geo-referenced and edited in ArcGIS. The elevation dataset were processed from 90m meters Digital Elevation Model (DEM), which is provided by CGIAR Consortium for Spatial Information (2012) GeoPortal. The DEM was interpolated using ordinary Kriging to fit the model resolution.

3.3 Submodels

3.3.1 Goal selection

Agents determine their current activity based on their personal attributes (e.g. age, sex), and their current water need. They also consider the time and distance into consideration when they make their goal choices. Agents who are less than 5 years old are always staying at home. They do not engage in any activities unless they are sick and need to visit the health facilities. They utilize water from the family reserve and use the home latrine. The other age groups can engage in any of the activities depending on their needs and abilities. Below, we will describe the characteristics of each activity. The assignment of all facilities location is based on proximity. In the model, agents give the highest priority to the nearest facility when choosing between two facilities of the same kind.

Table 1 Input parameters and variables.

Parameter	Value	Reference
Agent		
Initial number of agents	50,000	
Daily water consumption	15 liter/day	UNHCR (2006); Sphere (2004); CARE(2012).
Dehydration rate	0.003 liter/ minute	
% of symptomatic agents	99 %	
Max number of relatives	4 families	
% of asymptomatic agent	1%	
Max body resistance	1.0	
Health depreciation rate	0.004/ minute	
Mortality	Up to 50 % of infected (untreated)	
	Up to 1% of infected (treated)	Nelson et al (2009).
Min number of vibrios to cause cholera infection	1000/ml	Franco et al (1997) ; Nelson et al (2009)
First symptoms after infection revealed	12-72 hours	Nelson et al (2009)
Infected person fluid loss	1000ml/hr	Nelson et al (2009); Codeco (2001)
Vibrios per gram of stool of infected person	107 -109/ml	Nelson et al (2009) ; Franco et al (1997)
Vibrios per gram of stool of uninfected person	102 to 105/ml	Codeço (2001)
Symptoms continuation period	1-2 weeks	Nelson et al (2009)
Facilities		
Health facilities capacity	500 patient /day	
Borehole Maximum capacity	15,000 liter/day	
Borehole recharge rate	10 liter/minute	
Rainfall		
Rainfall amount	0.3 mm/ minute	
Duration	20 minute/ day	
Frequency	Every 1 week	

3.3.1.1 Goal selection - School

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

3.3.1.2 Goal selection - Mosque

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

3.3.1.3 Goal selection - Market

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

3.3.1.4 Goal selection - Food Distribution Center

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

3.3.1.5 Goal selection - Health Center

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

3.3.1.6 Goal selection - Water Points

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

In reality, neither all family members are tasked to collect water or utilize the water independently. Rather water is collected by any member of the family and is equally used by each family member. This notion of utilization of water by family members is represented in the model. Agents fetch water from water source and accumulate the water in the family bucket and utilize it from there. Agent should fulfill their water requirement each time they visit a water point. The daily consumption of each agent is 15 liters per day (CARE, 2012). This amount includes all possible uses of water: drinking, cooking and cleaning. We consider two types of water utilization: water for drinking and water for all other use. This distinction could introduce complexity in the behavior of the agents and help us to explore the refugees' exposure to contaminated water and its consequence. There is a notion of dehydration in the model. Each time step, agents check both their body water level and their family water levels to make decision whether to fetch water or not. Agents fetch water from the nearest water sources and utilize the water. Agent can fetch up to 25 liters in a single visit depending on the availability of water from the sources. If the source is dried up or very crowded, agents will visit other nearby water points.

3.3.1.7 Goal selection - Latrine

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

3.3.2 Hydrology

Within our model we use rain as a proxy for climatic events as previous research has shown that there are strong correlations between seasons and outbreaks of cholera (see Reiner *et al.,* 2012). Secondly water is one of the main methods that cholera is transmitted (Codeço, 2001). In this model, we use rainfall both as source of water for the refugee as well as a carrier of pollutants (disposal and feces). We utilize elevation surface data (DEM) to model rainfall in the model. We apply simple hydrologic model that only consider elevation gradient to model surface runoff. Rainfall flows downhill according to the elevation gradient. As the water flows from uphill to downhill, it carries pollutants. The concentration of pollutants in the water depends on the amount of pollutant per volume of water in the field unit (parcel).

The dynamics of surface runoff is designed as follows. When there is rain, all cells get equal amount of rainfall. At each time step, each cell will check if it has water to flow to its Moore neighbors. If the cell has water and the neighboring cell is a sink (i.e. at a lower elevation), it will give the water until it fills the sink depending on the elevation and water gradient. If the volume of water is less than the sink, all the water will flow to the sink. However, if the volume of water is greater than the sink, the water will flow until the two cells reach to equal level. As rainfall flows from cell to cell, the concentration of pollutant in the water varies accordingly.

3.3.3 Cholera SEIR Model

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

All the refugee agents are considered as susceptible hosts as shown in Fig. 7. The infectious dose of *V. cholerae* in humans varies greatly depending on the bacterial strain and the host. In many cases, a bacterial cell concentration of 103/ml of water is necessary to infect the host (Nelson *et al.,* 2009). We assume that a susceptible agent who ingests contaminated water with a bacterial cell concentration of 103/ml of water or above will be become exposed to the cholera disease. We distinguished two types of infectious agents: symptomatic and asymptomatic. A symptomatic agent shows the symptom between 8 to 17 hours (Nelson *et al.,* 2009[3]). Infants show the symptom more quickly than adults. However, if the agent is asymptomatic, the agent will pass the infection phase to the recovery stage without showing any sign of symptom. In reality as most infected people have no or mild symptoms but still are carriers of the disease for up to two weeks (Longini *et al.,* 2007) others who show symptoms can die from dehydration within hours.

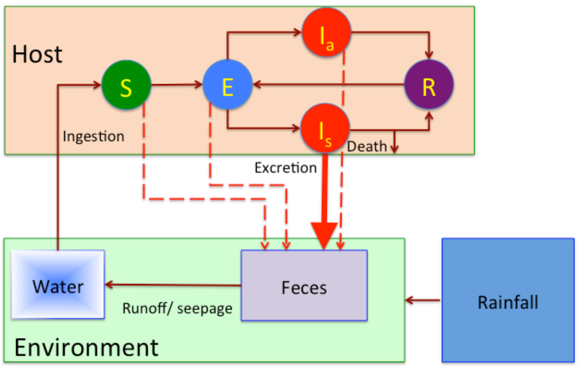

Fig. 7. Cholera transmission through the interaction of host and the environment. The progress of cholera transmission is represented as SEIR model. S = susceptible, E=Exposed, Ia= Infected (asymptomatic), Is= Infected (symptomatic).

An infected agent spread *V. cholerae* through excretion of feces to the environment until he is treated and recovered. According to Franco *et al*. (1997), an infected individual can spread 109/ml of *V. cholerae* through excretion of feces to the environment and the bacteria can survive in the environment for long period of time (Franco *et al.,* 1997). Hence in the model, we assume that any infectious agent could

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

spread a 109/ml of *V. cholerae* through excretion. If the latrine is of OFL type, the bacteria are accumulated in the field and can be transported by rain. If the latrine is of VIP type, the bacteria will be contained in the latrine even during the rain. Unless there is seepage, the potential of VIPL to contaminate a water source is very minimum. For simplicity, we did not consider contamination of water sources through seepage in this model unless it is exogenously introduced.

A recovered agent will stay as recovered for some time before the agent becomes susceptible again. The rate of transition from recovery to susceptible depends on the value set in the parameter. A susceptible and recovered agent may also spread *V. cholerae* through excretion of feces to the environment with a much lesser amount as set in the parameter.