ABSTRACT

Title of dissertation: ENVIRONMENTAL CHOLERA TRANSMISSION

TRIGGERS IN DELTA BENGAL

Wessam Mahmoud Elnemr, Doctor of Philosophy, 2014

Dissertation directed by: Dr. Rita Colwell, Professor, Marine and Estuarine

Environmental Sciences (MEES), Molecular Biology and

Biotechnology Area

Prediction of water-borne diseases is a critical aspect especially for developing countries. The current study focuses on cholera since it is considered to be a continuous public health threat. *Vibrio cholerae,* causative agent of cholera, is an autochthonous bacterial inhabitant in the aquatic environment and it is highly unlikely that cholera will ever be fully eradicated. Consequently, to reduce the disease burden, enhanced cholera prediction models that include several months' lead-time are still needed to further the development of effective mitigation and intervention strategies. Both regional and large-scale environmental conditions can aid in understanding and predicting how and when outbreaks may occur. The overall goal of the research reported here was to develop a quantitative cholera prediction model with high quality, using regional and remote-sensing data from endemic and epidemic regions, respectively, in the Bengal Delta. This research involved four separate supporting objectives: 1) Determination of the role of environmental factors associated with the seasonality and modulating dynamics in a cholera outbreak; 2) Development of a physically plausible hypothesis of how local

environmental factors modulate cholera outbreak dynamics; 3) Identification of the major environmental controls triggers sporadic cholera outbreaks in epidemic regions ; 4) Construction of an accurate model for the Bengal Delta simulating and predicting the two transmission routes of cholera (primary and secondary). The modeling results show that, for a high quality model $\geq$ 70% Pseudo-R Square, Bengal Delta cholera in coastal regions is characterized by a single spring peak, whereas Bengal Delta cholera in inland regions occurs in bimodal peaks, with distinct hydroclimatological explanations for the geographical differences. Results confirm that spring season cholera is associated with coastal seawater intrusion, and fall cholera outbreaks are driven by floods related to the monsoon. This is the first study that demonstrates the relationship between *in situ* environmental conditions with regard to cholera outbreaks. Furthermore, results from remote-sensing data show that ambient temperature followed by high rainfall periods are the main triggers of cholera outbreaks in epidemic regions. These findings provide important steps and contributions toward development of environmental factor-based predictive models for cholera outbreaks in the Bengal Delta region.

ENVIRONMENTAL CHOLERA TRANSMISSION TRIGGERS IN DELTA

BENGAL

by

Wessam Mahmoud Elnemr

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2014

Advisory Committee:
Dr. Rita Colwell, Chair
Dr. Anwar Huq
Dr. Fawzy Hashem
Dr. Antar Jutla
Dr. Pat Millner
Dr. Najib Al-Sayed (Dean's Representative)

ACKNOWELDGMENT

First and foremost, all praise and thanks are due to ALLAH the Almighty. He is the one who blessed me with the ability to achieve this success. During hard times in the course of my PhD, I always return to him and ask him in my prayers, and he always answers my prayers and supplications.

I would like to express my deepest gratitude and appreciation to my advisor, Dr. Rita Colwell, for her excellent guidance, caring, patience, and for providing me with all facilities connections of unlimited help that led to the successful completion of this dissertation. Dr. Colwell has played many significant roles during my stay in the US. First, she played an essential and critical role in the successful completion of my dissertation. Second, she supported me at the personal level during my personal hardships that made my life in the US more enjoyable. I greatly appreciate her excellent assistance, guidance and her spiritual support for me and my family including my mother during my PhD study. Dr. Colwell has imprinted values in me that will stay with me for the rest of my life. Dr. Colwell has gone above the call of duty on my behalf on more than one occasion without hesitation. Patience, foresight, humility, and respect; these are just a few to name.

I would like to express my cordial thanks to my Co-advisor, Dr. Anwar Huq, for all his enormous help, support, and encouragements during my PhD journey. His continuous advice, help and encouragement have always been shedding light in my way to the PhD. Dr. Huq was the first to introduce me to the world of modeling in order to accommodate my personal family situation. His office, lab, and warm welcome were

always available to me. I would like to deeply thank him for their understanding of my family situation and for caring and believing in me.

My other Co-Advisor, Dr. Fawzy Hashem has also played an integral role in the successful completion of this dissertation, and I would like to express my sincerest thanks to him. Dr. Hashem has gone above the call of duty on my behalf on more than one occasion without any hesitations. He has never given up on me even at times when I wanted to give up; he shared words and actions of encouragement and support. I wholeheartedly believe this dissertation would not have been completed without the my advisor and Co-advisors; Dr. Colwell, Dr. Huq and Dr. Hashem.

Very special thanks are also due to Dr. Antar Jutla. Without Dr. Jutla's help and teaching me about modeling and cholera, there is no doubt this dissertation would not have been successfully completed. Many times Dr. Jutla who has gone above the call of duty to make the necessary arrangements to get the research completed. I also would like to thank Dr. Pat Millner for serving on my committee. The guidance and support she provided has been instrumental to my success, and I greatly appreciate her help. Having Dr. Jutla and Dr. Millner in my committee was a blessing and an honor for me.

I would also like to recognize and thank the many people who have helped me in various ways with collaborative efforts, logistical support, or guidance on the research that was conducted. These people include Dr. Azhar Nizam at Rollins School of Public Health, Emory University, Atlanta, Dr. Guillaume Constantin de Magny at Institut de Recherche pour le Développement, Montpellier, France, Dr. Nur Hasan at Maryland Pathogen Research Institute, UMD). I would also like to acknowledge my friends

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction and Context

1.1 Cholera dynamics:

Cholera: Cholera, an acute water-borne diarrheal disease, remains a constant public health threat in the developing countries (World Health Organization (Morens *et al*., 2004; Ali *et al*., 2012; WHO 2012 Bulletin). The ongoing seventh pandemic of cholera, which started in 1960s, has been reported in over 50 countries affecting over seven million people (Gleick, 2008). The disease continues to be a serious threat in many tropical regions of the world, specifically in coastal areas of South Asia, Africa, and Latin America (Figure 1.1).

Cholera has been a subject of major research interest, not only for its high incidence rate in the developing countries, but also because of inherent complexities and interactions associated with humans and the environment and as a re-emerging disease (see Table 1.1, Committee on Emerging Microbial Threats to Health, 1992 and 2003). The causative agent of cholera, *Vibrio cholerae,* can survive and proliferate in distinctively different environments: a) the micro-environment of the human intestine as a pathogenic bacterium; and b) the macro-environment of coastal and estuarine waters as both pathogenic and non-pathogenic bacteria. These bacteria are autochthonous to riverine, estuarine, and coastal waters, and they live in association with phyto- and zooplankton, algae, and crustaceans (Colwell, 1996; Huq, and Colwell1996.; Stine *et al*., 2008; Hasan *et al*., 2012). Copepods, amphipods, and other crustaceans are dominant among zooplankton populations that have a chitinous exoskeleton, the most abundant polysaccharide in the marine environment (Tamplin *et al*., 1977; Huq *et al*., 1983;

Colwell *et al*., 1981). Large populations of *Vibrio* spp., including the pathogenic species of *V. cholera*, are harbored in the gut and attached to the copepod exoskeleton (Colwell and Huq, 1999; Tamplin *et al.,* 1990; Huq *et al*., 1983; Colwell *et al*., 1981). *Vibrio* spp. play an essential role in recycling the chitinous insoluble polysaccharide in the aquatic ecosystem (Yu *et al.,* 1991). Plankton is widely reported as a natural reservoir of *V. cholerae* (Colwell, 1996). On the other hand, phages may play a role in cholera epidemics, specifically the CTX phage-encoded cholera toxin and the toxin-coregulated pilus (TCP) genes associated with attachment. Lateral transfer of these genes leads to new strains containing virulence gene combinations (Waldor and Mekalanos 1996). When environmental conditions are unfavorable for active growth and cell division, the bacterium, including virulent strains of *V. cholerae,* has a selective advantage by entering into a dormant stage, called the viable but nonculturable (VBNC) state, (Colwell and Huq 1994; McDougald *et al*., 1998; Kaneko and Colwell, 1975).
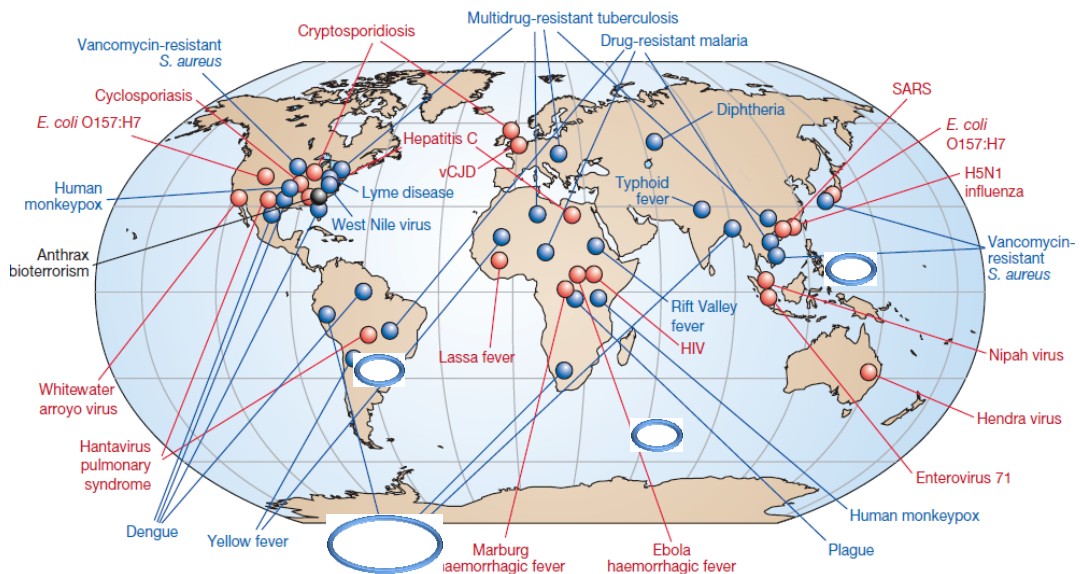
Figure 1.1 Global examples of emerging and re-emerging infectious diseases. Red represents newly emerging diseases; blue, re-emerging/resurging diseases; black, a 'deliberately emerging' disease. (Fauci, 2001)

Table1.1 Factors involved in the emergence of infectious diseases Committee on Emerging Microbial Threats to Health 1992 and 2003 (David *et al*., 2004)	
• Microbial adaptation and change	• International travel and commerce
• Human susceptibility to infection	• Technology and industry
• Climate and weather	• Breakdown of public health measures
• Changing ecosystems	• Poverty and social inequality
• Human demographics and behavior	• War and famine
• Economic development and land use	• Lack of political will
	• Intent to harm

Historically, cholera incidence has been linked to environmental and climate variables. Survival of *V. cholerae* as an aquatic bacterium is dependent on the physical characteristics of the water including alkalinity, salinity, and iron concentration, which influence virulence gene expression, *e. g.,* genes that regulate production of cholera toxin responsible for watery diarrhea (Miller *et al*., 1982 and Islam *et al*., 2004). In general, *V. cholerae* does not grow at temperatures below 15°C. However, aquatic water bodies with temperatures consistently above 20°C and salinities of 0.5–3.0% may harbor *V. cholerae* in endemic areas (Stephens *et al.,* 1998, Louis *et al*., 2003). In the early 19th century, the Bengal Delta was considered to be the native homeland of cholera (Pollitizer, WHO, 1959). Bangladesh is one example of the coastal regions of Asia affected by the pandemics of cholera, including the current seventh pandemic. Although cholera has been reported in Africa, Australia, Europe, and in the Americas (Figure 2), Bangladesh has one of the highest rates of cholera in the world, with an incidence rate of about 2.0 (range, 0.10 - 4.0) cases per 1000 people in endemic areas, and the mortality rate is about 6.3 deaths per 100,000 people at risk (*Bulletin of the World Health Organization* WHO 2012). There are two unique peaks of cholera epidemics in Bangladesh, in spring (minor peak: January-April/May) and in fall (major peak: September-November)

Figure 1.2 Countries affected by the Seventh Pandemic of Cholera (compiled from World Health Organization (WHO), Center for Disease Control and Prevention (CDC), and various news sources, Countries shaded have reported cholera outbreaks)

Akanda *et al.,* (2009) reported on the hydro-climatic factors affecting the outbreak of cholera in the Bengal Delta and proposed that intrusion of coastal water with low river discharge and flooding caused by high river discharge are the main forces responsible for spring and autumn outbreaks in the Bay of Bengal (BoB). In Bangladesh, various factors are known to facilitate spreading of the disease, including the characteristic heavy seasonal rainfall (Hashizume *et al*., 2008), seasonal flooding (Akanda *et al*., 2011 and Carrel *et al*., 2010), high population density concentrated near aquatic environments, and a high poverty level (Ali *et al*., 2002). Tauxe *et al.,* (1994) proposed that cholera outbreaks usually occur when the temperature increases.

Neilsen (1994) showed copepod production has two seasonal peaks. One peak is in February-April and the other is in August-September. These two peaks are relevant to earlier findings on the survival and multiplication of *V. cholerae* in the environment (Huq *et al.*, 1983 and 1984). Historically, cholera has been associated with a range of environmental and climate variables, including sea surface temperature (SST) (Lobitz *et al*., 2000; de Magny *et al*., 2008), precipitation (Pascual *et al*., 2002; Hashizume *et al*.,

2008), floods (Koelle *et al*., 2005), coastal salinity (Miller *et al*., 1982), peak river level (Schwartz *et al*., 2006), sea surface height (Lobitz *et al*., 2000) and fecal contamination (Islam *et al*., 2006). These studies did not focus on quantifying cholera incidence in Bangladesh and relating the seasonal hydroclimatological processes with biannual peaks of cholera.

1.2 Dynamics of cholera space in epidemic and endemic regions

A review of the literature indicates that the global understanding of cholera is mostly based on the Bengal Delta. In addition, most of the studies are focused on large-scale environmental conditions either measured by remote sensing or other large-scale measuring methods that ignore regional processes of the water bodies. These studies linked historical incidence of cholera in the Bengal Delta with environmental and climate variables, such
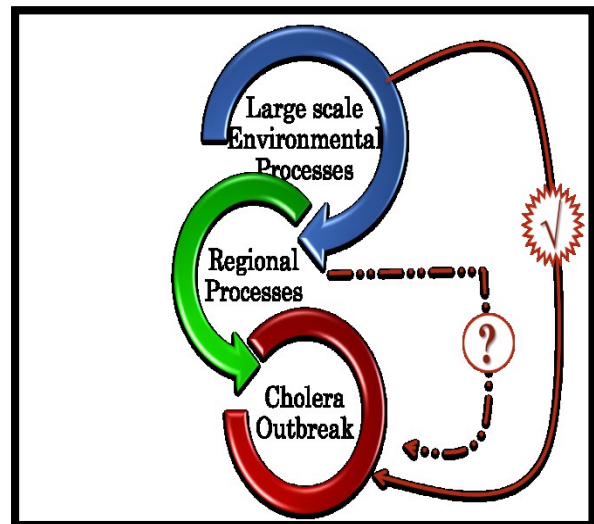
Figure 1.3 Relationship between large-scale environmental factors and cholera outbreaks: the missing outbreak link scheme for understanding cholera and the environment condition

as precipitation (*Hashizume et al. 2008*), floods (*Koelle et al., 2005*), river level (*Emch et al., 2008*), sea surface temperature (*Lobitz et al., 2000*), coastal salinity (*Miller et al., 1982*), dissolved organic matter (*Worden et al., 2006*), and fecal contamination (Islam *et al., 2006*). Remote sensing measurement of large-scale climate variables provide a limited understanding of what it can offer and explain as the effect of coastal and

hydrological connections with cholera incidence, using measurement of chlorophyll based on phytoplankton abundance. The coastal zone of the BoB shows high phytoplankton concentrations as well as high chlorophyll variability when compared with values further from the coast (Jutla *et al*., 2009). The effect of inter- and intra-annual variability of chlorophyll in coastal regions and the space-time variability of phytoplankton as an indication of subsequent zooplankton on cholera dynamics are still limited and incomplete.

There is no doubt that large-scale geophysical variables such as SST, precipitation, and coastal chlorophyll (measured by satellite remote sensing) are often related to both local environmental factors in ponds as well as the BOB, and conditions favorable for bacterial growth. Local environmental factors, such as biological activity in the ponds from where the bulk of the population in endemic regions derives water for daily usage, are often neglected or oversimplified in disease transmission dynamics. *How do local environmental factors affect cholera dynamics? Do local environmental factors have the same effect as large-scale environmental factors on cholera dynamics? What are the main local environmental drivers of a cholera outbreak?*

The continuing dilemma for both understanding (simulation) and prediction of cholera outbreaks derives from the following factors: 1) *V. cholerae* is naturally occurring in riverine, estuarine and coastal waters; 2) the difference in cholera disease between endemic and epidemic regions; and 3) emergence of new biotypes that can increase the incidence (or intensity) of disease (Morens *et al*., 2004). Simulation models should be transparent in quality for improved understanding of the main drivers of cholera incidence and spread. Examples of studies focused on models that lack the

essence of simulation models (transparency) are as follows: 1) Identifying the dominant zooplankton groups (rotifers and cladocerans) associated with detection of *V. cholerae* without correlating the data to its specific ecological drivers (Magny *et al.,* 2011); 2) Identifying the ecological drivers, e.g., chlorophyll, without estimating the plausible hypothesis behind the prediction of seasonal double peak cholera outbreaks for the Bengal Delta (Magny *et al.,* 2008); 3) Explaining the seasonal double peaks of cholera outbreaks in the Bengal Delta through the role of rivers, without taking into account large-scale processes such as river discharge (Akanda *et al.,* 2009); and 4) Lack of identification differentiating between the main environmental drivers in endemic and epidemic regions.

A major question remaining unanswered about cholera in epidemic or non-endemic regions is: *What are the major environmental processes that may contribute to severe outbreaks of the disease.* In this study, we seek to determine the main ecological drivers of the cholera outbreak in the Bengal Delta and construct a plausible hypothesis to explain the seasonal dual peaks of cholera in endemic regions and the main environmental driver of cholera outbreaks in epidemic regions.

2 Aims and Objectives:

2.1 Aims

Four key observations were the motivation for developing and exploring cholera simulation and prediction models for cholera outbreaks in both endemic and epidemic regions. These observations include coastal ecology and use of recent local environmental factors, as well as remote-sensing data specifically: 1) Large-scale geophysical processes such as SST and coastal chlorophyll, have been implicated with

cholera outbreaks in several parts of the globe (Huq *et al*., 2005; Magny *et al.,* 2008); 2) Local environmental factors, such as biological activity in ponds from where the bulk of the population in endemic regions derives water for daily use are often neglected or oversimplified in disease outbreak dynamics (Emch *et al*., 2008); 3) Appearance of new cholera cases in non-endemic regions like New Delhi; and 4) Rapid spread of cholera among human populations in some regions (WHO 2012). Establishing a connection between large- and local-scale environmental conditions enhanced our thinking toward identifying key environmental parameters driving cholera outbreaks. Further, developing a good quality model which includes the main environmental parameters associated with the emergence of cholera disease in epidemic/non-endemic regions is needed. Finally, simulation and prediction models that identify the major environmental cholera drivers in both endemic and epidemic regions would be extremely useful in developing strategies for preventing the disease.

2.2 Objectives:

The overall goal in the proposed research was to develop cholera simulation/prediction models that can simulate the dual peaks of the cholera outbreak in the Bengal Delta and provide insight on how a non-endemic region becomes endemic. To achieve this main goal, investigation was carried out based on the following research objectives:

a. Determination of the role of environmental factors associated with seasonality and modulating the dynamics of a cholera outbreak.

b. Identification and development of a physically plausible hypothesis of how local environmental factors modulate cholera outbreak dynamics.

c. Identification of the major environmental controls triggers sporadic cholera outbreaks in epidemic regions.

d. Constructing a new model that accounts for both primary and secondary transmission routes

This approach was employed to gain a comprehensive understanding that would allow identification of the major environmental conditions, either local or large scale, in both endemic and non-endemic regions. Furthermore, this approach provided a means to use the data to develop a good quality simulation model, for eventual use in predicting a cholera outbreak.

Figure (1.4) illustrates the schematic outline and progression of my research. How local environmental factors can be related to large-scale environmental factors derived by remote sensing and their effect on shaping cholera outbreaks in endemic regions is discussed in Chapter 2. That chapter will also include an up-to-date literature review about cholera and its relationship with hydrological processes. My approach in chapter 2 starts with a straightforward objective: The development of a simulation model for two distinct endemic regions with consideration of their unique settings. Such a premise will allow exploration of a transformational approach to protect vulnerable and resource-limited regions against cholera. This was a unique opportunity to use the three year data set that includes local environmental parameters collected directly at the pond site, extremely valuable to determine their relationship to cholera incidence. Chapter 3 deals with identification of major environmental factors affecting cholera outbreaks in epidemic regions and provides an explanation of how cholera can emerge in non-endemic

areas. In this chapter, another data set was used, employing large-scale environmental conditions measured by remote sensing of a non-endemic region.

	What do we know about cholera and hydrology?	Chapter 2[a]

Objective 1	What is the role of environmental factors in creating seasonality and modulating the cholera outbreak dynamics in a cholera outbreak?	Chapter 2[a]

Objective 2	What is the physically plausible hypothesis that can explain how local environmental factors modulate cholera outbreaks dynamics?	Chapter 2[a]

Objective 3	What is the major environmental control responsible for turning a non-infected area to an endemic one?	Chapter 3[b]

Objective 4	Building a model that accounts for both primary and secondary transmission routes	Chapter 3[b]

Figure 1.4 Research Pathway for cholera outbreak simulation and prediction

[a]**Elnemr W. M.**, Antarpreet S Jutla, Guillaume Constantin de Magny, Nur A. Hasan, Munirul Alam, R. Bradley Sack, Anwar Huq, Fawzy Hashem and Rita R. Colwell: Environmental Parameters related to Endemic Cholera (submitted).

[b]**Elnemr W. M.**, Antarpreet S Jutla, Anwar Huq, Fawzy Hashem, and Rita R. Colwell: Role of elevated Land surface temperature in altering cholera incidence dynamics in New Delhi, India (To be submitted)

Chapter 2

Local Environmental Factors and Endemic Cholera

2.1 Abstract:

Cholera remains a major public health threat. Since *Vibrio cholerae*, the causative agent of the disease, is autochthonous to riverine, estuarine, and coastal waters, it is unlikely cholera bacteria can be eradicated from its natural habitat. However, prediction of impending cholera outbreaks, in conjunction with effective public health measures, can reduce incidence and/or intensity of cholera outbreaks. Understanding environmental controls of growth and proliferation of the cholera bacteria is an essential first step in developing warning systems for the disease. Large scale geophysical processes, such as river discharge, sea surface temperature, and coastal chlorophyll, have been implicated with the occurrence of cholera. Local environmental factors, such as biological activity of ponds from which communities in cholera endemic regions derive their water for daily use, are often neglected or oversimplified. Here, using *in situ* data collected from 14 sites in two geographically distinct locations in the Bengal Delta, we show that cholera in coastal regions is characterized by a spring peak associated with seawater intrusion of cholera bacteria to inland water bodies. However, cholera occurs inland from the Bengal Delta in bimodal peaks with two distinct hydroclimatological patterns. The spring season cholera is associated with coastal seawater intrusion, whereas fall cholera outbreaks are driven by floods related to the monsoon season. This study is the first to demonstrate a relationship between *in situ* environmental conditions and large-scale hydroclimatological events with respect to cholera outbreaks, providing yet another step toward the development of prediction models for cholera.

2.2 Introduction

Microorganisms are the most ancient organisms on the planet and the largest reservoir of biodiversity on earth (Konstantinidis, 2006; Staley, 2006). Cohan (2001) stated that "Wherever there is life there are bacteria". Consequently, bacteria are found in all kinds of environments. In those environments with or without eukaryotes, even in environments with no animals or plants are known to have several kinds of bacteria either as commensals or pathogens. The cholera bacterium is one example of an anecdotally reported bacterium. Bishagratna (1963) reported cholera as early as 400 BC. Cholera is considered as the longest known water-borne disease in the history of mankind that can cause an epidemic. *V. cholerae* was first known as Comma bacillus that was discovered by F. Pacini in 1854. The discovery made by John Snow that cholera spreads through contaminated water was a big aid to understanding how cholera disease is transmitted. The pathogenic organisms are present in the aquatic ecosystem of Bangladesh throughout the year (Huq *et al.,* 1990) in a special dormant state, called nonculturable state where surviving strains can multiply when the environment becomes favorable (Huq *et al*., 2005; Chun *et al*., 2009). This finding was useful for further understanding of how cholera is transmitted among humans.

Cholera, an acute water-borne diarrheal disease, continues to be a major public health threat for less developed regions of the world. The cholera outbreak in Haiti in 2010, affecting more than one million people, indicates that our understanding of the occurrence of cholera and its links with the environment is still evolving. The current cholera pandemic is considered to be localized in South Asia, but cholera outbreaks are reported from other regions including Africa, Australia, Europe, and the Americas.

Cholera can cause significant mortality, as high as 60% if untreated, while mortality decreases to less than 1.0 % with adequate treatment and intervention, such as oral rehydration therapy supplemented with appropriate antibiotics (Mahalanabis *et al.,* 1992). *Vibrio cholerae,* the causative agent of the disease, is associated with zooplankton (Huq *et al.*, 1990, Huq *et al.,* 2005) and copepods act as vectors for cholera bacteria in the aquatic environment (Rawlings *et al.*, 2007). The bacterium is autochthonous in riverine, estuarine, and coastal ecosystems (Colwell *et al.*, 1977, Huq *et al.*, 1983 and Islam *et al.,* 1990b), hence cannot be eliminated. However, prediction of an impending cholera outbreak is possible by devising suitable mitigation and intervention strategies for endemic regions (defined as regions where cholera cases are reported throughout the year and have a characteristic seasonal pattern).

Two observations motivated the exploration of possible connections among the large scale hydro-climatic processes and aquatic environment factors: 1) local environmental factors, such as biological activity in ponds from where the bulk of the population in endemic regions derives water for daily use are often neglected or oversimplified in disease outbreak dynamics (Emch *et al.*, 2008); and 2) The proposed hypothesis (Akanda *et al.*, 2009) that explains how large-scale geophysical processes play a major role in cholera outbreaks

The Bengal Delta continues to be prone to endemic cholera. Many studies have postulated environmental links between large-scale hydroclimatic processes and cholera occurrence in the Bengal Delta, Specifically salinity (Miller *et al.*, 1982), temperature (Lobitz *et al.*, 2000; Constantin de Magny *et al.*, 2008), precipitation (Pascual *et al.*, 2002; Hashizume *et al.*, 2008), and chlorophyll (Lobitz *et al.*, 2000; Jutla *et al.*, 2010,

2011, 2012, 2013). Huq *et al.*, 2005 showed water temperature, water depth, rainfall, conductivity, and copepod counts were associated with occurrence of cholera in the Bengal Delta region. Several investigators have hypothesized links between large-scale hydroclimatology and regional-scale environmental processes (defined as processes at a pond scale or bodies of water of approximately 500 m^2) (de Magny *et al.*, 2008; Emch *et al.*, 2008), but none have been validated. Two hypotheses for occurrence of cholera existence in endemic regions of the world are dominant. The first is that when there are two seasonal peaks of cholera, the first peak is a result of contamination of inland water bodies by intrusion of bacteria laden coastal seawater, followed by a second peak related to massive flooding that disrupts the essential safe water infrastructure (Akanda *et al.*, 2009). The second of the dominant hypotheses is that when only a single seasonal outbreak occurs, it is during low river discharge with subsequent intrusion of coastal water, with the consequent transfer of bacteria to inland water bodies (Akanda *et al.*, 2009, 2011). A key question remains and that is whether a similar phenomenon can be observed on a local scale, such as where ponds serve as the household water source for the local inhabitants. The goal of this chapter is to develop a quantitative understanding of the role of local environmental factors in outbreaks of seasonal cholera, with the objective of determining whether local environmental factors and large-scale hydrodynamic factors associated with endemic cholera (Akanda *et al.*, 2009) could be linked.

2.3 Materials and Methods:

2.3.1 Study region

In situ data from two rural sites, Bakergonj and Mathbaria in Bangladesh, covering a time period from March, 2004, to September, 2007, were used in the study. Environmental factors (pH, total dissolved oxygen (TDO), total dissolved solids (TDS), conductivity, salinity, water temperature, air temperature, chlorophyll-a chlorophyll-b chlorophyll-c, phaeophytin, phosphate, silicate, nitrate, alkali, heterotrophic bacteria and coliform count) were analyzed from bimonthly water samples obtained at Mathbaria (six pond sites) and monthly water samples obtained from Bakergonj (eight pond sites).

Bakergonj is an inland region located at the upper edge of the southern estuary of Bangladesh, a cholera-prone area in the district of Barisal and situated at the southwestern part of Bangladesh, near the Bay of Bengal. The eight sampling sites in Bakergonj are North Varpasha Helipad, Mukharjee Bari, Mid Varpasha Jam-E Mosque, Thana Health Complex, Tolatoli River, Harun Dakua, Bara-Aolia Mazar, and Bairam Kha Lake.

Mathbaria is a coastal region situated at the southwestern part of Bangladesh. It is in Upazila Parishad (UZP), a sub-district located in the Pirojpur District. There are six different sampling sites (ponds) in Mathbaria, South Mithakhali Govt, Jotish Kant Common, Ishaq Akand's Common, Mathbaria Bazaar, BRAC Pond, and Arshed Bara Mia's Pond.
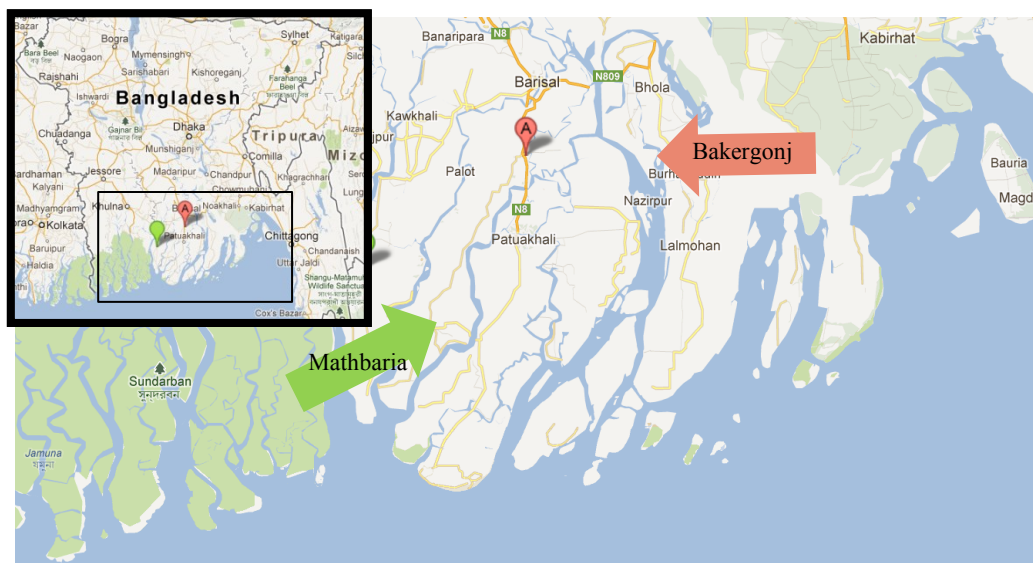
Figure 2.1 Map of the location of both Mathbaria and Bakergonj in Bangladesh (Google earth)

2.3.2 Environmental specimen collection and source of clinical data:

Details on specimen collection have been published earlier (Huq *et al*, 2005). Two plankton nets were double stacked by placing a 64μm mesh net inside a 20μm mesh net. *Ca.* 100 L of water sample was passed through the double stacked nets using a 15L bucket fixed onto a large stand. A total of seven buckets of water collected from seven distance points in a pond was used to obtain a heterogeneous mixture of bacterioplankton. Additional liters of water (105L) were used to adjust losses when a bucket of water was poured through the stacked net. Between the 2nd and 4th bucket flows, 500 mL of plankton-free water was collected by allowing the filtrate from the stacked nets to flow into a bottle (labeled MWA). All samples were collected aseptically in sterile dark Nalgene bottles (Nalgene Nunc International, USA), placed in an insulated plastic box, and transported at ambient air temperature from the site of collection to the central

laboratory at the International Center for Diarrheal Research, Bangladesh (ICDDR, B) in Dhaka. All samples were processed the following morning, with approximately 20 hours elapsing between sample collections in the field and processing in the laboratory. Air and water temperature, water conductivity, total dissolved solids (TDS) and salinity were measured using a portable meter (HACH model sens*ION*156 portable pH/conductivity/dissolved oxygen meter, Yellow Spring, Ohio, USA). Dissolved oxygen (DO) and pH were also measured using the same sens*ION*156 portable meter connected to a platinum pH electrode and DO probe, respectively. All parameters were measured by inserting the probe directly into the pond water and, when it was raining, water from the pond was collected in a 5L plastic beaker and measured therein. The pH probe was calibrated by employing three points of calibration, using pH 4, 7, and 10. The DO probe was calibrated using saturated air and the conductivity probe using 1000 µS/cm conductivity standard. A graduated rope, with a heavy iron ring attached, was used to measure water depth at a fixed point in the center of each pond. Calibration frequency was a measurement made prior to each round. One Liter of unfiltered water was collected to measure chlorophyll-a (µg/L), chlorophyll-b (µg/L), and chlorophyll-c (µg/L). Phaeophytin (µg/L), silicate (mg/L), nitrate NO_3 (µg/L), phosphate PO_4 (µg/L), and alkalinity (meq/L) were also measured. For fecal coliforms, unfiltered water (UFWA) was used, with 10ml and 1ml samples filtered using a 47mm #?µm pore size nucleopore polycarbonate filter and 100 µl of filtered water was directly spread, or spot-plated (approximately 25-50 spots), onto m-FC agar plates. The plates were sealed with parafilm, placed into petri dish bags (double layer) incubated in a plastic container overnight at 44.5°C, and observed after incubation for 24 hrs. Results are reported as

CFU/100-ml. Rectal swabs were obtained every two weeks from patients admitted to the local sub-district Health Complex of Bakergonj and Mathbaria. (Personal communication, Sack *et al*, 2014, in preparation). The rectal swabs from suspected cholera patients, i.e., those with watery stool, were collected for three consecutive days and placed into Cary Blair medium (in grams per liter: sodium thioglycolate, 1.5 g; disodium phosphate, 1.1g; calcium chloride, 0.1 g; sodium chloride, 5 g; agar, 5 g; pH 8.4) immediately after swabbing, and transported at ambient temperature to the central laboratory of the ICDDR, B in Dhaka, Bangladesh.

2.4 Statistical analysis:

Clinical cholera case and environmental data were collected from March, 2004 to September, 2007. Cholera cases were averaged creating two time series for Bakergonj and Mathbaria. A generalized linear model (GLIM) with Poisson distribution and log link was used to model the data and build a simulation model for cholera, following the method of Cameron and Trivedi (1998). Choice of model was based on the overall regression Chi-square statistic, significance of individual variable coefficient estimate at the 95% confidence level, and pseudo R^2, as well as low error values.
The multiple Poisson regression model (GLIM) for k predictors is as follows:

$$Y_i \sim \text{Poisson} (\mu_i)$$

$$\ln (Y_t) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 \ldots\ldots\ldots\ldots.. + \beta_k X_k + e_i \qquad [1]$$

Where k predictors are X_1, X_2, ..., X_k and β are the corresponding model coefficients. The model choice was based on the best fit with the count data and because we were interested to explore the role of each environmental factor. The model fit was

conducted for the seasonal data set and was examined by comparing observed and predicted numbers of cases in each area. The entire dataset was divided into two seasons:1) Spring season inclusive of January through May; and 2) Fall season inclusive of observations averaged from August through December. The diagnostic experiments and data analysis were focused on identifying the significant variables correlated with *in situ* data (collected directly from the ponds) and large-scale data (obtained by satellite remote sensing) and cholera incidence and dynamics. Results from these analyses were expected to provide new knowledge about how local environmental conditions of two endemic regions, with distinct regional settings, can aid in development of a cholera simulation model.

2.5 GBM (Ganges-Brahmaputra- Meghna) Basin Hydroclimatology

The Ganges-Brahmaputra-Meghna (GBM) system is considered one of the largest river basins that have three rivers with distinct characteristics and flow through very different regions. This system runs through Tibet, Nepal, and Bangladesh and is composed of Himalayan Rivers, the Ganges and the Brahmaputra. The GBM is a very complex system not only because each one of its individual rivers are large, but also because each one of them has its tributaries, important in many ways including water availability and use. Part of the GBM complexity derives from the transboundary nature of many of its tributaries (Ahmad *et al*., 2001; Biswas and Uitto, 2001).

The GBM rivers are joined by the Meghna in Bangladesh. The Ganges-Brahmaputra-Meghna (GBM) basin has a highly seasonal hydroclimatic nature that is characterized by annual precipitation occurring during June–September (Chowdhury and Ward, 2004). Akanda *et al*., (2011), however, showed that during the prolonged dry

20

season (December–May), only a fraction of the average flow reaches the Bay of Bengal (BOB). This leads to a salinity akin to those in estuarine regions in the inland freshwaters (Rahman *et al*., 2000), thus providing an optimum environment for *V. cholerae* growth (Louis *et al*., 2003). A large part of coastal Bangladesh will exhibit brackish water conditions in spring (Miller *et al*., 1982; Louis *et al*., 2003; Wahid *et al*., 2007; Islam and Gnauck, 2008).

During the monsoon rain (June through September), the river rises rapidly, causing flooding in Bangladesh and large-scale contamination of the water system, including the autochthonous cholera bacteria that are in the water ecosystem (Schwartz *et al*., 2006; Akanda *et al*., 2009). Chowdhury and Ward, 2007 showed that in 1988 and 1998, at 60 percent high flood nearly 20 percent of the land area of Bangladesh is inundated. Under flood conditions, all water system including reservoirs, and submerged areas, will contain *Vibrio cholerae*, with 2/3 of the samples collected over a three-year study testing positive for *V. cholerae* O1 (Huq *et al*., 1990). In addition, during the 2004 floods about 62.5% of water samples were contaminated by cholera bacteria in the suburban reservoirs around Dhaka (Islam *et al*., 2006). Mirza *et al*., 2001 studied the submerged area December–January and found that it remained positive for *Vibrio cholerae*, and served as an ideal habitat for its growth and multiplication after the flood. This can be explained as the river levels falls from September through November in most of the parts of the river, the adjoining flood plain water level decreases slowly because of low gradients, congested drainage, and substantial depression areas.

2.6 Dynamics of cholera seasonality:

Cholera is dynamic throughout different parts of the world. The disease varies in frequency, severity, duration, and endemicity. In South Asian countries line Bangladesh, cholera is endemic, that is, cholera occurs every year while cholera in other regions, such as parts of South America and Africa has sporadic epidemics. Although endemic cholera regions are known as areas of persistent cholera, i.e., cholera occurs there every year but the seasonal rate varies greatly from year-to-year. Since the early 19th century, the Bengal Delta estuary formed by the Ganges and the Brahmaputra rivers has been considered as the native homeland of cholera (Bouma and Pascual, 2001). The International Center for Diarrhoeal Disease Research (ICDDR) in Bangladesh has a cholera surveillance program that provides some of the longest and largest records available in the world. The cholera surveillance program serves as the main treatment center for the most concentrated population center in Bangladesh that carries out a systematic sub-sampling of all patients visiting the hospital.

The seasonal cycle for cholera outbreaks in Bangladesh is described by several studies. These studies focused on describing the seasonal cycle for specific cholera strains such as classical (Samadi *et al*., 1983), El Tor (Khan *et al*., 1984), and O139 (Heidelberg *et al*., 2002). Different cholera strains (classical, El Tor, and O139) show differences in their seasonal cycles although the symptoms of the cholera they cause are similar. *V. cholerae* El Tor has caused most cases in a seasonal cycle from September to November, just after the monsoon while the classical cholera strain has a dominant seasonal cycle after the peak of the newer strain (Merson *et al*., 1980; Glass *et al*., 1982; and Samadi *et al*., 1983). The El Tor cholera seasonal cycle is described further in several

additional studies as having two annual pattern peaks, namely an April peak before the monsoon (smaller spring outbreak) followed by a September to December peak after the monsoon (larger fall outbreak) (Lipp *et al*., 2002; Sack *et al*., 2003; Islam *et al*., 1993; Baqui *et al*., 1992; Huq and Colwell, 1996). When *V. cholerae* 0139 first appeared in Bangladesh in 1993, it showed a similar seasonal pattern (Emch and Ali, 2001). Cholera seasonal cycle patterns are evident in other parts of the world. Classical cholera in Pakistan typically has two peaks: fall peak (November to January) and spring peak from (April to May) (Martin *et al*., 1969 and McCormack *et al*., 1969). On the other hand, cholera in Kolkata, India, shows several seasonal cholera cases peaks, in April, May, and June (Gangarosa and Mosley 1974; Kaper *et al*., 1995).

The seasonality information of cholera outside of the South Asia region is limited since there are a few surveillance systems that can collect detailed cholera incidence from these regions. Seasonal cholera peaks in South America occur in the summer months (Lipp *et al*., 2002; Cockburn *et al*., 1960; Kaper *et al*., 1995) with the rise in waters following the rainy season. In east African nations including Djibouti, Kenya, Mozambique, cholera outbreaks are reported in summer season following rainfall and/or floods (WHO, 1998). Cholera reports from 1979 to 1983 show two peaks in Dar es Salaam and Tanzania; one peak [from October to December] followed by a second peak [from March to May], both of which coincide with periods of increased rainfall (Mhalu *et al*., 1987). The seasonal peak of cholera epidemic in rural southern Tanzania is slightly later in June and July (Acosta *et al*., 2001). The high cholera season for the northern parts of southern Africa in 2002 occurred from the end of January to mid-March, while the cholera peak in Mozambique has a slightly longer period that occurs during the hot, rainy

months from December to May (Bateman *et al*., 2002; Folgosa *et al*., 2001; Aragón *et al*., 1994).

The cholera incidence climatology in our study is constructed by averaging *V. cholerae* O1 records across sites of the 2004-2007 time series in two distinct places in Bangladesh (Mathbaria and Bakergonj), that exhibits significant seasonal and inter-annual variability (Figure 2.2). A closer look reveals that cholera in the Delta Bengal exhibits a distinct seasonal and spatial variation that is both complex and dynamic. Mathbaria and Bakergonj are located in two regions of Bangladesh, but they have different seasonal patterns. Cholera in Mathbaria has a single annual peak found in the spring season (April) while unique dual-cholera peaks typically are observed during spring and fall seasons in Bakergonj. We can see the highest number of cholera cases in October during the large peak in the fall, followed by a smaller peak in the spring (Figure 2.2). Akanda *et al*., (2009) used data (1980-2000) to show that some regions of the Bengal Delta, such as Dhaka and Matlab, have a dual cholera incidence while other regions such as Mathbaria (spring peak) and Chhatak (fall peak) have a single annual peak.
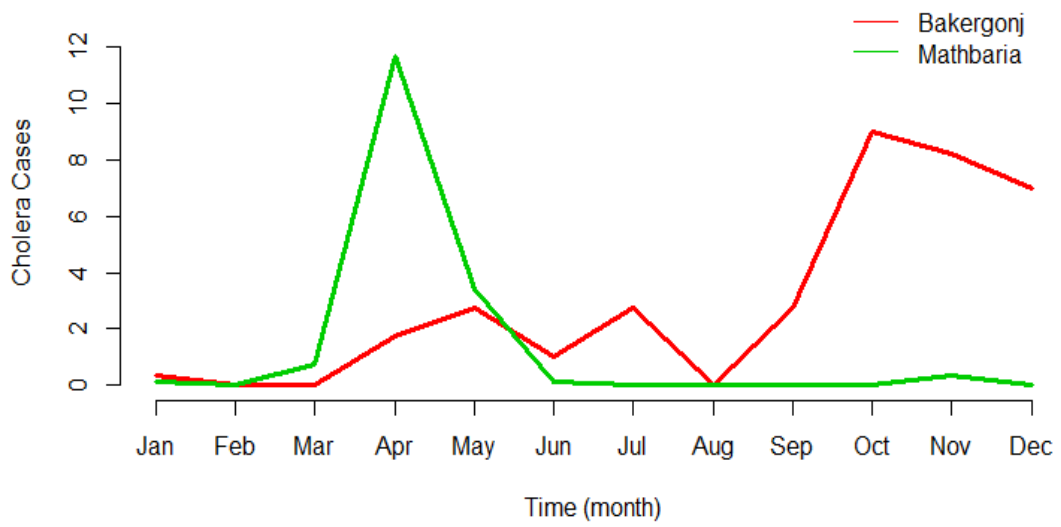
Figure 2.2 Seasonality of cholera for Bakergonj and Mathbaria, Bangladesh (2004-2007)

2.7 Cholera and Coastal Regions:

The modern history of cholera began in 1817, when explosive cholera epidemic spread out from the Ganges River delta to the entire world as the first of seven pandemics. The Seventh Pandemic first reached Africa, specifically Guinea, in August, 1970 (Echenberg, 2011) and spread along the African coast. The seventh pandemic global pattern and magnitude started in 1961 in Indonesia and was reported in over 50 countries, suggesting that coastal environments are the main origin of cholera outbreaks, spreading inland by secondary means (Colwell, 1996). Travel ease and the increasing number of travelers is considered to contribute to the spread of the pandemic. The main source of cholera transmission is along the African and Asian coasts, whether it is estuarine regions or not such as Accra and Cape Coast in Ghana (WHO, 2011), Dakha in Bangladesh (WHO, 2008), Luanda in Angola (WHO 2007and 2009; Grestl and Alberti, 2006, and Quito in Ecuador). However, these regions are also characterized by densely populated urban settings. There are many studies that explain the emergence of cholera from coastal regions. However, Siddique *et al*. (1994) show a plausible cholera outbreak progression pathway from Bangladesh coastal regions to inland areas of the country during a major pandemic. Moreover, studies of Epstein (1993) and Colwell and Huq (2001) explained the cholera-coastal connection by *Vibrio cholerae* strains, the causative agent of cholera, being found mainly in marine plankton. Some evidence for historical coastal link to cholera mortality in the Bengal Delta formed by the joining of Ganges and the Brahmaputra rivers, were provided by Lobitz *et al*., (2000) and later by Bouma and Pascual (2001). Lobitz *et al* (2000) showed significant correlation between sea surface temperature (SST) and spring cholera deaths in the coastal districts (Lobitz *et al*., (2000);

Bouma and Pascual, 2001). Sack *et al*. (2003) used a detailed cholera epidemiological study from Bangladesh, including both Bakergonj and Matlab to show that inland regions have mostly fall outbreaks while coastal regions have spring peaks. In most regions, the initial outbreaks of cholera show a strong and significant correlation with the coastal areas, indicating the important role of the coastal marine environment in transporting cholera to other regions. Since the initial coastal cholera outbreak in Africa in 1970, the disease has spread along maritime, terrestrial, and aerial routes from coastal regions to other parts of the continent. Hence, several inward and cross-border epidemics have been identified such as the spread of cholera from Comoros to Madagascar in 1999 (Duval *et al*., 1999) and from Guinea-Bissau to Senegal in 1995 (Aidara *et al*., 1998). On the other hand, the outbreak of cholera in 1991 in about 20 Latin American countries was first reported in coastal villages near Lima, Peru, with a total of 5000 deaths. Recent studies investigated genetic comparisons of cholera strains isolated in coastal African countries and showed that both new and atypical strains of *V. cholerae* El Tor which had emerged in the early 1990s in the Bay of Bengal could be isolated along the are coast of Africa (Olsen *et al*., 2009; Lam *et al*., 2010, 2012; Safa *et al*. 2010). Considering the importance of coastal cholera outbreaks, Jutla *et al*. (2010) provided a modeling framework that enhanced our understanding of cholera prediction dynamics, connecting the ''macro'' (hydrological, ecological, climatic, and coastal processes) and the''micro'' (microbiological, genetic, and human intestinal-scale processes) dimensions.

2.8. Role of environmental factors in seasonality and modulating cholera outbreaks.

The life cycle of *Vibrio cholerae* is linked to two distinct environments, the microenvironment, and the macro-environment (Jutla, *et al*., 2010) and they are quite

different and affected by multiple factors. *Vibrio cholerae* occurs in a commensal relationship with zooplankton in brackish waters of the macro-environment (Huq *et al*., 1984; Epstein, 1993; Alam *et al*., 2006). Zooplankton act as vectors for *Vibrio cholerae* by providing nutrients and also physically carry the bacterium, playing an important role in facilitating survival, multiplication, and transmission of *Vibrio cholerae* in the natural aquatic environment (Lipp *et al*., 2002). *Vibrio cholerae* growth, multiplication, and ability to cause disease is governed by many processes, including microbiological and genetic processes. The macro-environment is governed by many climatic, ecological, and hydrological processes. Numerous studies have focused on the microenvironment, development of new vaccines, antibiotics, and protocols for disease treatment while macro-environmental processes have focused on the ecology of *Vibrio cholerae* and developing prediction tools.

In general, the continuous development of newer types of remote-sensing devices will allow more data to be available to study the emergence of water related diseases. Prediction of many water-related diseases has already been significantly enhanced, such as Rift Valley fever (Linthicum *et al.,* 1999), malaria (Hay *et al*., 1998; Adim *et al*., 2010), and schistosomiasis (Malone *et al*., 2001). Studying cholera dynamics by using remote sensing is a promising research area that will provide large and information dense datasets (Harvell *et al.,* 2002). The first remote-sensing data were used by Lobitz *et al*., (2002) to investigate the connection between SST, phytoplankton, and cholera. Many studies have demonstrated an association between environmental and climate-derived satellite data with cholera, such as coastal salinity (Miller *et al*., 1982), fecal contamination (Islam *et al*., 2006), precipitation (Pascual *et al*., 2002; Hashizume *et al*.,

2008), floods (Koelle *et al*., 2005), river discharge (Jutla *et al*., 2010; Akanda *et al*., 2009), sea surface temperature (SST) (Lobitz *et al*., 2000; de Magny *et al*., 2008), chlorophyll and river discharge (Jutla *et al*., 2010). Table 2.1 shows some important examples of cholera relationships with climate and how they were analyzed.

Table 2.1 Available environmental data and their relation to cholera

Source.No	Author	Data analysis Type	Variables	Large/Regional Environmental variables	Scale (spatial , temporal)
1	Jutla *et al*., 2010	Analytical/ Statistical based	Chlorophyll and river discharge	Large-scale Environmental variables	Monthly, three main rivers/coasts
2	Emch *et al*., 2008	Statistical based	Large Scale: chlorophyll, Sea surface temperature, Sea surface height Regional Scale: Rainfall, Temperature, River discharge/height	Large and regional-scale Environmental variables	Monthly (three cities)
3	Magny *et al*., 2008	Statistical (regression) based	Previous season cholera incidence, chlorophyll, precipitation, SST	Large scale Environmental variables	Monthly, (two cities)
4	Akande *et al*., 2009	Analytical/ Statistical based	Sea surface temperature, river discharge	Large-scale Environmental variables	Monthly, two main rivers

Emch *et al.,* (2008) used several environmental variables, including SST, river discharge, sea surface height, as well as chlorophyll measurements, from three coastal regions in South Asia and reported a two month lag between plankton blooms and cholera outbreaks in Bangladesh. Magny *et al*., (2008) used coastal plankton data derived from Bengal Delta to establish a prediction model between cholera epidemic and

plankton. Jutla *et al*., 2010 focused on large-scale hydroclimatic processes, such as coastal chlorophyll and how it can affect cholera outbreaks. However, none of these studies investigated successfully the role of local environmental processes and how they are correlated with large-scale environmental processes. Local environmental processes are identified as the biological activity in the ponds from where the bulk of the population in endemic regions derives water for daily household usage. The analysis reported here was designed to understand disease transmission dynamics affected by local environmental processes as well as how it connects the local environmental processes to large-scale environmental processes.

The following questions were addressed: (i) What are the local environmental factors that will help understand disease transmission dynamics? (ii) Are there other kinds of chlorophyll that can be linked to cholera incidence other than chlorophyll A? (iii) What is the mechanism of cholera spreading among different regional settings?

Preliminary construction of the correlation analysis for all available local environmental and physicochemical variables collected from the ponds in both Mathbaria and Bakergonj suggests that the major local environmental factors affecting cholera outbreak are water depth and chlorophyll-c (Figure 2.3 and 2.4; Table 2.2). This step is very important as it is done to determine: 1) What local environmental and physicochemical variables are significantly correlated with cholera incidence; 2) What variables with multi-collinearity should be excluded from the model to construct a good quality model. The second step in my analysis was construction of simulation models to describe how these major environmental variables affect a cholera outbreak.

Each region (Mathbaria and Bakergonj) is described separately:

In the Mathbaria model, water depth and chlorophyll-c were the two major environmental parameters responsible for cholera outbreaks since they describe about 64.17% of the model variance. A moderate negative correlation was observed between cholera cases and water depth (r = - 0.47), while a positive association was found with chlorophyll-c (r = 0.29) (Table 2.3). The increase in pH and dissolved O_2 concentration provide a favorable environment for *V. cholerae* (Oppenheimer *et al*., 1978, Huq *et al*., 1984a and Isalm *et al*., 1994). Interestingly, pH has a weak negative significant correlation (r = - 0.07) and a negative coefficient in the model. This can be explained by its confounded effects that arise from its low variability, since our data have pH values of 7.5 and higher, and the optimum pH for *Vibrio cholerae* is 8.5 (Huq *et al*., 1984; Oppenheimer *et al*., 1978), thus, a one point increase above the optimum pH will be inversely proportional to cholera.
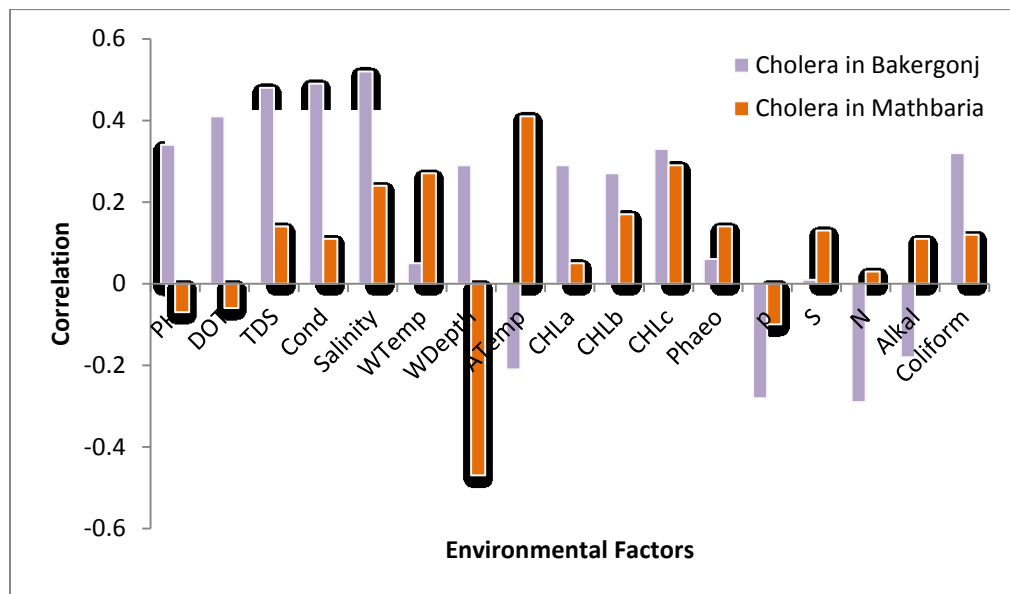
Figure 2.3 Cross-correlation between cholera cases and different local environmental factors in both Mathbaria and Bakergonj, Bangladesh

The pH values are included in the model because pH is a critical factor affecting *Vibrio cholerae* growth. Model performance without it is poor since it explains only

68.18% of the variance, compared with performance of the model which explains 77.7%. Air temperature showed a positive correlation with cholera cases (r = 0.41) as the temperature increased. Especially above 30°C since *V. cholerae* multiplication increases with temperature increase (Singleton *et al.,* 1982; Paz, 2009). Cholera bacteria can survive and thrive in brackish waters, particularly in the presence of abundant zooplankton and phytoplankton, suggesting high correlation between plankton abundance and disease outbreaks (Huq *et al.*, 1984; Epstein, 1993; Alam *et al.*, 2006). Phytoplankton and zooplankton act as vectors for *Vibrio cholerae* by providing nutrients and also physically serve as a host for the bacterium, playing an important role in facilitating survival, multiplication, and transmission of *Vibrio cholerae* in the natural aquatic environment (Lipp *et al.*, 2002).
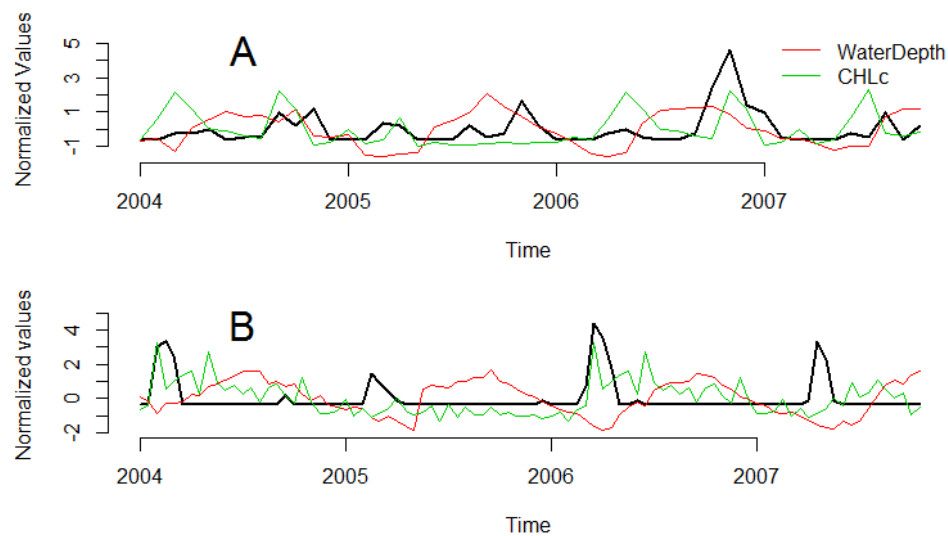
Figure 2.4 Seasonal correlations between Water Depth, Chlorophyll-c (CHLc) and cholera cases (black line) from 2004 to 2007, (A) Bakergonj, (B) Mathbaria

In the Bakergonj model, conductivity explains almost double (22%) the variance explained by water depth, chlorophyll-c, phosphate, and nitrate, individually. A positive correlation was observed between water depth (r = 0.29) and cholera cases as well as positive association between chlorophyll-c (r = 0.33) and cholera cases (Figures 2.3 and

2.4). In addition to water depth and chlorophyll-c, other environmental parameters, such as nitrate, phosphate and conductivity were found to be correlated with cholera outbreaks in Bakergonj. Both phosphate and nitrate showed a negative significant correlation with cholera cases $r = - 0.28$ and $r = - 0.29$, respectively.

The pseudo R^2 (appendix 1) value for the fitted model of Bakergonj is equal to 48.5%, which is lower than the fitted model for Mathbaria. This can be explained by the fact that twice as many samples per month were collected from Mathbaria (bimonthly) as compared to those (sample number/frequency) for Bakergonj (monthly). Chlorophyll, an indicator of phytoplankton, and phytoplankton a surrogate for zooplankton, has received extensive attention among scientists in several related studies since the late 1990s (Olson, 1996; Huq and Colwell, 1996). Lobitz *et al.,* (2000) explored the potential role of chlorophyll measured by remote-sensing satellites to understand the phytoplankton-zooplankton-cholera relationships. From Table 2.3, we can see that chlorophyll c after time delay calcination is significantly correlated with cholera incidence in both Bakergonj and Mathbaria ($r = 0.33$ and $r = 0.29$, respectively).

Most of the recent cholera models, demonstrated in Table (2.5) are based either on regression or semi-mechanistic processes and developed using large-scale environmental variables (Magny *et al*., 2008; Pascual *et al*., 2008; Matsuda *et al*., 2008). **A *cholera simulation model (CSM) using in-situ data collected from several sites/ponds in two geographically distinct locations in Bangladesh is presented in Table 2.3.***

The proposed CSM for Mathbaria takes the following functional form:

$$Y_i \sim \text{Poisson}(\mu_i)$$

$$\ln Y_i = \beta_0 - \beta_1 X_{WDepth} + \beta_2 X_{ATemp} + \beta_3 X_{CHLc} - \beta_4 X_{PH} \qquad \text{(Equation 2)}$$

The proposed CSM for Bakergonj may take the following functional form:

$$Y_i \sim \text{Poisson}(\mu_i)$$

$$\ln Y_i = \beta_0 + \beta_1 X_{WDepth} + \beta_2 X_{Cond} + \beta_3 X_{CHLc} - \beta_4 X_{Nitrate} - \beta_5 X_{Phosphate} \qquad \text{(Equation 3)}$$

Table 2.2 Correlation between 17 local environmental, physicochemical variables and cholera cases in both Mathbaria and Bakergonj, Bangladesh (2004-2007).

Independent variables for Multiple Poisson regression modeling

	O1 in Bakergonj	PH	DOT	TDS	Conductivity	Salinity	Water Temp	Water Depth	Air Temp	CHLa	CHLb	CHLc	Phaeophytin	Phosphate	Silicate	Nitrate	Alkal	Coliform	O1 in Mathbaria
PH	0.34	1	0.06	0.25	0.23	0.16	0.35	-0.07	0.36	0.24	0.14	0.14	0.25	-0.04	0.07	-0.16	0.03	0.13	-0.07
DOT	0.41	0.34	1	-0.0079	-0.07	0.02	-0.001	0.1	-0.001	-0.15	-0.12	0.08	-0.18	0.26	-0.15	-0.03	-0.09	-0.19	-0.06
TDS	0.48	0.55	0.33	1	0.98	0.94	0.37	-0.46	0.3	0.3	0.24	0.37	0.29	-0.05	0.41	-0.04	0.55	-0.33	0.14
Conductivity	0.49	0.54	0.38	0.98	1	0.93	0.36	-0.43	0.28	0.31	0.27	0.32	0.29	-0.06	0.39	-0.09	0.53	-0.34	0.11
Salinity	0.5173	0.53	0.323	0.97	0.94	1	0.32	-0.5	0.283	0.37	0.34	0.32	0.27	-0.04	0.45	-0.01	0.61	-0.32	0.24
Water Temp	0.05	0.08	0.082	0.1	0.18	-0.01	1	-0.18	0.89	0.2	0.07	0.31	0.36	0.11	0.3	-0.22	0.13	0.003	0.27
Water Depth	0.29	-0.11	0.072	-0.2	-0.17	-0.09	0.08	1	-0.35	-0.37	-0.45	-0.12	-0.32	0.15	-0.52	-0.11	-0.54	0.04	-0.47
Air Temp	-0.21	-0.04	0.15	-0.09	-0.011	-0.25	0.8	-0.15	1	0.12	0.1	0.19	0.3	0.01	0.26	-0.17	0.15	0.12	0.41
CHLa	0.29	0.58	0.079	0.42	0.44	0.38	0.3	-0.32	0.1	1	0.68	0.25	0.73	0.01	0.17	-0.12	0.18	-0.009	0.05
CHLb	0.27	0.48	0.027	0.28	0.3	0.23	0.05	-0.25	0.04	0.65	1	-0.17	0.49	-0.18	0.12	-0.14	0.27	-0.08	0.17
CHLc	0.33	0.37	0.17	0.24	0.26	0.2	0.27	-0.21	0.16	0.7	0.2	1	0.22	0.21	0.14	-0.03	-0.008	-0.02	0.29
Phaeophytin	0.06	0.34	0.056	0.13	0.15	0.03	0.2	-0.5	0.36	0.6	0.48	0.59	1	-0.06	0.15	-0.16	0.06	0.001	0.14
Phosphate	-0.281	-0.15	-0.13	-0.33	-0.31	-0.37	0.32	-0.05	0.41	0.09	-0.17	0.33	0.27	1	-0.13	-0.12	-0.16	-0.2	-0.1
Silicate	0.009	0.31	0.08	0.03	0.03	0.03	0.054	-0.27	0.02	0.39	0.19	0.31	0.31	0.02	1	0.16	0.52	-0.16	0.13
Nitrate	-0.29	-0.27	-0.26	-0.31	-0.3	-0.28	-0.11	0.2	-0.09	-0.15	-0.35	-0.03	-0.2	0.13	-0.19	1	0.09	0.004	0.03
Alkal	-0.18	0.37	0.13	0.46	0.44	0.38	-0.1	-0.76	0.01	0.23	0.27	0.08	0.28	-0.26	0.26	-0.17	1	-0.22	0.11
Coliform	0.32	0.2	0.033	0.32	0.26	0.35	-0.13	0.13	-0.17	0.13	0.25	0.002	-0.08	-0.09	-0.42	-0.11	-0.2	1	0.12

Mathbaria correlation matrix is in the upper diagonal while Bakergonj correlation matrix is in the lower diagonal. The red color indicates strong correlation, -0.6 > r > 0.6, (multi-collinearly if exists between the independent variables), the green color indicates moderate correlation, -0.4 > r > 0.4, the yellow color indicates weak correlation, -0.28 > r > 0.28, the white color indicates very weak correlation.

Table 2.3 Summary of models obtained for Mathbaria and Bakergonj in Bangladesh

Model Parameters

		Intercept	Water Depth	CHLc	PH	Air Temp	Conductivity	Nitrate	Phosphate	df
Mathbaria										60
Coefficient		14.388	- 7.7123	0.1125	-2.220	0.439				
P-value		0.001	0.000	0.000	0.000	0.000				
SE		4.3566	1.532094	0.02423	0.3583	0.10448				
95%	Min	5.849	-10.71524	0.065	-2.923	0.2344				
CI	Max	22.927	-4.7095	0.16	-1.518	0.6439				
Bakergonj										30
Coefficient		-2.554	1.9456	0.066			0.0024	- 0.003	- 0.1117	
P-value		0.009	0.000	0.000			0.047	0.001	0.002	
SE		0.97278	0.3461377	0.01351			0.0012361	0.00101	0.0036767	
95%	Min	-4.46	1.267194	0.03978			0.000029	-0.0052	-0.0183	
CI	Max	0.06477	2.624029	0.0927			0.0048	-0.0012	-0.00396	

Table 2.4 Available cholera prediction models

Author	Model Type	Variables
Magny *et al*., (2008)	Regression model	chlorophyll, precipitation, SST
Pascual *et al*., (2008)	Semi-mechanistic model	Population, biological variables (immunity levels, susceptibility rates), ENSO
Matsuda *et al*., (2008)	Regression model	rainfall, air temperature

2.9 Physical hypothesis for the role of local environmental factors in modulating cholera outbreak dynamics

Cholera incidence in Bengal Delta has been linked to large-scale environmental and climate variables. Akanda *et al*. (2009) linked cholera outbreak peaks with the flow season of the GBM Rivers; the first small spring cholera peak in Mathbaria and Bakergonj occurs during the dry season while the second bigger fall peak in Bakergonj only occurs in the wet season. Typically cholera shows a single incidence peak in a year through other affected regions in the world, such as Southeast Asia, sub Saharan Africa, southern Africa and South America (Emch *et al*., 2008; Hashizume *et al*., 2008; Bertuzzo

et al., 2008; Gil *et al*., 2004). The main objective of this study was to investigate whether the biannual cholera peaks in the Bengal Delta region (Mathbaria and Bakergonj) are governed by two spatially distinct seasonal transmission mechanisms, influenced by local hydroclimatic processes that match the large scale ones.

In this study, a working hypothesis was developed to explain how local environmental factors are related to outbreaks of cholera. During the spring seasons in this study period, cholera case values were found to be inversely related to water depth, i.e., spring cholera peaks in both Mathbaria and Bakergonj are seen in strong drought years. On the other hand, fall cholera outbreaks in Bakergonj are positively correlated with water depth, i.e., fall peaks occur in high flood years (Figure 2.5). The working hypothesis was that spring cholera outbreaks occur as a result of coastal plankton intrusion through low river discharge, whereas fall cholera outbreaks occur due to widespread flooding. If this hypothesis is valid, then a statistically significant negative correlation should be observed between the fall cholera outbreaks peak in Mathbaria with water depth (r^2 = - 0.47; P < 0.01) and statistically significant positive correlation between the spring cholera outbreaks peak in Bakergonj with water depth (r^2 = 0.29; P < 0.01). The question is whether local and large-scale drivers of cholera outbreaks are related to each other or not. This can be answered by the work of Akanda and his colleagues (2009) who originally suggested this hypothesis. They explained the role of large-scale river discharge in cholera outbreaks and stated that spring cholera in the Bengal Delta is hypothesized to be the result of coastal plankton intrusion through low river discharge, whereas fall cholera outbreaks occur due to widespread flooding in the

region. Thus the objective of this part of the study was to explore the role of local environmental water depth on cholera outbreaks.

Figure 2.5 Water-depth (on regional scale) as an important factor in seasonality of cholera outbreaks. Lower water depth can be related to coastal intrusion and spring cholera outbreaks. High water and flooding are factors in fall cholera outbreaks (Akanda *et al*., 2009) based on large-scale environmental processes.
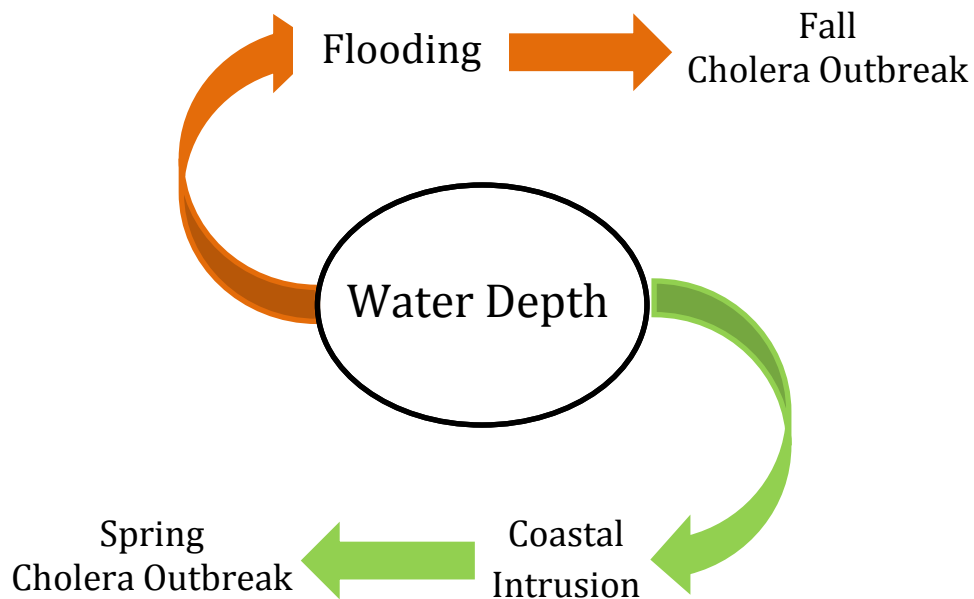
We further validated our hypothesis by separating Mathbaria and Bakergonj data into two seasons (spring and fall) to build a simulation model that can assess the validity of the hypothesis for the role of water depth on transmission of cholera throughout different seasons. Table 2.5 shows results for the best simulation model performance obtained (equation 4) for the spring and fall peaks, after analyzing all variable combinations, using a forward and backward selection method.

The model of a spring peak in Mathbaria had the highest pseudo R^2 (70.9%). It was obtained using water depth, chlorophyll-c, heterotrophic bacteria, and air temperature. The coefficient for water depth yielded a negative sign for ($\hat{\beta}_1$= - 7.74; p

value < 0.005), table 1), indicating an inverse relationship with cholera cases. Overall water depth explained about 57.3% of the model variance. A positive coefficient was observed for chlorophyll-c ($\hat{\beta}_2$ = 0.0768; p value < 0.005). The model of spring peak in Bakergonj showed water depth, coliform counts, and heterotrophic bacteria were associated with spring cholera cases. The coefficient for water depth was negative ($\hat{\beta}_1$ = -13.5; p value < 0.005, Table 1) also for Mathbaria, indicating an inverse relationship with cholera cases. The coliform and heterotrophic bacteria counts showed different associations with cholera ($\hat{\beta}_2$ = - 0.29; p value < 0.005 and $\hat{\beta}_3$ = 0.136; p value < 0.005 respectively). Compared with the spring cholera model for Mathbaria (Pseudo R^2 70.9%), the Bakergonj spring cholera model yielded a low Pseudo R^2 equal to 47.1%. The difference between the model fitness for the two sites may reflect the difference in the frequency of sampling and analysis sampling frequency. Sampling was bimonthly in Mathbaria and monthly in Bakergonj. ***The spring and fall peaks of the cholera simulation model (CSM) were revealed from in-situ data collected from several sites/ponds in two geographically distinct locations in Bangladesh (Table 2.5).***

The proposed CSM for Mathbaria spring peak takes the following functional form:

$$Y_i \sim \text{Poisson} (\mu_i)$$

$$\ln Y_i = \beta_0 - \beta_1 X_{\text{WDepth}} + \beta_2 X_{\text{CHLc}} + \beta_3 X_{\text{ATemp}} - \beta_4 X_{\text{Heteroplate}} \qquad [4]$$

The proposed CSM for Bakergonj spring peak takes the following functional form:

$$Y_i \sim \text{Poisson} (\mu_i)$$

$$\ln Y_i = \beta_0 + \beta_1 X_{\text{WDepth}} + \beta_2 X_{\text{Coliform}} + \beta_3 X_{\text{Heterotrphs}} \qquad [5]$$

On the other hand, the highest pseudo R^2 (54.5%) was obtained for fall cholera peak in Bakergonj relative to water depth, chlorophyll-c, salinity, nitrate and phosphate. The coefficient for water depth was positive ($\hat{\beta}_1$ = 0.403; p value < 0.005, table 1), indicating a proportional relationship with cholera cases (Figure 2.6). Water depth explained approximately 20% of the model variance. A positive coefficient was observed for chlorophyll-c ($\hat{\beta}_2$ = 0.07; p value < 0.005, Table 1), since it is indicator of phytoplankton abundance, food for the zooplankton that carry vibrios as commensals, with concomitant increase in bacterial numbers. In addition, nitrate, phosphate and salinity were associated with cholera, with both phosphate and nitrate showing a significant negative coefficient with cholera, ($\hat{\beta}_4$ = - 0.00045; p value < 0.005 and $\hat{\beta}_5$ = - 0.021; p value < 0.005, respectively), confounding effects that difficult to explain.

The proposed model for Bakerganj fall peak is expected to have the following functional form:

$$Y_i \sim \text{Poisson} (\mu_i)$$

$$\ln Y_i = \beta_0 + \beta_1\, X_{WDepth} + \beta_2\, X_{Salinity} + \beta_3\, X_{CHLc} + \beta_4\, X_{Nitrate} + \beta_5\, X_{Phosphate} \qquad [6]$$

The proxy validation and observed cholera cases shown in Figures 2.6 and 2.7 suggest that, in regions with a dual cholera peak, two models may need to be developed for the two peaks, and subsequent combining both to one modeling framework.

Table 2.5 Seasonal models for cholera cases in Mathbaria and Bakergonj, Bangladesh

Model Parameters

	Intercept	Water Depth	CHLc	Hetero trophic	Air Temp	Salinity	Nitrate	Phosphate	Pseu do-R²
Mathbaria (spring peak)									70.9 %
Coefficient	8.98	-7.74	0.0768	-0.105	0.2122				
P-value	0.011	0.000	0.009	0.006	0.002				
SE	3.53	1.39	0.0293	0.0388	0.07				
95 Min	2.06	-10.47	0.019	-0.181	0.074				
% Max	15.9	-5.018	0.134	-0.029	0.349				
CI									
Bakergonj (fall peak)									54.4 %
Coefficient	2.99	0.403	0.07			0.0005	-0.0045	-0.021	
P-value	0.211	0.003	0.000			0.731	0.001	0.001	
SE	2.39	0.926	0.014			0.0015	0.0012	0.0012	
95 Min	-1.702	-1.412	0.0428			-0.0025	-0.0072	- 0.034	
% Max	7.695	2.219	0.0987			0.0036	-0.0018	-0.0089	
CI									
Bakergonj (spring peak)		Coliform							47.1 %
Coefficient	20.34	-13.5	-0.29	0.136					
P-value	0.003	0.002	0.008	0.018					
SE	6.96	4.26	0.1107	0.069					
95 Min	6.698	-21.86	-0.512	0.0009					
% Max	33.99	-5.14	-0.078	0.2718					
CI									
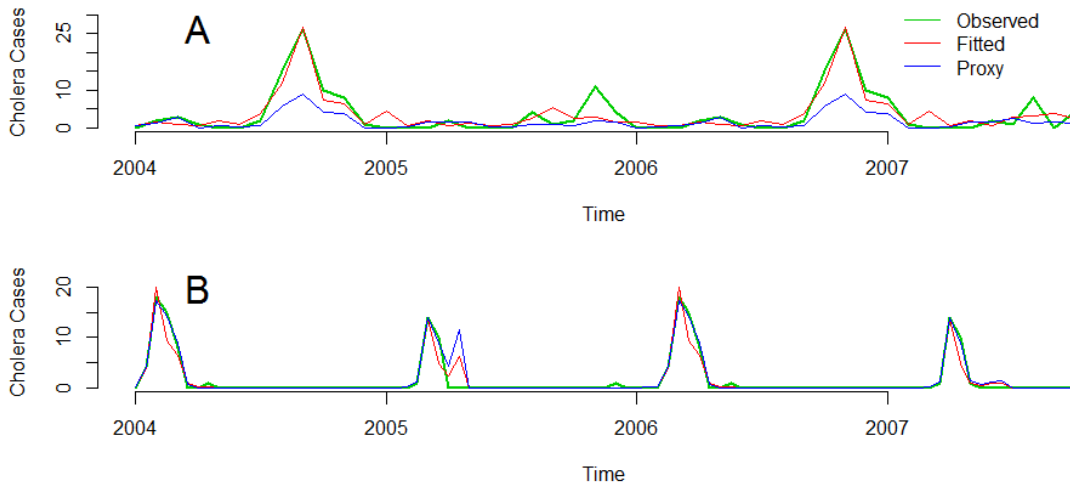
Figure 2.6 Epidemiological data of cholera seasonal peaks showing the temporal dynamics of observed cholera cases and models including fitted and proxy validation shown in green, red and blue respectively, (A) Bakergonj and (B) Mathbaria
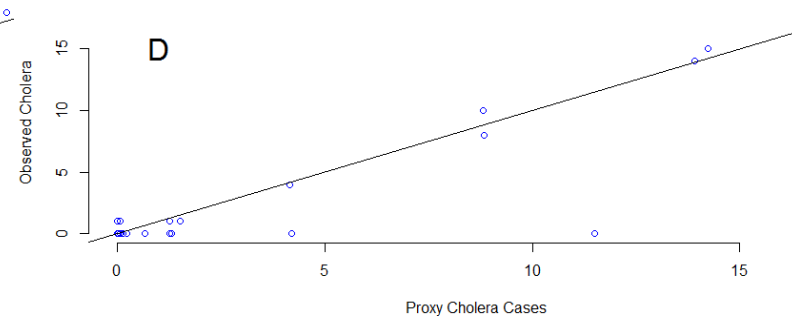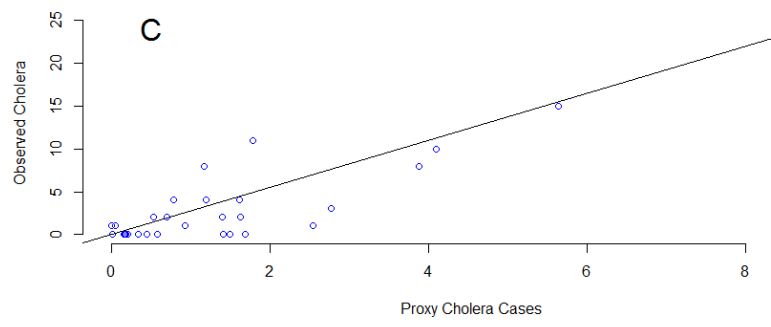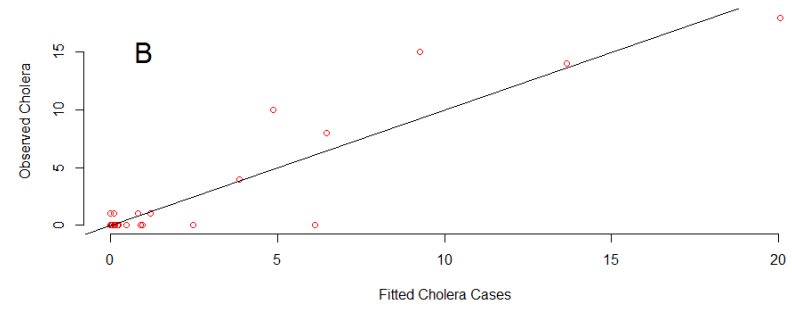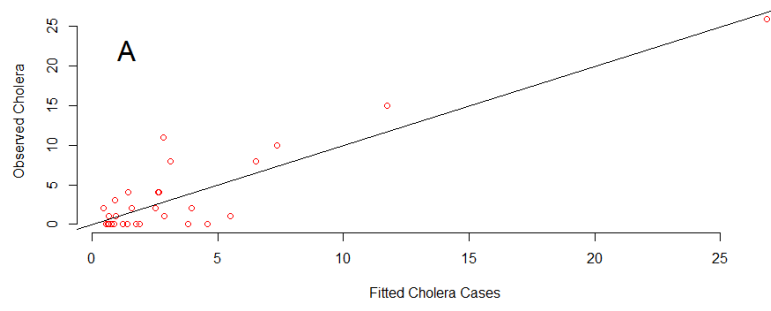
Figure 2.7 Epidemiological data of cholera seasonal peaks showing a scatterplot of observed cholera cases against simulated cholera cases by fitted model in red circles and cholera cases by proxy validation model in blue circles, Black line represents perfect agreement between simulated and observed cases. A & C: For Bakergonj and B & D: For Mathbaria

2.10 Discussion and Summary

The goal of this study was to extend the hypotheses of Akanda *et al*., 2009 by determining environmental factors associated with endemic cholera. Two endemic regions in the Bengal Delta were selected for study with data available for analysis. Ponds, in this study, are connected to large tributaries of the Ganges River prone to tidal intrusion of coastal seawater during spring season since river discharge during the spring season is extremely low. Mathbaria, a coastal town, has a single cholera peak in the spring, whereas Bakergonj located near the coast but inland has two seasonal peaks, one in the spring and another in the fall. The negative association observed in this study between water depth and spring cholera in Mathbaria is interpreted as an effect of coastal water intrusion, with contamination of inland water bodies accelerating cholera transmission by increasing salinity and nutrient concentration (Rao, 1973; Valsaraj, Rao, 1994). Akanda *et al.* (2009), using data for large-scale hydroclimatic processes, proposed that intrusion of coastal water during low river height creates conditions for cholera during the spring season in the inland regions of the Bengal Delta. This, followed by widespread flooding during high water level river periods, has the effect of cross contamination in effective water infrastructure, resulting in a fall peak of cholera cases. Our analysis for both Mathbaria and Bakergonj supports and strengthens this hypothesis. First, the Mathbaria modeling results are consistent with those of Akanda *et al.* (2009), namely that pond water depth is a major environmental factor associated with cholera, explaining 59% of the model variance, with a negative water depth coefficient. In Bakergonj, the spring peak is related to coastal intrusion, and the fall peak to widespread flooding and cross contamination rising from inadequate sanitation. Using seasonal proxy

46

models, a negative association was observed between spring cholera and water depth, but positive for fall cholera and water depth. To explain the microbiology of organisms in lacks; as the depth of water decreases, bacteria concentrate as well as the nutrients and alternate hosts such as algae and plankton. Then as the depth of water increases either due to high rainfall or increase in river discharge, the sediments, plankton and other organisms become resuspended along with the bacteria. Chlorophyll-c showed a positive relationship with cholera in both regions for all seasons. Higher values of chlorophyll-c with seasonal increase in number of cholera cases most likely reflects the role of plankton blooms and subsequent growth of zooplankton populations and related increasing *V. cholerae* population numbers (Colwell, 1996, Nalin *et al*., 1979, Rawlings *et al*., 2007). Two additional environmental variables, nitrate and phosphate appear to be related to fall cholera outbreaks in Bakergonj. The confounding effect of salinity will need to be investigated further. A key implication of this analysis is the potential for a cholera prediction model using information from large-scale hydroclimatic processes readily obtained by satellite remote sensing. The negative and positive relationships of water depth with cholera during the spring and fall cholera shown in Fig. 4 corroborate the finding of Akanda *et al*., (2009) and Jutla *et al*., (2013), who proposed intrusion of coastal seawater during the spring season initiates cholera, while a flood during the fall season accelerates cross-contamination of bacteria into the water where there is a faulty sanitation infrastructure.

Regional differences between the two locations confirm the importance of the local environment relative to large-scale hydrologic and climatic drivers of seasonal cholera dynamics. The vast space-time coverage provided by satellite remote sensing of

coastal and terrestrial ecosystems where *Vibrio cholerae* populations exist offers the potential of achieving spatially distinct prediction of impending cholera outbreaks in other countries of South Asia and also in Sub-Saharan Africa (Jutla *et al*., 2013). Stream flow from the major rivers of the Bengal Delta and remotely sensed plankton abundance in the coastal Bay of Bengal together explain over 75% of the variability of cholera in the Bengal Delta (Jutla *et al*., 2013). Thus, this predictive capability would allow deployment of an operational cholera warning system to identify vulnerable populations at the regional scale a few months in advance of a cholera epidemic. Key benefits could be achieved with minimal installation and operating cost, yet allowing timely implementation of preventive measures to contain the spread and magnitude of outbreaks. While results of this study are encouraging, future studies will benefit from biweekly sampling to validate and strengthen predictive capacity. In summary, timely prediction for cholera, coupled with access to clean water and adequate sanitary infrastructure, will benefit regions of the developing world with respect to public health.

2.11 Future Prospects

The most notable limitation in a local environmental study is the sampling process. Since sampling was biweekly in Mathbaria and monthly in Bakergonj, we have a better quality model for Mathbaria (Pseudo r^2 = 77.7%) and more power to predict the outcome variable (a cholera outbreak) while Pseudo r^2 = 48.5% for the Bakergonj model. Also we can see the validated model (Table 2.5) for the fall peak in Bakergonj does not show the effect of water depth, a major factor of our hypothesis. For future studies, a consistent biweekly sampling effort is strongly recommended.

Chapter 3

Trigger and Transmission of Sporadic Cholera Outbreaks in Epidemic Regions

3.1 Abstract:

Cholera remains a major public health threat in developing countries. Since the causative agent is *Vibrio cholerae*, which is native to the aquatic environment, it is unlikely that the disease can be eradicated. Therefore, developing hydroclimatology enforced disease models may be useful to predict or determine trigger and transmission mechanisms of cholera in epidemic and endemic regions. Part of South Asia, particularly in the Indus River Basin, periodically experiences sudden and sporadic outbreaks of cholera. Using data for a period of fifty years from 1950 to 2012, the episodic variability of cholera was found to be linked to hydroclimatic factors. Our results show that warmer air temperatures, followed by high rainfall, leads to increased risk of cholera in the region. By developing a mechanistic understanding of the disease system using models forced with hydroclimatic processes, it was observed that environmental factors play a role in triggering the disease whereas the human to environment route can increase the risk of cholera transmission in highly populated regions.

3.2 Introduction:

Cholera as a serious diarrheal deadly disease described in history books as pandemic in some countries, and continues to represent a serious public health threat for poorer countries around the world. There are few diseases in history that resemble the severity of cholera. The mortality rate of diarrheal diseases, including cholera, is about 2.2 million per year according to the global disease burden article published by WHO in 2008. There are about 3-5 million cholera cases and 100 000–120 000 deaths every year.

Recently, cholera has re-emerged as a major infectious disease with a major increase in global incidence. Under International Health Regulations, cholera is one of the three diseases that require notification. WHO in 1995 recorded the highest numbers of cholera cases in 94 notified countries in 1994. Cholera is endemic throughout the African continent, particularly West Africa, where sanitation and waste disposal either poor or non-existent. About 85% of officially mortified cholera cases occur in Africa. This trend was continued in 2006, with 99% of the total number of cases were reported globally in Africa (Kindhauser, 2003). Zuckerman *et al*., (2007) reported that there is an under estimation of the actual cholera burden on the Indian subcontinent due to lack of surveillance and under-reporting. The main triggers of a cholera outbreak in most developing countries are unsafe water supply and inadequate sanitation (Lee 2004). The International Institute for Population Sciences and Macro International 2007 reported about 73% of the rural Indian population does not have proper water disinfection, and about 74% do not have sanitary toilets. Available freshwater in India is expected to decrease by 2025, according to the Intergovernmental Panel on Climate Change 2007,

which is from 1,820 m^3 per capita to < 1,000 m^3, in response to the combined effects of population growth and climate change.

Figure 3.1 Estimated cholera incidences for entire populations in endemic countries (Ali *et al*., 2012 & WHO 2012 Bulletin)

Spread of cholera occurs through the fecal-oral route, from an infected human reservoir. However, evidence also shows that *V. cholerae* inhabits seas, ponds, and other aquatic environments (Colwell *et al*, 1977; Colwell *et al*, 1990), where it is capable of introduction into vulnerable populations. Miller *et al* (1985) identified two general routes for cholera transmission based on these reservoirs. The ***Primary transmission*** route occurs via marine water bodies where *V. cholerae* exists, spreading from estuarine environments to humans when humans have some form of contact with contaminated water or, alternatively, when they consume contaminated shellfish or aquatic plants . The ***Secondary transmission*** route is defined as the spread of cholera from an infected individual to susceptible individuals in a population. Improving our understanding of cholera transmission, by both primary and secondary routes, is critical for prevention purposes. Primary transmission (figure 3.2) is determined by many processes including

hydrological, ecological, and climatic, while secondary transmission is determined by many other processes such as number of contacts between humans, population size, birth rate and other factors. We recognize the importance of secondary transmission on cholera diffusion through a given population. However, *V. cholerae* is autochthonous in riverine, estuarine and coastal waters, living in a wide range of natural conditions in association with zooplankton, and crustaceans, so it is unlikely that *V. cholerae* can be eradicated from its natural habitat. Consequently, cholera transmission should be assessed with macro-environmental factors, including ecological, hydrological, and climatic, which vary from region to region.
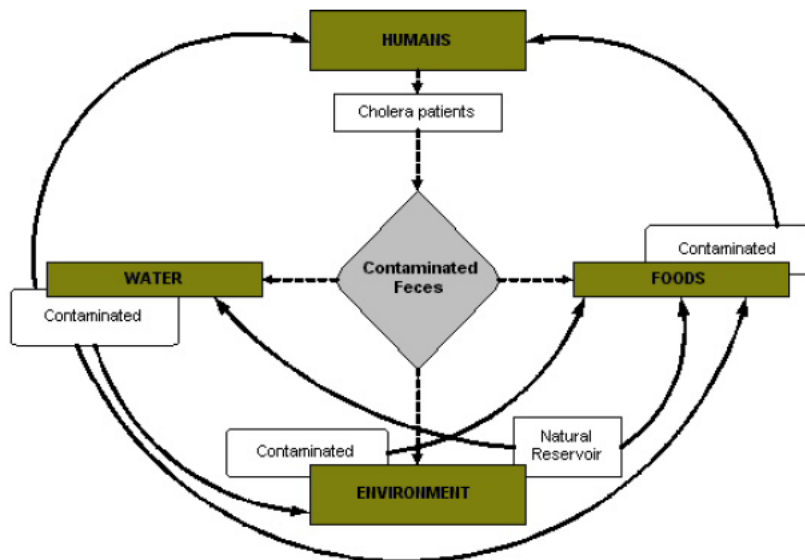
Figure 3.2. Major Transmission Routes of Cholera (modified from Mintz *et al*., 1994)

Minimization and/or prevention of cholera spread require early detection, prediction, and early warning of outbreaks where it may occur. Such early warning systems are considered as a first and important step to allow an affected country or population to

adapt to climate change. Successful climate change adaptation must occur on several scales, including physiological, behavioral, social, institutional, and organizational. We must draw adaptation strategies in response to temporal and spatial scales of climate variability for each region (Ebi *et al*., 2006). In India, there are several potential adaptation strategies to control infectious diseases, including: 1) Removal of vector breeding sites; 2) Reducing contact between vectors and humans; 3) Improving sanitation and drinking water operations and infrastructure; and 4) Monitoring pathogens, as well as disease burden. An additional and important potential adaptation strategy includes monitoring floods and heat waves (Bush *et al*., 2011), which are addressed in this chapter.

Cholera outbreaks in New Delhi, India show a single changing seasonal peak. The peak can occur at any time from April through September (Balakrish, 2007). Cholera outbreaks have been associated with a wide range of environmental variables, such as sea surface temperature (SST) (Lobitz *et al*., 2000; Cash *et al*., 2008), sea surface height (Lobitz *et al*., 2000), monsoon precipitation (Hashizume *et al*. 2008), coastal plankton (Magny *et al*. 2008; Emch *et al*. 2008; Lobitz *et al*., 2000; Tamplin *et al*., 1990), air and water temperature (Islam *et al*. 2009; Huq *et al*. 2005), and coastal salinity (Miller *et al*., 1982).

Existence of such processes for cholera and other complementary observations motivate us to explore the utility of satellite-derived macro-environmental variables to develop a cholera prediction model. Perspectives that prompt use of satellite-acquired data include the following: (i) Almost all cholera outbreaks originate near coastal areas, including the reemergence of cholera in Latin America in 1991 (Jutla *et al*., 2010, Magny

et al., 2008; Emch *et al*., 2008; Lobitz *et al*., 2000; Tamplin *et al*., 1990), whereas New Delhi is a noncoastal region; (ii) Remote sensing provides unprecedented coverage of space-time measurements of many environmental factors around the world (Uz and Yoder, 2004; Jutla *et al*. 2011); and (iii) The sporadic seasonality of cholera outbreaks in New Delhi. Our main objective here was to develop a cholera prediction model using remote-sensing information with two months' prediction lead time; and suggest a plausible pathway by which the variables used for development of the cholera prediction model may provide an explanation for an environment conducive to cholera outbreak and transmission.

3.3 Materials and Methods:

The current data set contains large-scale environmental variables for New Delhi (January, 1999, to December, 2007), in India (Figure 3.3); Two environmental variables were collected from New Delhi 1) rainfall data: from the Global Precipitation Climatology Centre and monthly mean at 0.5 degree Lat-Long spatial resolution on a global grid; and 2) land surface temperature: from NCEP global reanalysis, monthly mean at 2 degree Lat-Long spatial resolution.

Figure 3.3 shows the location of both New Delhi and Chennai (Dr. de Magny)

New Delhi is the capital of India. It has many government branches of the judicial, legislative and executive of India. New Delhi is one of the eleven districts of

Delhi National Capital Territory and situated within the metropolis of Delhi. It has about 1% of the population of Delhi metropolis. Delhi metropolitan is considered to be the largest city in India and the world's second most populous city with a population of about 22 million in 2011. It is ranked as the second wealthiest city in India after Mumbai.

3.4 Statistical analysis:

Clinical cholera case and environmental data were collected from Jan, 1999, to December, 2007. Cholera cases were collected on a monthly pattern from New Delhi, India. A generalized linear model (GLIM) with logistic distribution and log link was used to model the data and build a prediction model for cholera, following the method of Cameron and Trivedi (1998). Choice of model was based on the overall regression Chi-square statistic, Goodness-of-Fit Tests including Deviance, Pearson, Hosmer-Lemeshow and Measures of Association including Concordant, Discordant, Somers' D, Goodman-Kruskal Gamma and Kendall's Tau-a as well as low error. We also examined the correlations between cholera cases and environmental factors and used stepwise regression to identify the impact of the significant factors. In a logistic regression, we retained the variables that were significant in the stepwise regression as well as their estimated effective direction that was consistent with our expectation. The $\log_{10}$ occurrence of cholera cases in humans as the outcome variable, were regressed against the predictors at various months lags periods.

The multiple Poisson regression model (GLIM) for k predictors is as follows:

$$Y_i \sim \text{Bern}(\pi_i)$$

$$ln\left[\frac{\pi_i}{1-\pi_i}\right] = \beta_0 + \beta_1 X_{1,\tau-\tau1} + \beta_2 X_{2,\tau-\tau2} \ldots\ldots + \beta_k X_{k,\tau-\tau k} + e_\tau, \tau \geq \max \{\tau_1, \tau_2, .. \tau_k\}$$

Equation (1)

Where k predictors are X_1, X_2, ……., X_k and β are the corresponding model coefficients. The lags (months before the outbreak happen) for the *k* predictors are τ_1, τ_2, ….. τ_k. Y_t was assumed to follow a binomial distribution. $log\left[\frac{\pi_i}{1-\pi_i}\right]$ is the natural log of the odds ratio of $Y_i=1$ versus $Y_i = 0$. π_i is the probability of a 1 (the proportion of 1's, the mean of Y): *Vibrio cholerae* being an infective agent. The model choice was based on: (i) the best fit with the data; (ii) prediction purpose; (iii) the difficulty to estimate the actual count of cholera cases; and (v) explore the relationship between the environmental factors. The model fit is conducted for the seasonal data set and is examined by comparing observed and predicted numbers of cases in each area. Results from these analyses will provide new knowledge of how macro-environmental factors of sporadic cholera outbreaks in non-endemic region can aid in the development of a cholera prediction model.

3.5 Remote Sensing:

Generally in modern studies, remote sensing refers to detection and classification of objects either on the surface of earth, or in the atmosphere or oceans by using aerial sensor technologies. In 1978, the first oceanic remote sensing device for the Coastal Zone Color Scanner was launched on Nimbus7, and this was followed by the SeaWiFS mission in 1997. SeaWiFS has been used extensively in several studies to measure chlorophyll, e.g., in oceanic processes (Tang *et al*., 2003; Yoder *et al*., 1987; Danling *et al*., 2002,

Yuras *et al* 2005), in land-ocean interactions (Lopez and Hidalgo, 2009; D'Sa and Miller, 2003; Jutla *et al*., 2009a), and in assessment of coastal pollution (Chen *et al*., 2007). There are about eight channels of SeaWiFS at: 412, 443, 490, 510, 555, 670, 765, and 865 nm (nanometers: 1μm = 1,000 nm), each with bandwidths of 20 or 40 nm (O'Reilly *et al*., 2000).

Monitoring and quantifying large-scale environmental factors is done perfectly by satellite remote-sensing techniques. It is the preferred way when *in-situ* data is difficult and prohibitively expensive to obtain. In chapter 2 study, we observed a correlation between large scales; remote sensing-derived river discharge and in-situ water depth to draw an explanation of cholera outbreak in Bangladesh. Lobitz *et al*., (2000) suggested that cholera is influenced by climatic changes, which can be indirectly measured using satellite imagery. They illustrated that sea surface temperature (SST) and sea surface height (SSH) in the Bay of Bengal were associated with temporal fluctuations of cholera in Dhaka, Bangladesh from 1992 to 1995.

Remote sensing measurements of many relevant environmental variables (e.g., sea surface temperature (SST); sea surface height; rainfall; chlorophyll) are used extensively to develop our understanding of the possible controls of a cholera outbreak in the coastal regions and its links to terrestrial hydrology. In various ocean basins across the globe, both satellite measured chlorophyll and SST show an inverse relationship (Smyth *et al*., 2001, Uz and Yoder, 2004, Legaard and Thomas, 2006). However, In Bay of Bengal, a positive relationship is observed between phytoplankton and SST (Lobitz *et al*., 2000; Chaturvedi, 2005; Emch *et al*., 2008; Magny *et al*., 2008). Our preliminary

analyses, using remote sensing data, suggest that land surface temperature is a major player for cholera outbreak in New Delhi, India.

3.6 Shifts in Seasonality of Cholera in New Delhi, India:

Over six decades cholera patterns have been studied extensively to show seasonal changes in India. The highest *V. cholerae* infection was in April and May, but recently it has been shifted to September and October with two peaks annually (Rogers 1926 and 1928). In 1950s, Politzer (1959) described New Delhi as cholera free region for a considerable period of time. In about 56 years span, Delhi has transformed from free to rare to seasonal cholera occurrence region and finally to region that have cholera cases throughout the year. Recently, Sharma *et al.*, (2007) had this striking observation about cholera outbreak in New Delhi. They stated that cholera has a changing seasonality. In New Delhi cholera spread is related to its ecosystem and other physical environmental factors that favors *V. cholerae* proliferation (Rogers 1928). However, cholera incidence has not been correlated well with many conducive environmental factors such as rainfall, temperature, and relative humidity.

Our preliminary results from New Delhi data that go back to 1999 through 2008 show the changing seasonality phenomenon described by other studies (Figure 3.4). The figure shows that New Delhi has a single cholera outbreak peak throughout the year but with changing seasonality. There was a peak in May for 1999 and 2000 while cholera peak was on July for 2001 and 2002. Sharma *et al.* (2007) show the changing seasonality trend in his study. They observed that during 2003, Delhi had cholera peak during April while 2004 and 2005 cholera pattern occurred throughout the year. They stated that

cholera pattern during 2004 and 2005 was completely different than Delhi cholera previous trends.
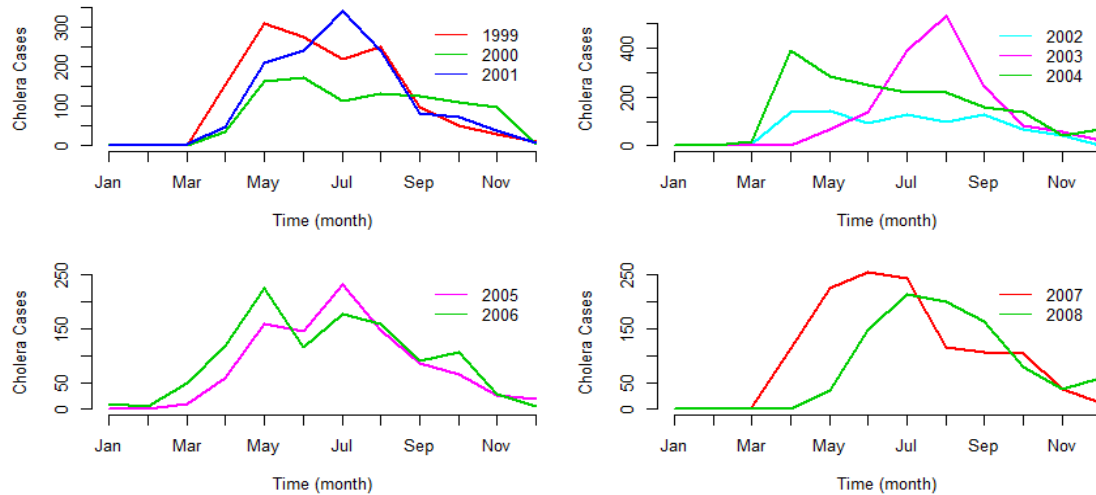
Figure 3.4 Epidemiological data of cholera seasonal peaks showing the temporal dynamics of observed cholera cases from 1999 through 2008 in New Delhi, India.

Our results show similar trend as those of Sharma *et al*. (2007) during 2004 and 2005 that cholera occurs throughout the year. In 2006, cholera peaks occurred throughout the year, but with the highest peak occurring during March. Both 2007 and 2008, cholera has mostly a single peak during June and July respectively. We further calculated the average of cholera peaks through the studied years (1999-2008). Figure 3.5 shows that the highest cholera incidence occurs in July. However, cholera is present from the beginning of May.
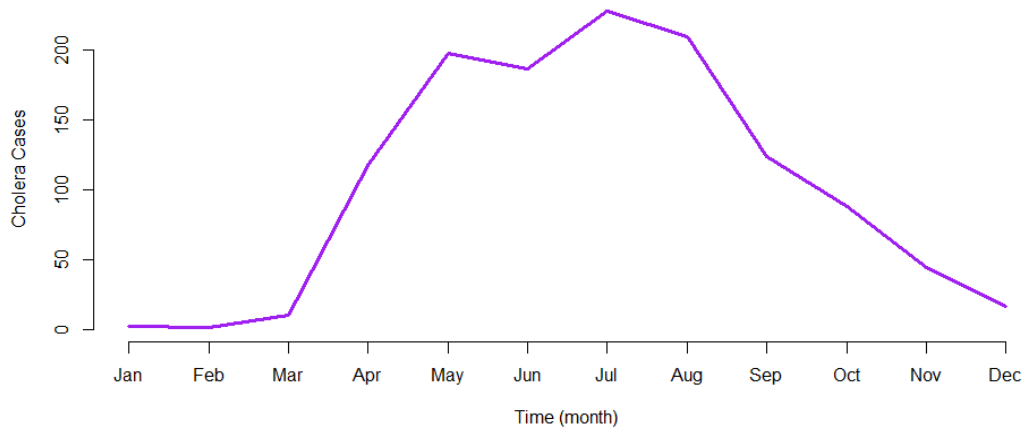
Figure 3.5 The average of Seasonal cholera peaks on monthly scale from 1999 through 2008 in New Delhi, India.

Lastly, wavelet analysis is investigated and run for three times over each signal (figure 3.8). The analysis has two main steps; the first step is investigating ***the larger frequency domain*** cycles in the time-series that is represented between 4 months (or 0.3-years) to 5 years. The second step is splitting the larger frequency domain into ***two frequencies***: 1) frequencies from 4 months (or 0.3-years) to 2-years to focus on annual cycle or cycle with higher frequency and 2) frequencies from 2-years to 5-years focusing on inter-annual cycles to avoid possible effect of stronger cycles that may mask others.
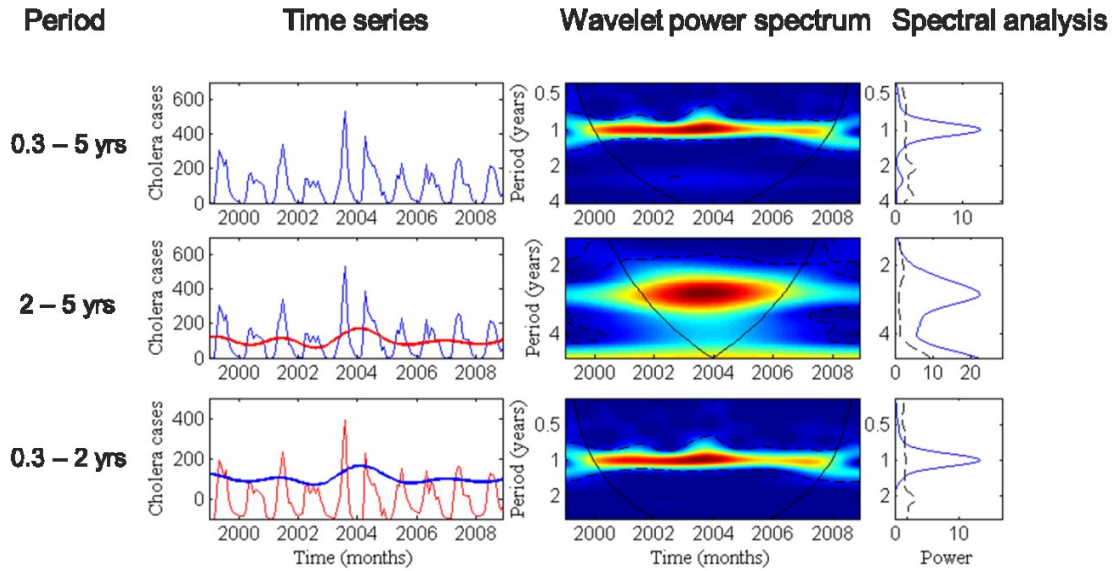
Figure 3.6 Wavelet analyses for the New Delhi cholera cases time-series. The Left panel represents the time-series; the middle panel is the wavelet power spectrum. The spectrum has the period of the cycle unit is expressed in years on the Y-axis and the time on the X-axis. The wavelet power, in other words the detection of a frequency in the time-series, is coded from dark blue, no frequency detected, to dark red frequency strongly detected. As the power expresses a correlation between a specific frequency and the time-series, the significance of the correlation is tested. When the correlations are continuously significant, they are delimited by the dashed line. (Analysis is done by Dr. Magny)

In the larger frequency domain step, a light blue to dark red significant band of frequencies is detected. This band of high correlations is centered and stationary on the 1 year period (illustrated on the 1st and 3rd row of graphics, 2nd column). This will indicate we have a single cholera outbreak peak every year. In the 2nd row, the row that explore only if 2- to 5-years cycles may be present in the time-series, the analyses show a big patch of correlation centered on 3 years period. Interpreting that in addition to have an outbreak every year, cholera outbreaks are more intense every 3 years. However, we need to be cautious when making that inference since the length of the time-series is short. When we are tracking cycle that appears every 3 years over a period of time of only 10 years, which means we are observing it 3 times. The significance of phenomenon observed only three times reaches the statistical limits. If it was observed over 30 years, it would to provide 9 observations that would offer robust statistics.

3.7 The main Triggers of Delhi Cholera

Cholera disease impact and spread is greatly influenced by the environmental factors especially with season and space. In any ecosystem, the environmental factors represent the biological, chemical and physical factors. Cholera spread in India was investigated in relation to its ecosystem, which favors *V. cholerae* (Banerjee B, Hazra J., 1974). In addition to these environmental factors, there are other important factors such as: social customs, human behavior, drainage and economic status that were studied extensively in terms of diarrheal disease spread and persistence (Rajendran *et al.*, 2008).

In Bengal, climate is a very important factor for *V. cholerae* persistence and spread. The relationship of cholera incidence with SST remains a subject of interest. Several studies indicate that cholera outbreaks are related to coastal SST (Colwell, 1996;

Bouma and Pascual, 2001; Magny *et al*., 2008). The first indication of a statistically significant association of the SST cycle with annual bimodal peaks in cholera outbreaks were demonstrated qualitatively by Lobitz *et al.,* (2000). However, Lobitz *et al.,* (2000) did not cite the strength of the correlation (Table 3.1). Since the work of Lobitz *et al.,* (2000), several studies have reported varying levels of association of cholera incidence with coastal Bay of Bengal SST. Bouma and Pascual (2001) reported that cholera in Bangladesh is moderately related with coastal SST, while Emch *et al.,* (2008) and De Magny *et al.,* (2008) suggested that there is no relationship between cholera and coastal SST but strong relationship with chlorophyll. Our current study represented in figures 3.6 and 3.7 (both seasonal and time series graphs) shows that air temperature and cholera incidence are highly correlated ($r = 0.73$, $P < 0.05$). On the other hand, rainfall shows moderate significant correlation with cholera incidence in New Delhi ($r = 0.483$, $P < 0.05$).

Table 3.1 SST- cholera analysis for Bangladesh

Author	Year	Type of analysis	Finding
1. **Lobitz *et al., *(2000)**	1992-1995	Qualitative	Strong relationship
2. **Bouma & Pascual (2001)**	1841-1940	0.3-0.5 (Quantitative)	Moderate relationship
3. **Emch *et al., *(2008)**	1985-2003	0 (Quantitative)	No relationship
4. **De Magny *et al., *(2008)**	1997– 2006	0 (Quantitative)	No relationship
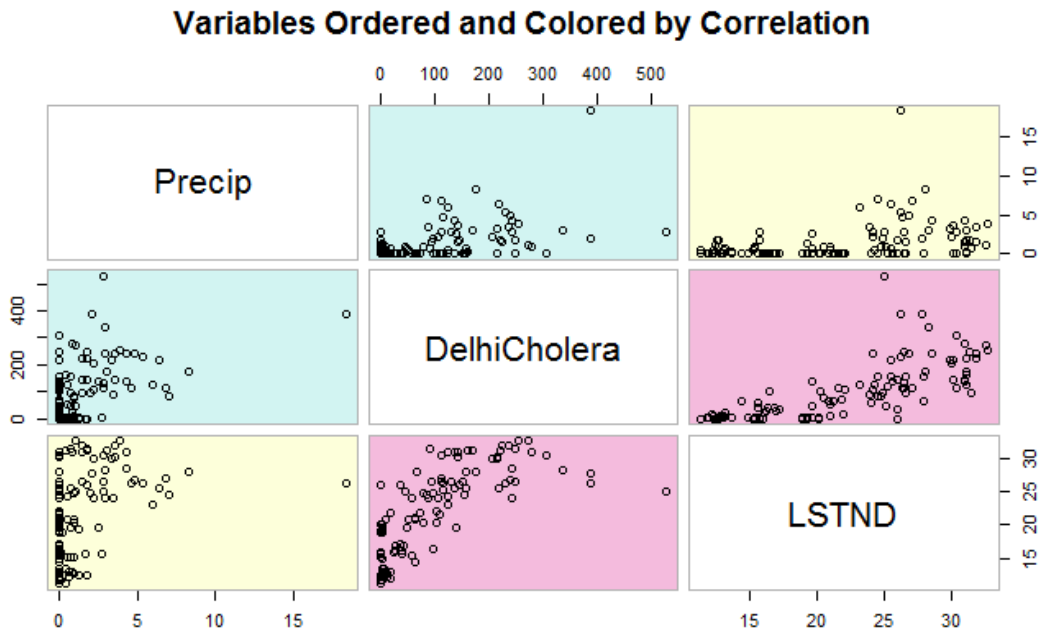
Figure 3.7 A correlation matrix for Land Surface Temperature, Precipitation and New Delhi Cholera (The highest correlation variables are closest to the diagonal)
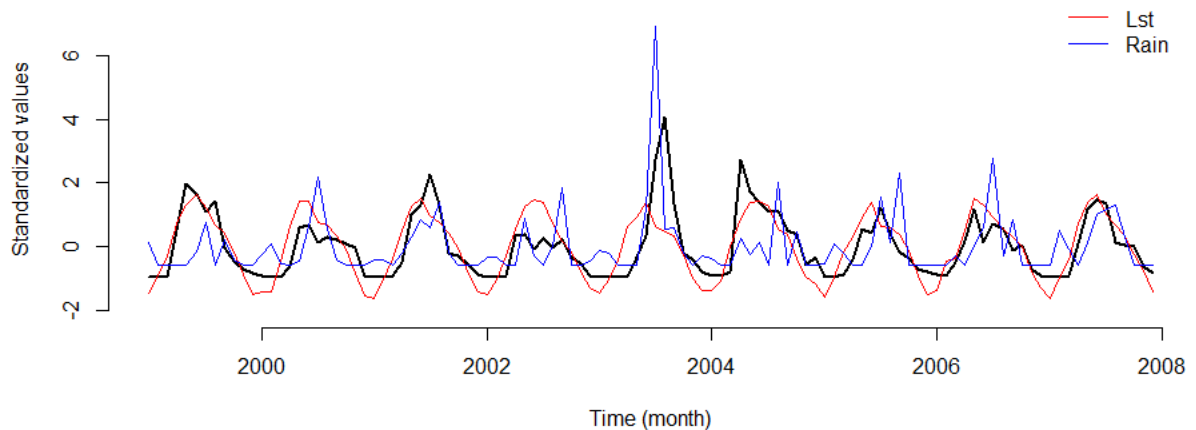
Figure 3.8 Seasonal correlation between Land Surface Temperature, Rainfall and cholera cases from 1999 to 2008 in New Delhi, India

We collected land surface temperature over the past 60 years and observed elevation of Delhi temperature compared to the 1950s through 1990s (Figures 3.8 and 3.9). The above studies show a marked difference of opinion in terms of understanding the relationship between SST and cholera outbreaks. The question remains; *how does air temperature affect cholera outbreaks in South Asia? And more specifically, how does a non-infected region to become cholera endemic?*

Plankton serves as the ecological niche for cholera bacteria (*Huq and Colwell, 1996*). Therefore, a possible role for zooplankton in causing cholera outbreaks has been emphasized in several studies (*Colwell, 1996; Huq and Colwell,* 1996; *Emch et al., 2008; Constantine de Magny et al., 2008*). With availability of chlorophyll data, a surrogate for zooplankton and measured by satellite, some of the recent studies have attempted to correlate chlorophyll values in the coastal Bay of Bengal with Bengal cholera incidence (*Emch et al., 2008*; *Constantine de Magny et al., 2008*), but did not elaborate which cholera outbreak peaks are related to zooplankton. Suggested physical roles of the coastal plankton processes remain largely unexplained in these studies. However, Lobitz *et al* (2000), Jutla *et al*., (2011) suggested a physical role for SST in cholera outbreaks through the zooplankton link. Their hypothesis is that the rise in SST may increase phytoplankton blooms and; followed by zooplankton blooms and bacteria are associated with zooplankton, therefore, cholera outbreaks.
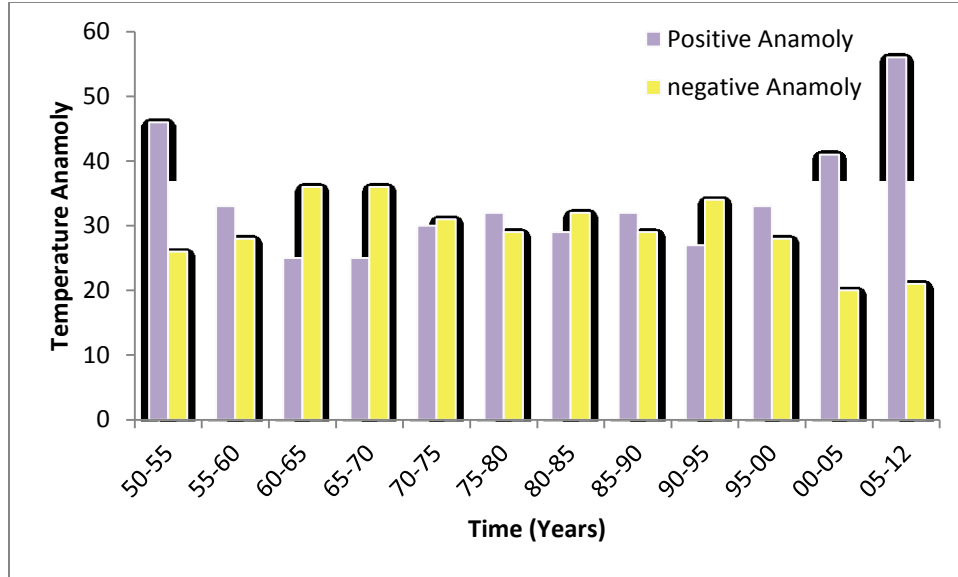
Figure 3.9 Calculated Land surface temperature anomalies over the past 60 years from 1950 till the present in New Delhi, India

To improve our understanding of how sporadic cholera outbreaks transmission dynamics are affected by large environmental processes in New Delhi and quantitatively estimate the effect of each environmental driver, we conducted logistic regression analysis for both land surface temperature and rainfall at different lag periods (Table 3.3). After examining both Goodness-of-Fit Tests and Measures of Association (Table 3.3) for several models combinations including: 1) M1: LST; 2) M2: LST and Rainfall; 3) M3: LST at one month lag and Rainfall; 4) M4: LST at one month lag and Rainfall at one month lag; 5) M5: LST at two months lag and Rainfall; and 6) M6: LST at two months lag and Rainfall at one month lag (table 3.3 and figure 3.14), **It was proposed Cholera Prediction Model (CPM) for New Delhi take the following functional form:**

$$Y_i \sim \text{Bern}(\pi_i)$$

$$\ln\left[\frac{\hat{\pi}(Cholera)_i}{1 - \hat{\pi}(Cholera)_i}\right] = \hat{\beta}_0 + \hat{\beta}_1(LST)_{i1} + \hat{\beta}_2(Rain)_{i2} \qquad \text{(Equation 2)}$$

$$\ln\left[\frac{\hat{\pi}(Cholera)_i}{1-\hat{\pi}(Cholera)_i}\right] = \hat{\beta}_0 + \hat{\beta}_1(LST)_{i1,t-1} + \hat{\beta}_2(Rain)_{i2}$$ (Equation 3)

In New Delhi model (M1), land surface temperature is the main major environmental signatures responsible for cholera outbreaks since it is perfectly associated with 87.5 % (concordant) of the model data. The concordant, one of the measures of association of the model, the percentage is increased to 88.7%, when rainfall is added to the model equation (equation 2). For prediction purpose, we have chosen model 3 that include land surface temperature at one month lag period and rainfall. This model is associated with about 74.6% of the data Figure (3.11) and Table (3.3). A significant positive coefficient is observed in equation 2 between cholera cases and land surface temperature ($\hat{\beta}_1$ = 0.317; P value < 0.005). Moreover, when we examined the cholera model (equation 3, Table 3.2), there is a significant positive coefficient ($\hat{\beta}_2$ = 0.201; p value < 0.005). Both models M2 and M3 fit well with the observed data since the p value for all deviance, person and Hosmer-Lemeshow are greater than 0.05 (table 3.2).

Table 3.2 Summary of the best models obtained for New Delhi in India

		Model Parameters					
	Intercept	LST	Rainfall	LST (t-1)	Goodness-of-Fit Tests		
M2							
Coefficient	-3.88	0.317	0.2		44.31 1.000	Deviance	P-Value
P-value	0.000	0.000	0.076		48.75 1.000	Pearson	P-Value
SE	1.78	0.1120	0.115		0.89 0.999	Hosmer-Lemeshow P-Value	
M3							
Coefficient	-0.2		0.201	0.0895	58.2 1.000	Deviance	P-Value
P-value	0.000		0.032	0.086	76.67 0.945	Pearson	P-Value
SE	1.14		0.131	0.0539	13.66 0.091	Hosmer-Lemeshow P-Value	

We further investigated the probability of expedience between precipitation and cholera incidence (Figure 3.9) without any lags to demonstrate that they are moderately correlated. Figure 3.12 shows how well all the models, including both M2 and M3, fit with the observed number of cholera cases. Models 1 and 2 probabilities perfectly match with cholera occurrence. Although model 3 is not perfectly matching, as 1 and 2, but it still can be used as a prediction model to describe the relationship between temperature and rainfall.
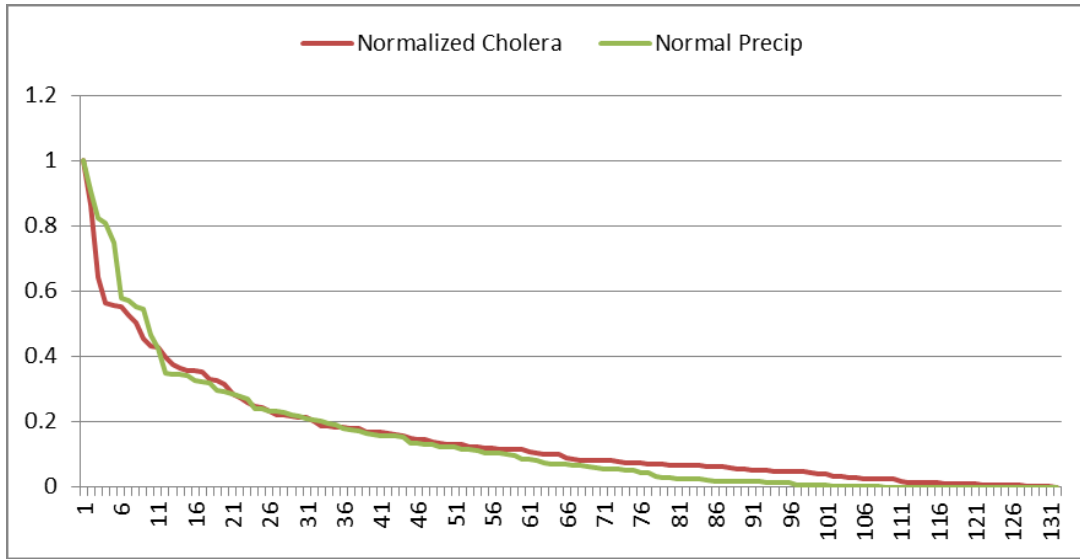
Figure 3.10 The relationship (Probability of Expedience) between cholera and precipitation in New Delhi, India from 1999 to 2007
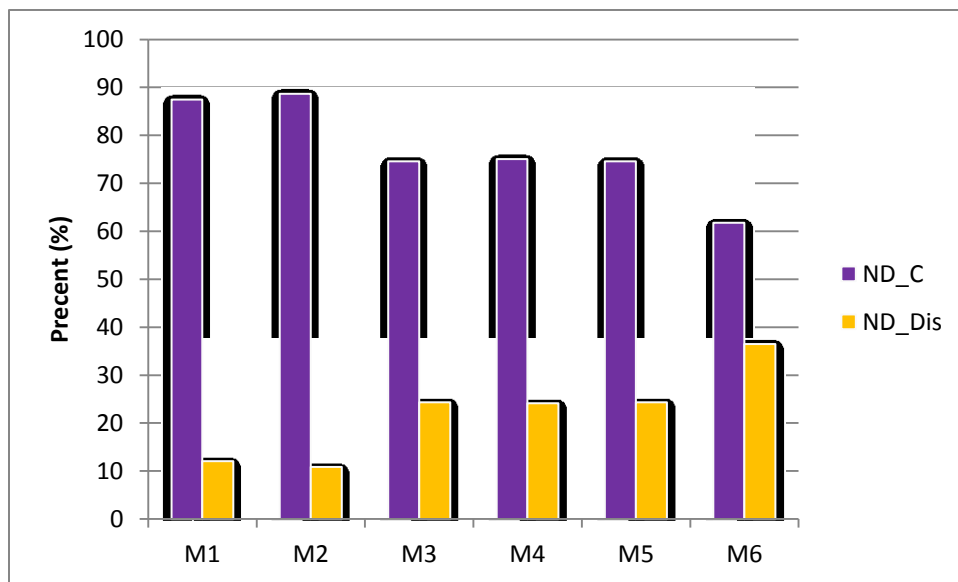
Figure 3.11 Model's performance for New Delhi illustrating measures of association including both concordant and dis-concordant.
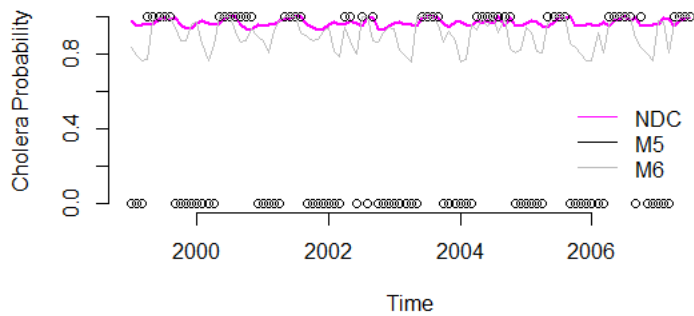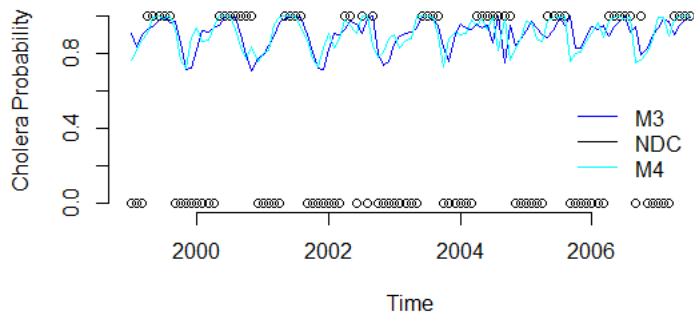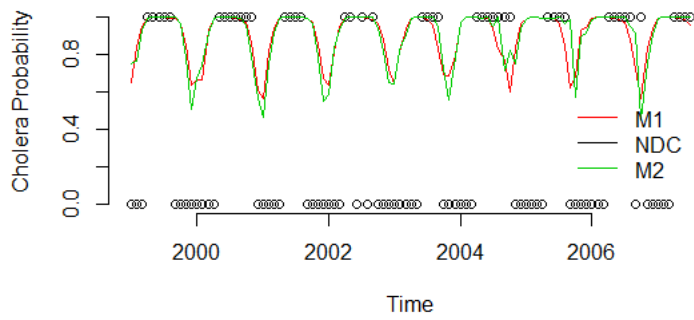
Figure 3.12 Epidemiological data of cholera seasonal peaks showing the temporal dynamics of observed cholera cases and fitted models for M1 through M6 shown in black, red, green dark blue, light blue, pink and grey respectively in New Delhi, India

Table 3.3 Measures of Association indicators and Goodness-of-Fit Tests for predicting cholera occurrence in New Delhi, India and the 9 other stations

New Delhi, India (The main site)	M1: T(t)		M2: T(t), R(t)		M3: T(t-1), R(t)		M4: T(t-1), R(t-1)		M5: T(t-2), R(t)		M6: T(t-2), R(t-1)	
Deviance P-value	47.46	1.000	44.31	1	58.2	1	55.89	1	56.71	1.000	60.94	0.999
Pearson P-value	58.39	1.000	48.75	1	76.67	0.945	67.97	0.991	76.98	0.943	84.80	0.827
Homser-Lemeshow P-value	4.78	0.781	0.89	0.999	13.66	0.091	2.95	0.938	4.17	0.842	7.41	0.493
Concordant	87.5		88.7		74.6		75.1		74.6		61.8	
Discordant	12.1		10.9		24.4		24.2		24.4		36.5	
Somers' D	0.75		0.78		0.5		0.51		0.5		0.25	
Goodman-Kruskal Gamma	0.76		0.78		0.51		0.51		0.51		0.26	
Kendall's Tau-a	0.14		0.14		0.09		0.09		0.09		0.05	
Central Delhi												
Deviance, P-value	257.02	0.642	252.91	0.693	271.16	0.384	259.45	0.585	291.12	0.130	276.20	0.305
Pearson, P-value	293.27	0.120	302.03	0.058	267.22	0.450	267.76	0.441	264.47	0.498	266.81	0.457
Homser-Lemeshow, P-value	12.55	0.128	6.29	0.615	6.75	0.564	5.01	0.756	5.83	0.666	8.11	0.422
Concordant (%)	76.3		77.8		71.9		75.8		64.7		69.4	
Discordant (%)	23.3		21.8		27.7		23.8		34.7		29.9	
Somers' D	0.53		0.56		0.44		0.52		0.3		0.39	
Goodman-Kruskal Gamma	0.53		0.56		0.44		0.52		0.3		0.4	
Kendall's Tau-a	0.21		0.22		0.17		0.2		0.12		0.15	
Central Lahore												
Deviance, P-value	345.12	0.061	342.78	0.067	371.19	0.006	374.11	0.004	390.94	0.001	395.10	0.000
Pearson, P-value	299.88	0.588	301.08	0.553	308.26	0.437	307.74	0.445	308.41	0.435	308.76	0.429
Homser-Lemeshow, P-value	9.10	0.334	11.96	0.153	9.14	0.331	4.61	0.798	9.63	0.292	11.91	0.155
Concordant (%)	74.3		75.5		70.1		68.5		61		55.6	
Discordant (%)	25.5		24.2		29.4		31.1		37.9		43.2	
Somers' D	0.49		0.51		0.41		0.37		0.23		0.12	
Goodman-Kruskal Gamma	0.49		0.51		0.41		0.38		0.23		0.13	

Kendall's Tau-a	0.23	0.24	0.19	0.17	0.11	0.06

South Ludhiana												
Deviance, P-value	245.54	0.992	239.93	0.995	245.25	0.991	246.89	0.989	257.37	0.964	258.67	0.959
Pearson, P-value	307.65	0.384	312.55	0.297	297.60	0.528	300.69	0.478	300.14	0.487	306.34	0.388
Homser-Lemeshow, P-value	12.38	0.135	5.32	0.723	9.73	0.284	7.22	0.513	8.60	0.377	2.30	0.970
Concordant (%)	71.7		73.5		73		70.9		66.9		64	
Discordant (%)	28		26.1		26.5		28.5		32		35.2	
Somers' D	0.44		0.47		0.47		0.42		0.35		0.29	
Goodman-Kruskal Gamma	0.44		0.48		0.47		0.43		0.35		0.29	
Kendall's Tau-a	0.12		0.13		0.13		0.12		0.1		0.08	
North Sialkot												
Deviance, P-value	201.27	0.996	202.82	0.995	209.30	0.987	210.91	0.984	209.22	0.987	212.05	0.981
Pearson, P-value	255.09	0.522	258.97	0.454	260.60	0.426	258.68	0.459	259.63	0.442	259.39	0.447
Homser-Lemeshow, P-value	7.16	0.519	7.55	0.479	7.67	0.467	18.41	0.018	5.16	0.741	8.40	0.395
Concordant (%)	62.4		64.7		57.5		56.2		59.3		54.1	
Discordant (%)	37.3		34.8		41.5		43.1		39.7		45	
Somers' D	0.25		0.3		0.16		0.13		0.2		0.09	
Goodman-Kruskal Gamma	0.25		0.3		0.16		0.13		0.2		0.09	
Kendall's Tau-a	0.06		0.07		0.04		0.03		0.05		0.02	
North Rawalpind												
Deviance, P-value	236.37	0.861	236.35	0.851	271.44	0.300	272.90	0.279	287.99	0.112	289.85	0.098
Pearson, P-value	230.30	0.915	230.83	0.903	262.74	0.441	260.92	0.472	263.67	0.425	264.08	0.418
Homser-Lemeshow, P-value	6.87	0.551	7.55	0.479	4.41	0.81	5.73	0.678	4.32	0.828	16.61	0.034
Concordant (%)	78.8		78.9		69.3		68		57.2		54	
Discordant (%)	20.8		20.8		30.2		31.5		41.8		45	
Somers' D	0.58		0.58		0.39		0.37		0.15		0.09	
Goodman-Kruskal Gamma	0.58		0.58		0.39		0.37		0.16		0.09	
Kendall's Tau-a	0.22		0.22		0.15		0.14		0.06		0.03	

West Peshawar

Deviance,	P-value	235.14	0.983	234.78	0.981	255.81	0.867	256.30	0.862	263.56	0.778	263.37	0.781
Pearson,	P-value	281.57	0.513	276.16	0.587	292.91	0.315	284.35	0.450	299.72	0.224	304.28	0.173
Homser-Lemeshow, P-value		7.41	0.494	5.40	0.714	6.81	0.558	7.49	0.484	9.01	0.341	16.15	0.040
Concordant (%)		73.3		73.3		63.8		62.6		59.6		61.5	
Discordant (%)		26.3		26.3		35.8		36.8		38.2		36.9	
Somers' D		0.47		0.47		0.28		0.26		0.21		0.25	
Goodman-Kruskal Gamma		0.47		0.47		0.28		0.26		0.22		0.25	
Kendall's Tau-a		0.14		0.14		0.08		0.07		0.06		0.07	

West Ismail Khan

Deviance,	P-value	177.81	1.000	177.73	1.000	186.15	1.000	186.40	1.000	186.00	1.000	186.79	1.000
Pearson,	P-value	267.63	0.460	269.12	0.418	269.07	0.419	268.52	0.428	270.76	0.391	268.41	0.430
Homser-Lemeshow, P-value		3.62	0.890	3.75	0.879	11.29	0.186	9.75	0.283	11.71	0.165	5.69	0.682
Concordant (%)		66.3		66.3		56.5		57.9		55.9		56	
Discordant (%)		33.2		33.1		42.6		41.4		43.3		42.9	
Somers' D		0.33		0.33		0.14		0.16		0.13		0.13	
Goodman-Kruskal Gamma		0.33		0.33		0.14		0.17		0.13		0.13	
Kendall's Tau-a		0.07		0.07		0.03		0.03		0.03		0.03	

West Multan

Deviance,	P-value	150.94	1.000	150.55	1.000	164.60	1.000	164.73	1.000	168.29	1.000	168.07	1.000
Pearson,	P-value	239.55	0.988	243.75	0.978	285.73	0.560	286.68	0.544	293.34	0.434	292.59	0.446
Homser-Lemeshow, P-value		4.70	0.790	9.29	0.318	10.27	0.246	15.46	0.051	5.19	0.737	8.16	0.418
Concordant (%)		73.8		73.9		64		63.6		55.5		55.5	
Discordant (%)		25.8		25.8		35.2		35.5		41.5		42	
Somers' D		0.48		0.48		0.29		0.28		0.14		0.14	
Goodman-Kruskal Gamma		0.48		0.48		0.29		0.28		0.14		0.14	
Kendall's Tau-a		0.08		0.08		0.05		0.04		0.02		0.02	

West Sirsa

Deviance,	P-value	218.81	0.855	218.79	0.845	227.19	0.729	226.56	0.739	235.87	0.581	234.63	0.603

Pearson, P-value	241.64	0.494	241.16	0.485	244.98	0.417	245.66	0.405	244.14	0.431	243.53	0.442
Homser-Lemeshow, P-value	14.17	0.077	13.33	0.101	12.49	0.131	15.68	0.047	5.12	0.745	12.34	0.137
Concordant (%)	69.8		69.6		65		65.8		57.7		60.8	
Discordant (%)	29.8		30		34.3		33.8		41.5		38.4	
Somers' D	0.4		0.4		0.31		0.32		0.16		0.22	
Goodman-Kruskal Gamma	0.4		0.4		0.31		0.32		0.16		0.23	
Kendall's Tau-a	0.13		0.12		0.1		0.1		0.05		0.07	

3.8 Temperature/ rainfall theory validation

Jutla *et al*. 2013 established that both temperature and rainfall are important factors for cholera outbreaks and the finding is evident this analysis. Therefore, to have a second look at the Jutla *et al*. 2013 hypothesis, we further investigated the role of temperature and rainfall on cholera outbreaks at nine stations. The nine stations show a single cholera peak each year (Figure 3.13). After fitting a logistic multiple regression model, we observed that land surface temperature is the main significant environmental factor that has positive coefficient across the nine stations ($\hat{\beta}_1$ = 0.1021, 0.0734, 0.0728, 0.0383, 0.0916, 0.0638, 0.0457, 0.0830, 0.0610) P value < 0.005; for Central Delhi, Central Lahore, Submontane Ludhiana, North Sialkot, North Rawalpindi, West Peshawar, West Ismail Khan, West Multan, West Sirsa stations respectively). Across all nine stations, more than 60% of the data is concordant in model 1 that has land surface temperature as the only environmental factor. Figure (3.14) and Table (3.3) illustrate the concordant percentage for M1 are 87.5%, 76.3%, 74.3%, 71.7%, 62.4%, 78.8%, 73.3%, 66.3%, 73.8%, 69.8 for Central Delhi, Central Lahore, Submontane Ludhiana, North Sialkot, North Rawalpindi, West Peshawar, West Ismail Khan, West Multan, West Sirsa stations respectively). When we added the rain to model 2, the concordance percentage increased at most of the stations to be 77.8, 75.5, 73.5, 64.7, 78.9, and 73.9 for Central Delhi, Central Lahore, South Ludhiana, North Sialkot, North Rawalpind, and West Multan, respectively, while it still the same percentage for West Peshawar (73.3%) and West Ismail Khan (66.3%). The above information indicates that both land surface temperature and rainfall are important environmental factors that drive cholera outbreaks.

In order to investigate the relationship between them, we conducted several experiments to build models 3, 4, 5 and 6. **It was proposed that a Cholera Prediction Model (CPM) for all nine different stations in India and Pakistan that may take the following functional form:**

$$Y_i \sim \text{Bern}(\pi_i)$$

$$\ln\left[\frac{\hat{\pi}(Cholera)_i}{1 - \hat{\pi}(Cholera)_i}\right] = \hat{\beta}_0 + \hat{\beta}_1(LST)_{i1} + \hat{\beta}_2(Rain)_{i2} \qquad \text{(Equation 4)}$$

$$\ln\left[\frac{\hat{\pi}(Cholera)_i}{1 - \hat{\pi}(Cholera)_i}\right] = \hat{\beta}_0 + \hat{\beta}_1(LST)_{i1,t-1} + \hat{\beta}_2(Rain)_{i2} \qquad \text{(Equation 5)}$$

Figure 3.13 Seasonality of cholera in India including 8 different stations (1875-1900)
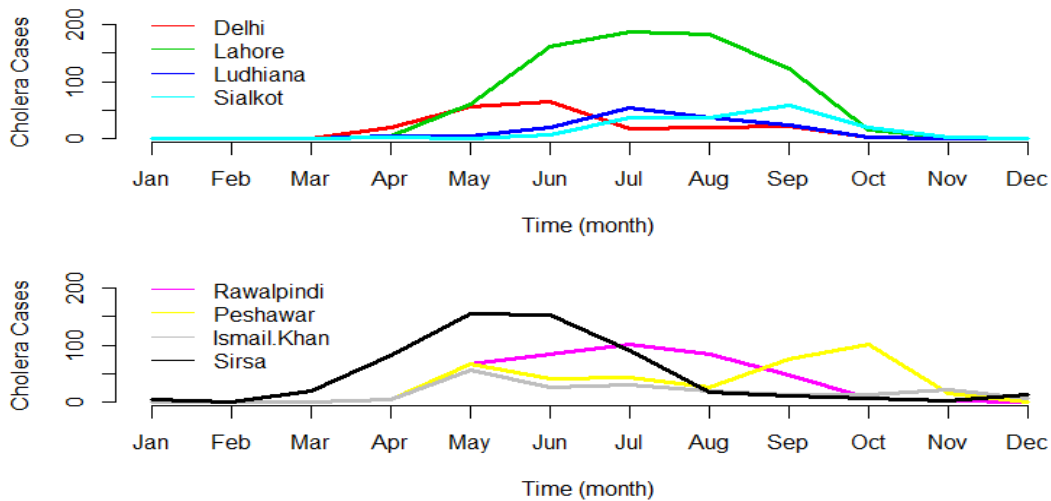
Of all the stations studied, M2 and M3 fit, well with the data, since the p value for deviance, Pearson, and Homser-Lemeshow are greater than 0.05. Although M3 showed a slightly lower concordance percentage for all the 9 studied stations, but it can be used as a good predictive model for cholera outbreak and it can describe the relationship between

the two main environmental parameters (Table 3.3). Figure 3.14 through Figure 3.22 show how well all the models, including both M2 and M3, fit with observed cholera cases. Models 1 and 2 probabilities match perfectly with cholera occurrence. Although, model 3 is not perfectly matching as 1 and 2, it can be used as a prediction model that can describe the relationship between temperature and rainfall.
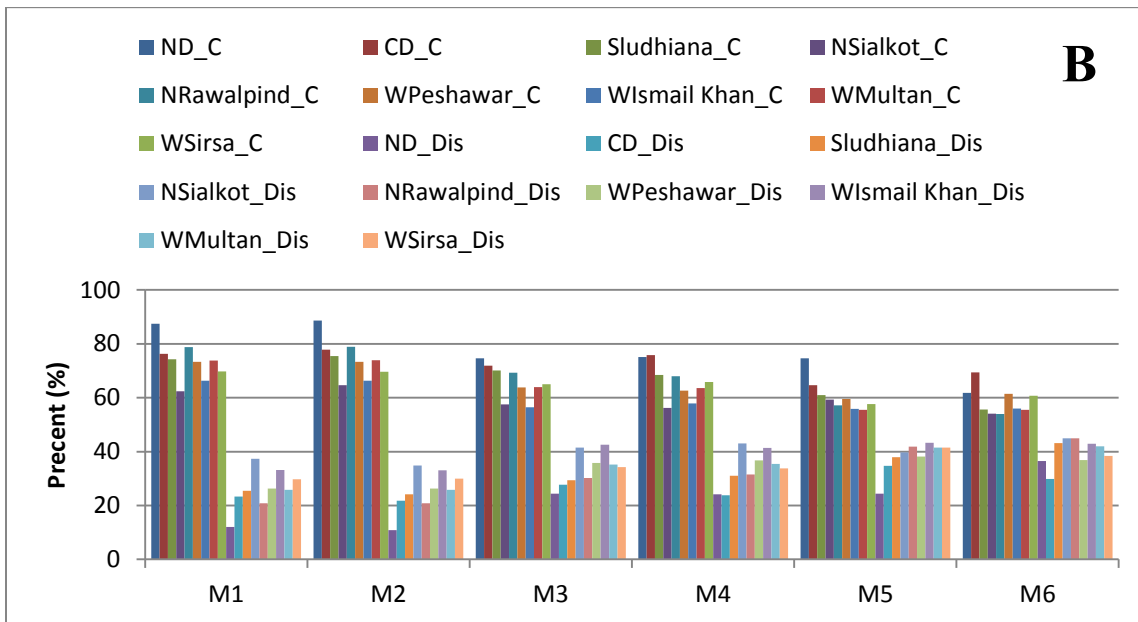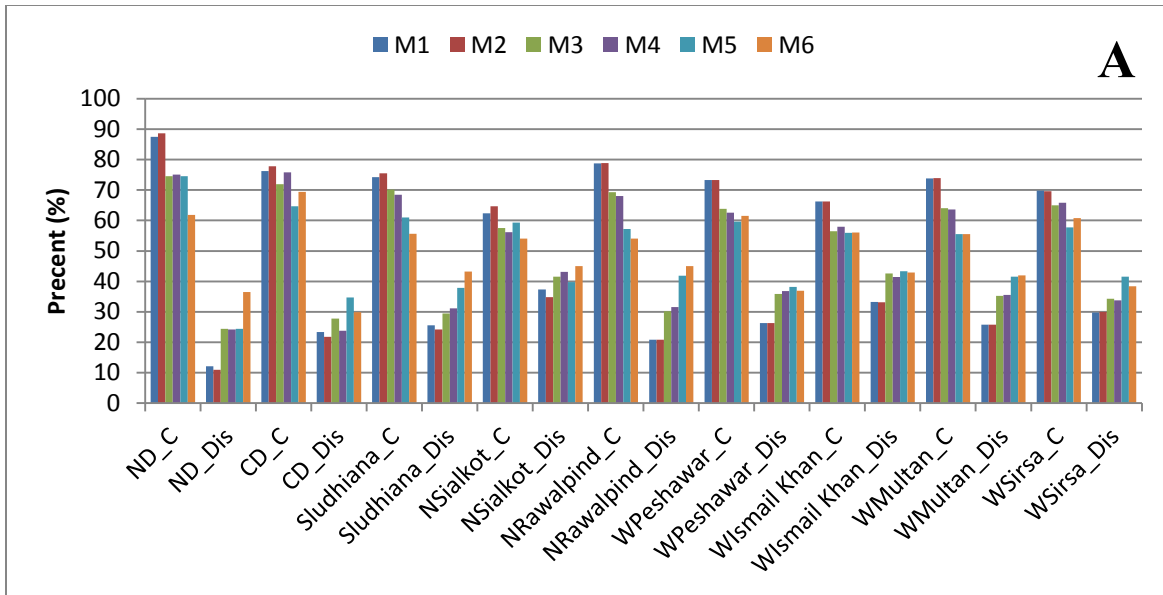
Figure 3.14 The concordant and dis-concordant percentage across the nine stations for six different models. Figure A is arranged by station while figure B is arranged by model.

The stations are; ND: New Delhi; CD: Central Delhi; SLudhiana: Submontane Ludhiana; NSialkot: North Sialkot; NRawalpind: North Rawalpind; WPeshawar: West Peshawar; WIsmail Khan: West Ismail Khan; WMultan: West Multan; WSirsa: West Sirsa. C: Concordant percentage; dis: Dis-Concordant percentage. The models are; M1: T(t); M2: T(t), R(t); M3: T(t-1), R(t); M4: T(t-1), R(t-1); M5: T(t-2), R(t); M6: T(t-2), R(t-1).

Table 3.4 Summary of the best models obtained for nine stations in India and Pakistan (1875-1900)

Model Parameters

	Intercept	LST	Rainfall	Intercept	LST (t-1)	Rainfall
S1: Central Delhi: M2				**Central Delhi: M3**		
Coefficient	-9.25	0.0986	0.0734	-6.9	0.0696	0.975
P-value	0.000	0.000	0.043	0.000	0.000	0.006
SE	1.52	0.180	0.0361	1.29	0.0155	0.036
S2: Central Lahore: M2				**Central Lahore : M3**		
Coefficient	-6.071	0.0692	0.0687	-4.001	0.04225	0.1047
P-value	0.000	0.000	0.126	0.000	0.000	0.018
SE	0.874	0.011	0.0457	0.76	0.00967	0.0454
S3: Submontane Ludhiana: M2				**Submontane Ludhiana: M3**		
Coefficient	-7.13	0.0667	0.0801	-6.35	0.0565	0.984
P-value	0.000	0.000	0.018	0.000	0.000	0.004
SE	1.33	0.0162	0.0337	1.26	0.0154	0.0339
S4: North Sialkot: M2				**North Sialkot: M3**		
Coefficient	-4.95	0.0392	0.35	-3.00	0.0137	0.0531
P-value	0.007	0.006	0.395	0.178	0.319	0.196
SE	1.21	0.0151	0.0402	1.07	0.0139	0.0397
S5: North Rawalpindi: M2				**North Rawalpindi: M3**		
Coefficient	-7.94	0.0911	0.0056	-4.577	0.0451	0.0581
P-value	0.000	0.000	0.894	0.000	0.000	0.157
SE	1.18	0.0153	0.0425	0.869	0.0115	0.0409
S6: West Peshawar: M2				**West Peshawar: M3**		
Coefficient	-6.28	0.0632	0.0568	-3.616	0.0297	0.118
P-value	0.000	0.000	0.547	0.012	0.005	0.254
SE	1.08	0.0132	0.0995	0.852	0.011	0.114
S7: West Ismail Khan: M2				**West Ismail Khan: M3**		
Coefficient	-5.68	0.0451	0.039	-3.27	0.0148	0.074
P-value	0.006	0.002	0.786	0.418	0.274	0.613
SE	1.32	0.0159	0.141	1.08	0.0138	0.141
S8: West Multan:				**West Multan: M3**		

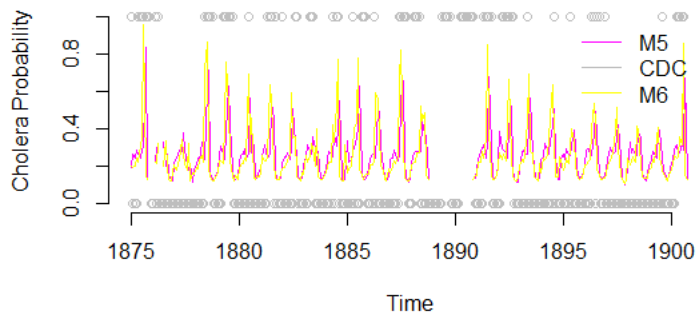| M2						
Coefficient	-9.2	0.0811	0.074	-5.03	0.0317	0.14
P-value	0.000	0.000	0.531	0.043	0.049	0.242
SE	2.04	0.0232	0.114	1.4	0.0168	0.112
S9: West Sirsa: M2				**West Sirsa: M3**		
Coefficient	-6.4	0.0615	0.0127	-5.02	0.0442	0.0254
P-value	0.000	0.000	0.875	0.003	0.001	0.749
SE	1.3	0.0155	0.0815	1.18	0.0142	0.0786

*LST: Land Surface Temperature

Figure 3.15 Epidemiological data of cholera seasonal peaks showing the temporal dynamics of observed cholera cases and fitted models for M1 through M6 shown in grey, red, green dark blue, light blue, pink and yellow respectively in Central Delhi, India
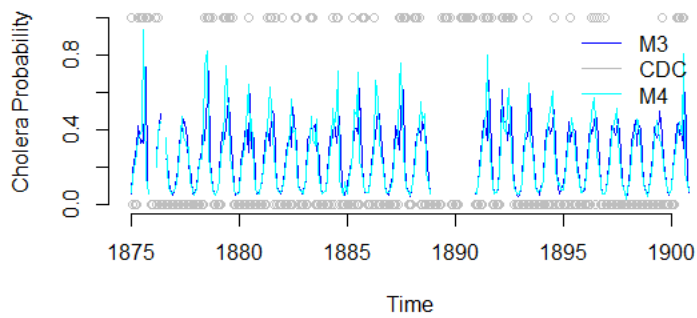
Figure 3.16 Epidemiological data of cholera seasonal peaks showing the temporal dynamics of observed cholera cases and fitted models for M1 through M6 shown in grey, red, green dark blue, light blue, pink and yellow respectively in Central Lahore.
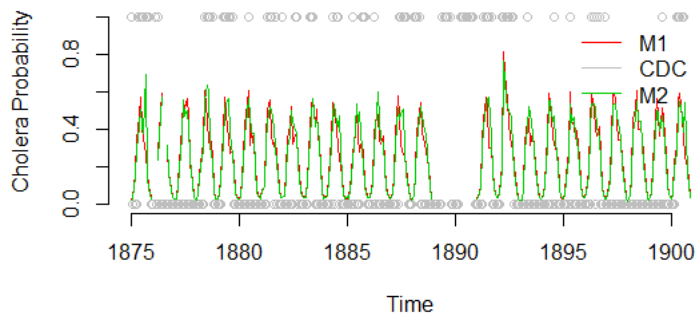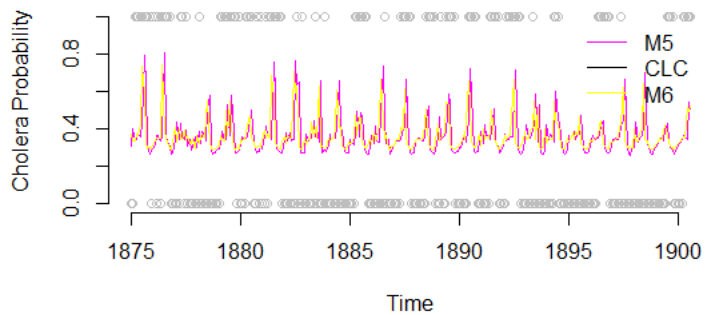
Figure 3.17 Epidemiological data of cholera seasonal peaks showing the temporal dynamics of observed cholera cases and fitted models for M1 through M6 shown in grey, red, green dark blue, light blue, pink and yellow respectively in Submontane Ludhiana.

Figure 3.18 Epidemiological data of cholera seasonal peaks showing the temporal dynamics of observed cholera cases and fitted models for M1 through M6 shown in grey, red, green dark blue, light blue, pink and yellow respectively in North Sialkot.
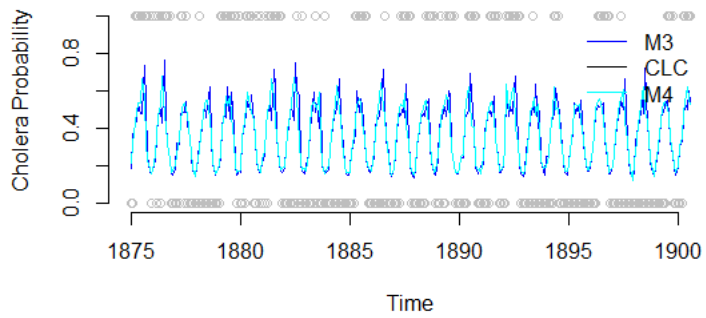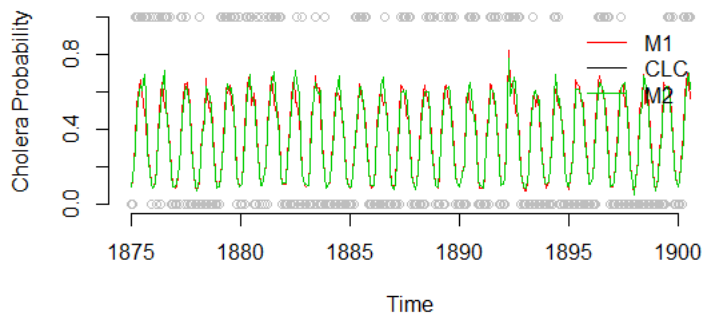
Figure 3.19 Epidemiological data of cholera seasonal peaks showing the temporal dynamics of observed cholera cases and fitted models for M1 through M6 shown in grey, red, green dark blue, light blue, pink and yellow respectively in North Rawalpindi.
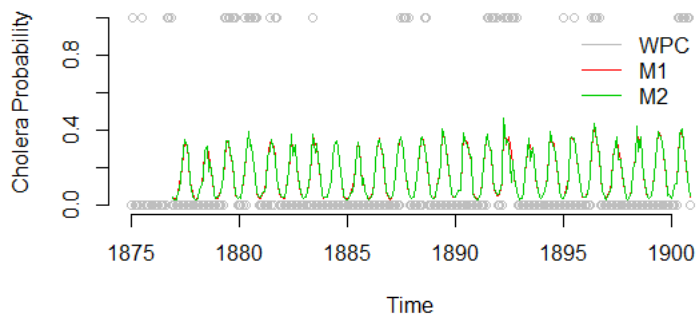
Figure 3.20 Epidemiological data of cholera seasonal peaks showing the temporal dynamics of observed cholera cases and fitted models for M1 through M6 shown in grey, red, green dark blue, light blue, pink and yellow respectively in West Peshwar.
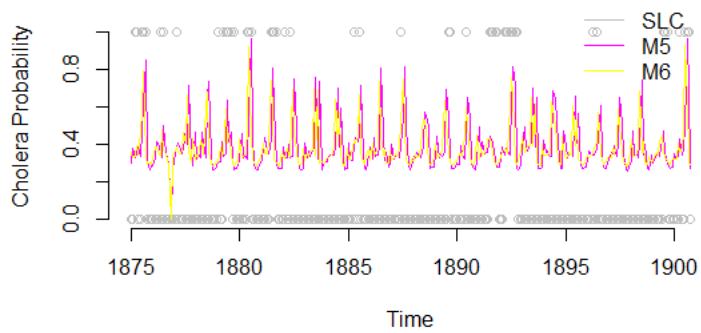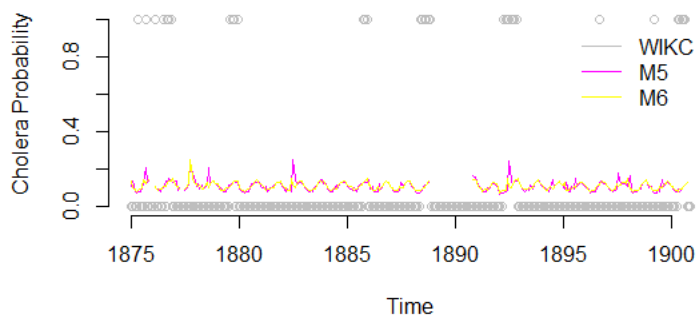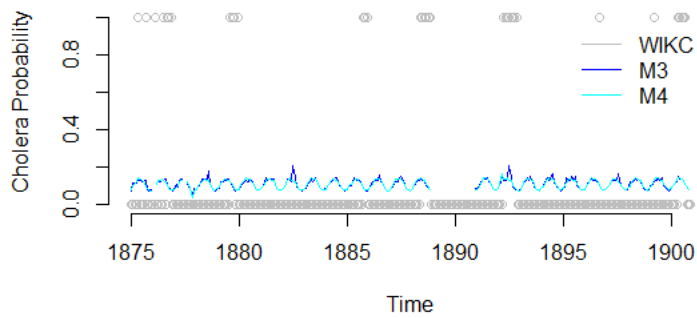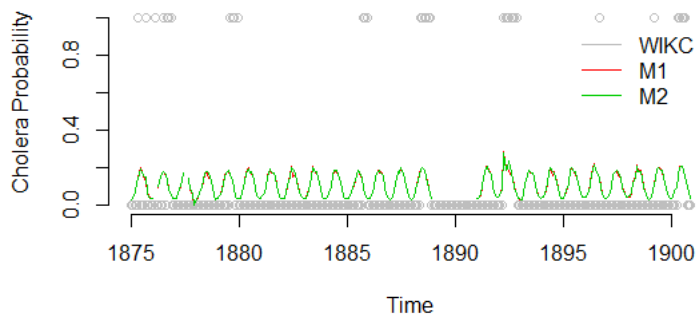
Figure 3.21 Epidemiological data of cholera seasonal peaks showing the temporal dynamics of observed cholera cases and fitted models for M1 through M6 shown in grey, red, green dark blue, light blue, pink and yellow respectively in West Ismail Khan.
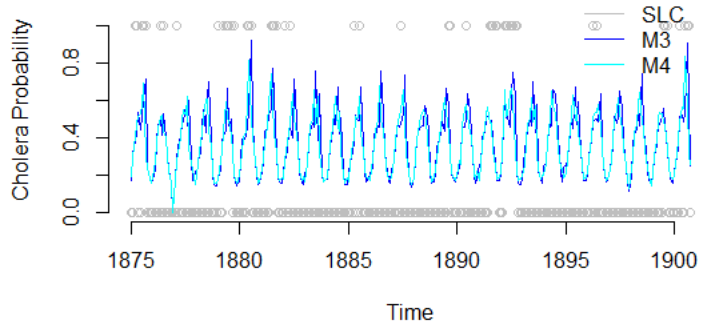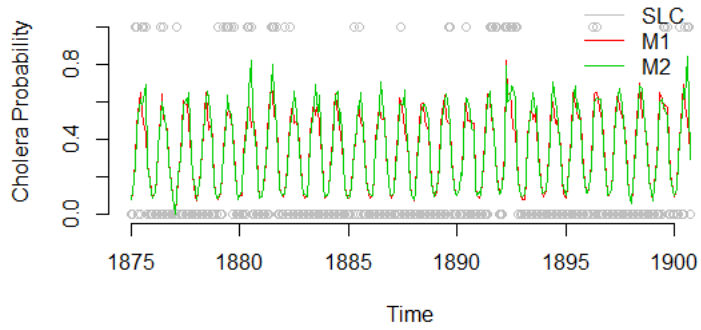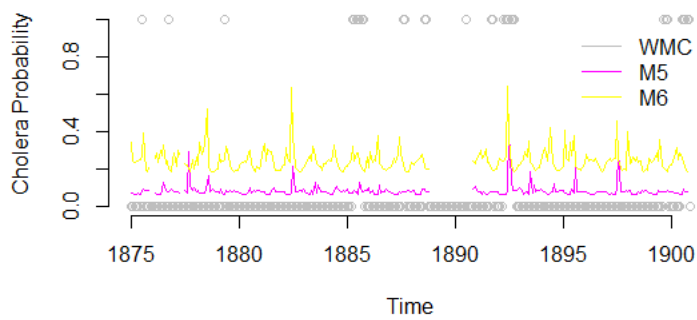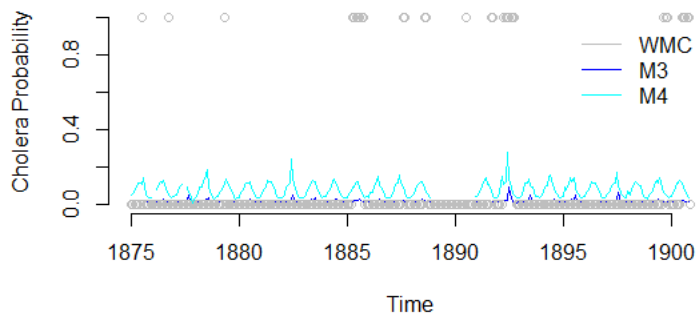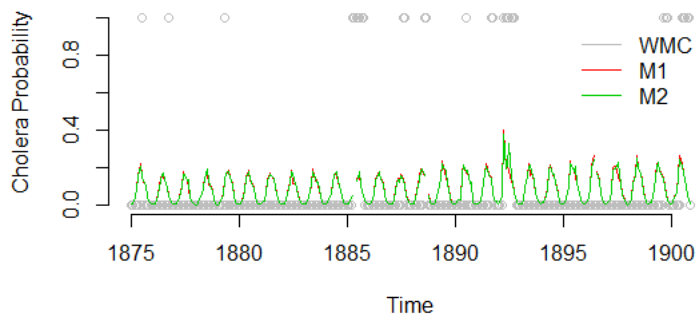
Figure 3.22 Epidemiological data of cholera seasonal peaks showing the temporal dynamics of observed cholera cases and fitted models for M1 through M6 shown in grey, red, green dark blue, light blue, pink and yellow respectively in West Multan.
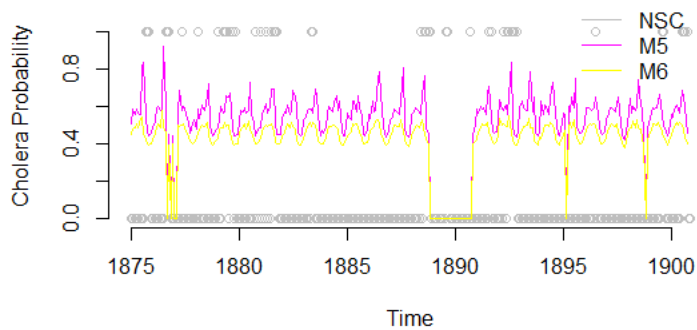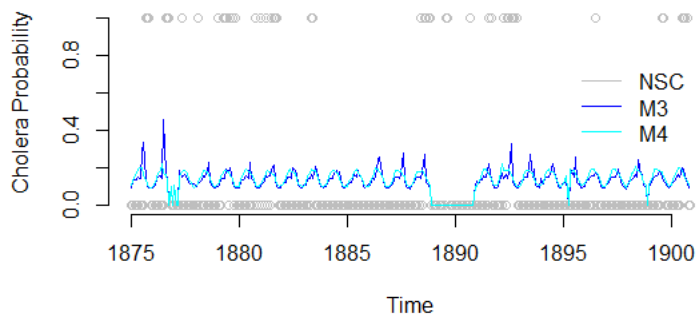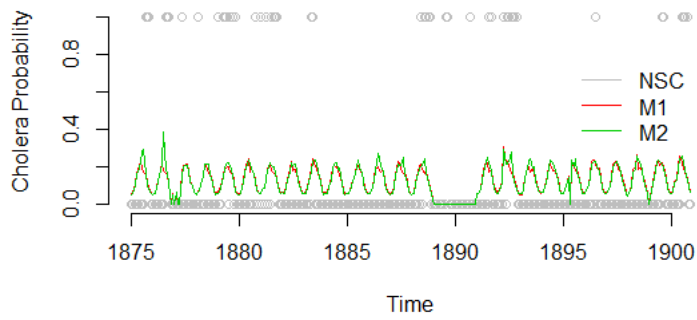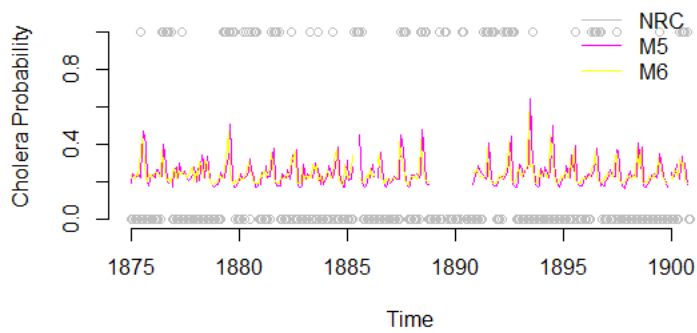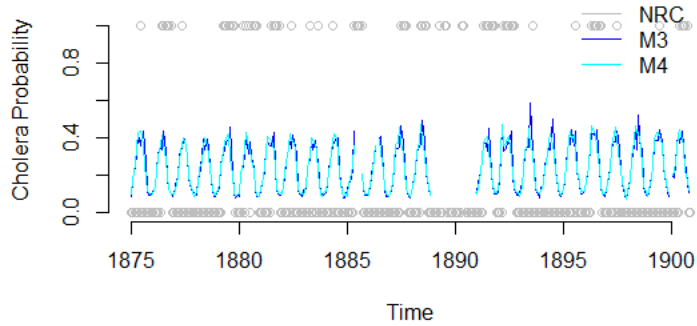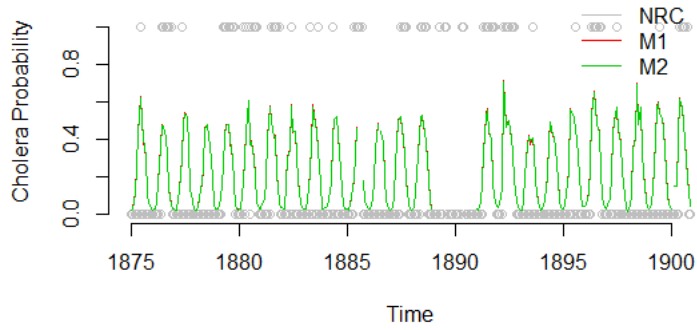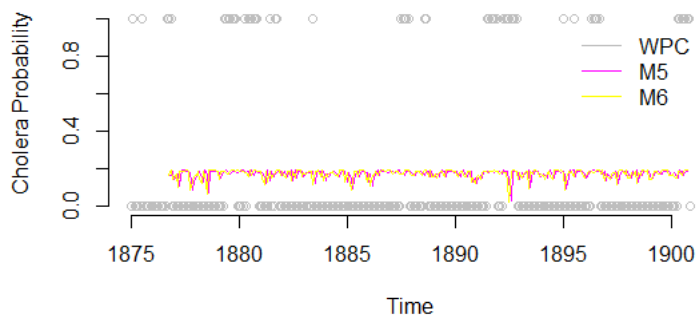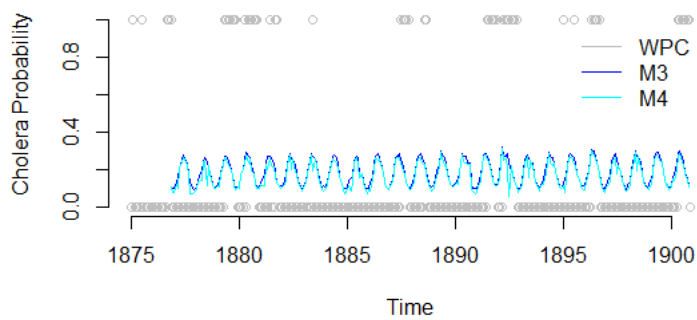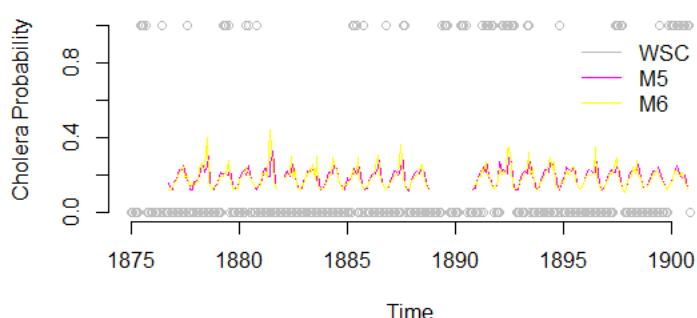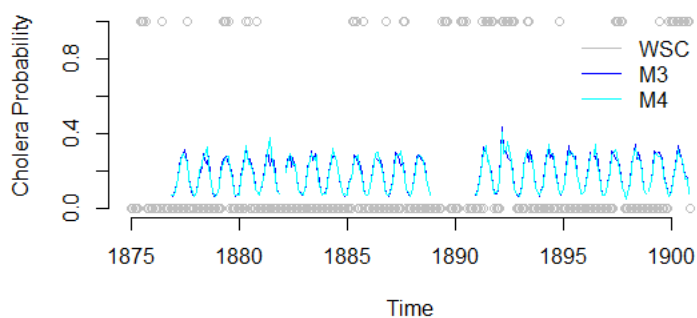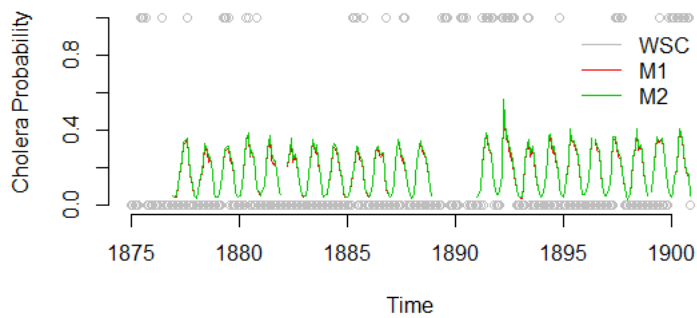
Figure 3.23 Epidemiological data of cholera seasonal peaks showing the temporal dynamics of observed cholera cases and fitted models for M1 through M6 shown in grey, red, green dark blue, light blue, pink and yellow respectively in West Sirsa.

3.9 Cholera Transmission/Spread modeling

3.9.1 SIR model background

Modeling of infectious diseases has two important roles; *prediction and understanding* which will help us to improve disease control and finally to eradicate the infection from the population, if possible. Prediction and understanding are related to model properties including accuracy and transparency. The common use of the model is the prediction which requires the model to be as accurate as possible. In order to build a good predictive model, we need to include all the known complexities and population level heterogeneities. Accurate predictive models can have an additional use as a statistical tool. The failure to accurately predict epidemic behavior in certain regions may be a warning that the underlying parameters in the model are different from the observable data. Models also can be used to understand how an infectious disease spreads in the real world and how various complexities affect the dynamics. As a consequence, it will provide epidemiologists with an ideal world to examine individual factors in isolation and decide which factors are important and which can be neglected. Finally, the understanding gained from modeling can help us to develop more deep accurate predictive models and gather more relevant epidemiological data.

"All models are wrong but some are useful" George E. P. Box (1919 – 2013). Models have their limitations. Because of infectious disease transmission dynamics, it is impossible to build a fully accurate model. There will always be some element of the host behavior or disease that is unknown. The best that we can do with modeling is to provide

confidence intervals on the epidemic behavior and determine the risk of infection for hosts.

The SIR model, developed by Kermack and McKendrick in 1927, is a simple model meant to quantitatively explain the dynamics of an epidemic. The name is an acronym, with S standing for 'susceptible individuals,' I standing for 'infected individuals,' and R standing for 'removed/recovered individuals,' or individuals who are no longer at risk for infection. A fixed population does not allow for removal or death, so at any time a member of a fixed total population falls into one of three categories: at risk for infection, infected, or recovered and now immune.

The Variables

The first step in finding the model of how the populations in these categories interact is defining our variables.

   I.   *Independent variable*: t (time) is measured in days.

  II.   *Dependent variables* that have two related sets, which, while they give the same information about the epidemic (differentiated by a factor of 1/N), sometimes are convenient in different ways for different equations.

       a. First, the number of people in each category are given as follows:

           i. *S=S(t)*, the number of susceptible individuals. Susceptible Individuals can be defined as those ones who never been infected and they are able to catch the disease. Once they have it, they move into the infected compartment.

          ii. *I=I(t)*, the number of infected individuals. Disease infection can spread from infected individuals to susceptible individuals. The

time they spend in the infected status is the infectious period, after which they enter the recovered compartment.

  iii. $R=R(t)$, the number of recovered individuals. Recovered Individuals are assumed to be immune for life.

 b. We also consider the fraction of the total population, N, that each category occupies:

  i. $s(t)=S(t)/N$, the susceptible fraction of the total population

  ii. $i(t)=I(t)/N$, the infected fraction of the total population

  iii. $r(t)=R(t)/N$, the recovered fraction of the total population

Without further information, it can be predicted that $s(t)$ will have the highest initial value out of the three variables and will decrease with time as susceptible individuals are infected, while $r(t)$ will have the lowest initial value and will increase with time as infected individuals recover.

Assumptions

- There is no positive growth of S (closed population); in other words, there is no birth or immigration into the population. The model also assumes that the only way an individual can leave S is by coming into contact with a member of I and becoming a member of I itself.

- Homogenous mixing of the population, where intricacies affecting the pattern of contacts are discarded, yielding $\beta\,S\,I$ are the transmission term.

- $S + I + R = 1$

- An individual in I comes into contact with an average number, b, of people per day. However, not all of those people are susceptible, only a fraction of the total

population given by *s(t)*. Therefore, the total number of newly infected people each day is *β? s(t)*.

- In terms of recovery, it is assumed that a fixed fraction, *γ*, of the infected population recovers every day, meaning that the total number of newly recovered people each day is *γ? i(t)*.

The rate of new infections can thus be defined as βSI, where β is a parameter for infectivity. Infected individuals are assumed to recover with a constant probability at any time, which translates into a constant per capita recovery rate that we denote with γ, and thus an overall rate of recovery γI. Based on these assumptions we can draw the scheme of the model:

$$\beta\,S\,I \qquad \gamma\,I$$

| S | → | I | → | R |

Equations:

$$\frac{dS}{dt} = -\beta\,S\,I$$

$$\frac{dI}{dt} = -\beta\,S\,I - \gamma\,I$$

$$\frac{dR}{dt} = \gamma\,I$$

Based on the first of the given assumptions, the only change that *S* can experience is negative, as susceptible individuals become infected individuals based on the fraction of the population that is infected and coming into contact with the susceptible fraction of the population and join population *I*. Similarly, the only way that *R* can change is to grow

as individuals recover and leave population *I* at a rate of γ. *I* reflects the changes in the other two populations, with the signs reversed.

3.9.2 SIR model and cholera transmission

Miller *et al.,* (1982) identified two routes of transmission for cholera disease: ***Primary transmission*** spread to the susceptible individual through some form of contact with water, local *V. cholerae* habitat, or consumption of shellfish or aquatic plants contaminated with *V. Cholerae*; and ***Secondary transmission***, diffusion of cholera, spread to susceptible individuals in a population through the infected individual.
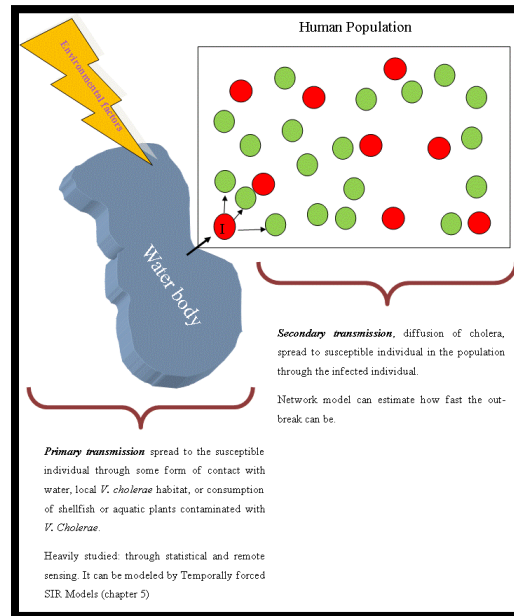
Figure 3.24 Two transmission routes for cholera

However, cholera is as a water-borne indirectly transmitted infectious disease is poorly studied. B*y using the SIR model, we aim to answer these questions: How many index cases should there be to cause an outbreak? What is the transmission rate of cholera disease in endemic and epidemic regions?* Based on the above information; there are two types of infected individuals that can be determined: 1) the primary infected individuals can be computed from a statistical model, and 2) the secondary infected individuals can be computed through the above differential equations.

Strategies to cholera SIR model:

We aim to combine both primary and secondary transmission routes in one model to investigate the spread of the disease.

I. **Model backbone:**

1. Calculating the primary infection through a statistical model from the New Delhi, India data.

    a. At each month of the year, there will be IP (primary infected individuals)

    b. Divide those numbers across the month days to estimate IP at each day to avoid the exponential increase in the infection.

2. Use real data for the model instead of theoretical numbers. So we used $N=10^6$.

3. Use small transmission rate (β) since the disease is not that contagious

4. Use I0 = 0

5. Set the max time as the length of IP

6. Feed IP with IS (Secondary infection) in order to calculate the total infection

    a. So, another term "IP(t)" was added to the actual model. Then, the same term is subtracted from the susceptible to preserve the consistency of the model.

    b. **The new Equations will be:**

$$\frac{dS}{dt} = -\beta\, S\, I - IP(t)$$
$$\frac{dI}{dt} = -\beta\, S\, I - \gamma\, I + IP(t)$$
$$\frac{dR}{dt} = \gamma\, I$$

7. Plotting the data
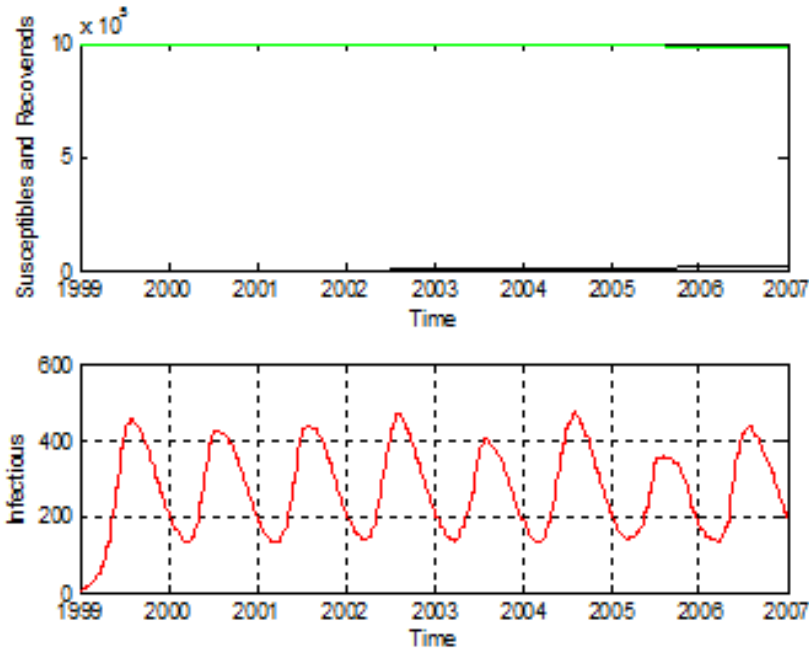
Figure 3.25 The total cholera infected individuals shown in red from a model that combines primary transmission route (a statistical model) and secondary transmission route (SIR model). Susceptible and recovered individuals are shown in green and black respectively.

II. **Optimization of the model:** By using different population number, different

beta (β) and gamma (γ).

1. Future work: use different β during the year.

III. **Model Validation strategy:**

1. Taking the model output that has the total infective individuals (IS +

IP)

2. Conduct statistical analysis to build the best fit model with the original

environmental data.

3. Compare the model quality with the original model.

3.10 Discussion:

Diarrhea is considered to be the second most leading cause of death in children below five years of age (Wardlaw *et al*., 2010). About 9 million under 5 years old children died in 2008 and 40 percent of death cases were due to pneumonia and diarrhea diseases (You *et al*., 2010). The estimated diarrheal disease burden from water, sanitation, and hygiene at the global level has revealed 4.0 percent of all deaths and 5.7 percent of the total diseases burden (Pruss *et al*., 2002). In Bengal, climate is a critical factor for *V. cholerae* survival, persistence, and spread. Russel studies in 1928 explained the decline in death rate by the decrease in the relative humidity in some districts of Burdwan and Malda. Hot and moist climate conditions in the Bengal basin have adverse effects on the general health of the people (Tromp, 1963). In the current study, the quantitative relationship between land surface temperatures and cholera incidence was established with statistical modelling approaches. In addition, results show that temperature with a one month lag period contributes to the emergence of a cholera outbreak. However, our main hypothesis is that air temperature is the main driver for the endemic cholera sporadic outbreaks in New Delhi, India. This hypothesis is further confirmed by: 1) Historical air temperature data for New Delhi, India from 1955 to the present. We can see that there is a gradual increase in air temperature to maximum in recent years, from 2000 to the present; and 2) Air temperature shows 87.5% concordance with the observed data.

During early 1920s in India, No correlation was observed between rainfall and cholera incidence because rainfall alone was not a critical factor for cholera incidence (Rogers, 1926). Our current study proved a significant correlation between rainfall and

cholera incidence as well as quantitatively estimated this relationship. However, in 2011 Goel and Jiang described how heavy rainfall can cause a rapid shift of *V. cholerae* genotype from one strain to another strain in one epidemic region in India. Other factors such as contamination of drinking water, defective drainage and surface outwash are important for cholera infection spread. Later in the 1920s, Russel and Sundararajan established that cholera incidence is greatly associated with high temperature and intermittent rainfall which act as ideal climatic conditions for *V. cholerae* multiplication. Our current scenario of cholera incidence in New Delhi, India revealed that ideal climatic condition for a cholera outbreak depends upon the connectivity between rainfall and temperature. The present study shows that both land surface temperature (LST) and rainfall had a greater concordance percentage (88.7%) than the model that has only LST. So both LST and rainfall are important factors for a cholera outbreak. The observed temperature and rainfall changes over a thirteen year period were not constant owing to either El Niño or La Niño. Interestingly, we established a relationship between LST and rainfall since a good quality predictive model was constructed with LST (one month lag period) and rainfall that has a 74.6% concordance. Our findings demonstrated that the climate and health relation in New Delhi and the relative land surface temperature plays a big role in increased *Vibrio*-mediated infections. By using this good predictive model, a model that combines both a statistical model as well as a SIR model was constructed and it has improved prediction ability for cholera outbreaks compared to a basic SIR model alone.

3.11 The expected pitfall and caveats?

We were not able to collect satellite-derived chlorophyll data to investigate any combined effect with LST (Land surface temperature) for cholera incidence and endemism in New Delhi, India.

Chapter 4

Summary, Research Contributions, and Future Research

4.1 Summary of the Research

The overall goal of the proposed research was to develop a quantitative framework for cholera outbreak high quality models either cholera simulated models (CSM) or Cholera prediction models (CPM) with two to three months lead time, using both *in-situ* and remote-sensing data. To achieve this goal, four closely-related research objectives were followed: (i) Determination of the role of environmental factors associated with seasonality and modulating the dynamics of a cholera outbreak, (ii) Identification and development of a physically plausible hypothesis of how local environmental factors modulate cholera outbreak dynamics, (iii) Identification of the major environmental controls triggers sporadic cholera outbreaks in epidemic regions, and (v) Constructing a new model that accounts for both primary and secondary transmission routes

Seasonality of cholera outbreaks differs from region to region in endemic areas such as Mathbaria and Bakergonj in Bangladesh. Such temporal and spatial variation between cholera outbreaks in those two spatially different regions implies that there are different driver mechanisms which affect cholera outbreaks. Ponds in both regions (Mathbaria and Bakergonj) are connected to large tributaries of the Ganges River and are prone to tidal intrusion of coastal seawater during the spring season since river discharge during the spring season is extremely low. Using *in-situ* data collected from both regions, it was established that water depth, explaining 59% of the model variance, is the main

environmental factor that drives cholera outbreaks in both regions but by a different mechanism. The negative association observed in this study between water depth and spring cholera in both Mathbaria and Bakergonj is interpreted as an effect of coastal water intrusion. Cholera transmission is increased by inland water body contamination that will increase salinity and nutrient concentration (Rao, 1973; Valsaraj, Rao, 1994). In the inland regions of the Bengal delta, intrusion of coastal water role during low river height on cholera during the spring season was proposed by Akanda *et al.,* (2009), using data for large scale hydroclimatic processes. Coastal water intrusion followed by flooding during high river height results in a fall peak of cholera cases. Our analysis for both Mathbaria and Bakergonj supports and strengthens this hypothesis. Modeling results for both Mathbaria and Bakergonj are consistent with those of Akanda *et al.,* (2009), namely that water depth has a negative coefficient in spring models for both regions, which implies that the spring peak is related to coastal intrusion. While it has a positive coefficient in the fall peak models, which implies widespread flooding and cross contamination arising from inadequate sanitation. Coastal seawater intrusion and flooding theory which initiates cholera during the spring and fall seasons respectively is proposed by both Akanda *et al*., (2009) and Jutla *et al*., (2013) using large scale environmental factors. Higher values of chlorophyll-c with seasonal increase in number of cholera cases most likely reflects the role of plankton blooms and subsequent growth of zooplankton populations and related increasing *V. cholerae* population numbers (Colwell, 1996, Nalin *et al*., 1979, Rawlings *et al*., 2007). Stream flow from the major rivers of the Bengal Delta and remotely sensed plankton abundance in the coastal Bay of Bengal together explain over 75% of the variability of cholera in the Bengal Delta (Jutla *et al*., 2013). An

important aspect of this study is the potential for a cholera prediction model using information from *in-situ* hydroclimatic processes readily obtained the ponds and the link that exists between large scale and regional scale environmental factors. The importance of the local environment relative to large scale hydrologic factors is evident when we study seasonal cholera dynamics in two locations with regional differences.

This is one of the first few studies that quantitatively link available *in-situ* data with large scale hydroclimatic factors to build up models for cholera outbreaks. Key findings from the phenomenological modeling framework are: (1) Cholera outbreaks can be predicted using two-seasonal modeling strategies depending on the choice of *in-situ* regional environmental variables, (2) Implementation of preventive measures to contain spread of outbreaks with minimal installation and operating cost, yet allowing timely, (3) Establishment of the link between macro-environmental variables and local environmental variables.

To achieve the third goal, cholera was evaluated in one of the epidemic regions (India, New Delhi). The aim was to determine the main triggers of cholera outbreaks and validate a transmission mechanism of cholera in epidemic regions. New Delhi is connected to large tributaries of the Indus and is recognized to have sporadic cholera outbreaks that can change from year to year. Cholera outbreaks in New Delhi have sporadic peaks with changing seasonality throughout the years. A good quantitatively predictive model with relatively high Pseudo R-Square (64.64%) is constructed using land surface temperature, rainfall, and rainfall with three months lag period. Jutla *et al.,* (2013), using data for large-scale hydroclimatic processes, proposed that cholera outbreaks are more likely to spread when rainfall is followed by high temperature. Our

primary and quantitative analysis for New Delhi, India strengthens this hypothesis. Moreover, this hypothesis is validated by collecting historical data back to 1875. We can see rising temperature nowadays as compared to 1950. On the other hand, rainfall and rising temperature hypothesis is confirmed by the quantitative estimation of cholera outbreaks at nine stations in India and Pakistan. We observed that both rainfall and temperature (one month lag) are significant factors in all the conducted models. A study carried out in Azerbaijan, covering a period of 28 years, also showed a strong correlation between air surface temperature and water temperature with the occurrence of *V. cholerae* O1 followed by cases of cholera (Gurbanov *et al.*, 2012).

4.2 Research Contributions

The major contribution of the research is the quantitative evaluation of the both *in-situ* and satellite-based prediction modeling architecture for cholera outbreaks with a lead time of two to three months. Although statistical in nature, the prediction models have shown that satellites have tremendous potential to predict cholera outbreaks (~60% prediction in Bengal) in the ground-based data scarce regions. The research also filled the gab that exists between large-scale environmental factors and *in-situ* environmental factors.

My research attempts to integrate findings from three related disciplines (hydrology, epidemiology, and microbiology), using the latest data *in-situ* local environmental conditions, as well as satellite remote-sensing measured variables. The importance and uniqueness of my research is in development of a comprehensive understanding of cholera dynamics using combined information from these three disciplines, from which development of an actionable prediction modeling framework for possible intervention

strategies is accomplished. The major contribution from my research will be development and validation of a prediction model for cholera incidence using both local and large-scale geophysical variables. I expect the following outcomes from this research: (i) Provision of an estimate of local environmental conditions useful for a cholera model relative to the Bengal Delta region; (ii) Development of a physically consistent hypothesis to validate the observed relationship between spring and fall cholera peaks in two geographically distinct regions, based on local environmental conditions; (iii) Explanation of the role of Land Surface Temperature and rainfall in altering cholera incidence dynamics in South Asia; (iv) Development and validation of a cholera prediction model with the capability to forewarn seasonal outbreaks of cholera; and (v) Construction of a good quality model that accounts for both primary and secondary transmission routes.

The research resulted in two publications, one research grant from National Institutes of Health, and 3 conference proceedings reports. To disseminate our understanding of the water-related diseases, we also organized a poster session at the American Society of Microbiology meeting (May 17-20, 2014) titled: A Mechanistic Model to Understand Trigger and Transmission of Cholera in Epidemic Regions

4.3 Future Prospects:

The model prediction and simulation results from the current research (more than 65% accuracy) are promising; yet, we are cognizant of possible caveats and limitations that warrant further investigation. Our results are based on nine years of both *in-situ* and remote-sensing data; as more data become available, the proposed approach needs to be further validated and refined. Colwell *et al*., 1977 proposed the idea of environmental

cholera transmission and subsequently reported on the topic in greater detail in a series of subsequent publications (e.g., Tamplin *et al*., 1990; Colwell and Huq, 1994; Akanda *et al*., 2009; Jutla *et al*., 2013). The role of environment in cholera transmission continues to be refined as more observational data become available and as the roles of different abiotic, biotic, and hydroclimatological factors affecting cholera transmission are clarified. *In-situ* environmental factors such as water depth have a direct connection to *V. cholerae* outbreaks in endemic regions. Although we have a good quality model, it seems that the sampling/analysis process at the local level is a critical factor in order to build a better predictive model. On the other hand, both rainfall and land surface temperature estimates from satellite remote sensing provide a quantitative measure of space-time distribution of cholera outbreak. However, rainfall and LST alone cannot provide the whole explanation of why cholera outbreak happens. Our future research should address if there are other factors that are important and increase the outbreak. As well as it should have consistent sampling process to have better prediction models. Our results demonstrate that remot sensing data measured by satellite over a range of space and time scales can be very effective in developing a cholera prediction model with several months' lead time. Such prediction lead time will have tangible 196 impacts to design and implement effective cholera intervention and mitigation strategies for various resource constrained and cholera affected regions of the world.

Appendix 1

Parameter	Formula and meaning
Test of goodness of fit	It establishes whether or not an observed frequency distribution differs from a theoretical distribution. It includes deviance, Pearson's chi-squared test ($\chi 2$), and Hosmer–Lemeshow test.
1. Deviance	It is used for statistical hypothesis testing to measure the quality of fit statistic for a model. By using the sum of squares of residuals in ordinary least squares to cases the model-fitting is achieved by maximum likelihood. $$D(y) = -2\Big(\log\left(p(y \mid \hat{\theta}_0)\right) - \log\left(p(y \mid \hat{\theta}_s)\right)\Big).$$ $\hat{\theta}_0$: is the fitted values of the parameters in the model $\hat{\theta}_s$: is the fitted parameters for the "full model" Both of them are functions of the observations y.
2. Pearson's chi-squared test ($\chi 2$)	It is a statistical test that can be applied to categorical data to measure if there is any observed difference between those sets arose by chance. It is first investigated by Karl Pearson in 1900. $$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$ χ^2 = Pearson's cumulative test statistic, that has χ^2 distribution. O_i= an observed frequency; E_i= an expected (theoretical) frequency, n= the number of cells in the table.
3. Hosmer-Lemeshow	This statistical test is especially used to assess the goodness of fit for logistic regression mode. It is used to measure how the observed event rates match with the expected event rates in subgroups of the model population. The Hosmer–Lemeshow test statistic is given by: $$H = \sum_{g=1}^{G} \frac{(O_g - E_g)^2}{N_g \pi_g (1 - \pi_g)}.$$ Og: is the observed event for the gth risk decile group. Eg: is the expected event for the gth risk decile group. Ng: is the number of observations for the gth risk decile group πg: is the predicted risk for the gth risk decile group

	G is the number of groups.
	The test statistic asymptotically follows a χ^2 distribution with $G - 2$ degrees of freedom. The number of risk groups may be adjusted depending on how many fitted risks are determined by the model. This helps to avoid singular decile groups.
Measures of Association	Two or more events are said to be associated with each other if the probability of occurrence of one depends on the occurrence of the others. For logistic regression, we will use: Concordant (%), Discordant (%), Somers' D, Goodman and Kruskal's gamma, and Kendall's Tau-a
Concordant (%) and Discordant (%)	Concordant (%): is the percentage of the cases that are ranked in the same order on both variables. Discordant (%): is the percentage of cases ranked differently on the variables. For example if (x_1, y_1), (x_2, y_2), …, (x_n, y_n) for the random variables X and Y respectively. • Any pair of observations are *concordant*, if (x_i, y_i) and (x_j, y_j) agree in the ranking. • Any pair of observations are *disconcordant,* if both $x_i > x_j$ and $y_i > y_j$ or if both $x_i < x_j$ and $y_i < y_j$. • Any pair of observations is neither concordant nor discordant. if $x_i > x_j$ and $y_i < y_j$ or if $x_i < x_j$ and $y_i > y_j$. If $x_i = x_j$ or $y_i = y_j$,
Somers' D	It measures ordinal association or the relationship between two ordinal variables (for example, low, medium and high exposure levels of antibiotic) Kruskal and William (December 1958).
Goodman and Kruskal's gamma	It measures the rank correlation or how much orderings of the data is similar when it is ranked by each of the quantities. Values range from −1 (100% negative association, or perfect inversion) to +1 (100% positive association, or perfect agreement). A value of zero indicates the absence of association. $$G = \frac{N_s - N_d}{N_s + N_d}.$$ N_s: is the number of concordant pairs that are ranked in the same order on both variables. N_d: is the number of discordant pairs that are ranked differently on the variables.
Kendall's Tau-a Or Kendall's tau (τ) coefficient	It is a non-parametric hypothesis test to measure rank correlation between two quantities (Maurice Kendall, 1938) $$\tau = \frac{(\text{number of concordant pairs}) - (\text{number of discordant pairs})}{\frac{1}{2}n(n-1)}.$$

$$R^2 = 1 - \frac{\ln \hat{L}(M_{Full})}{\ln \hat{L}(M_{Intercept})}$$

McFadden's Pseudo R-Square	M_{full} = Model with predictors $M_{intercept}$ = Model without predictors The log likelihood of the intercept model will represent the total sum of squares while the log likelihood of the full model will represent the sum of squared errors. The ratio of both likelihoods will indicate the level of improvement over the intercept model offered by the full model. When we compare two models on the same data, McFadden's Pseudo R-Square would be higher for the model with the greater likelihood. $\hat{L}$ = Estimated likelihood

References:

Acosta C.J., C.M. Galindo, J. Kimario, K. Senkoro, H. Urassa, C. Casals, M. Corachan, N. Eseko, M. Tanner, and H. Mshinda (2001): Cholera outbreak in southern Tanzania: risk factors and patterns of transmission. *Emerging Infectious Diseases*; 7(3):583-587.

Ahmad Q.K., A.K. Biswas, R. Rangachari, and M.M. Sainju (2001): Ganges-Brahmaputra-Meghna Region: A Framework for Sustainable Development. The University Press, Dhaka

Akanda A. S., A.S. Jutla, M. Alam, G.C. de Magny, A.K. Siddique, R.B. Sack, A. Huq, R. R. Colwell, and S. Islam (2011): Hydroclimatic influences on seasonal and spatial cholera transmission cycles: Implications for public health intervention in the Bengal Delta. *WATER RESOURCES RESEARCH*, VOL. 47, W00H07, doi:10.1029/2010WR009914, 2011

Akanda, A.S., A.S. Jutla and S. Islam, (2009): Bimodal Cholera Transmission in Bengal Delta: A Hydroclimatological Explanation. *Geophy. Res. Lett*. 36, L19401, doi: 10.1029 / 2009 GL039312.

Alam, M., *et al*., (2006), Seasonal cholera caused by *Vibrio cholerae* serogroups O1 and O139 in the coastal aquatic environment of Bangladesh, *Appl. Environ. Microbiol*., 72, 4096–4104, doi:10.1128/AEM.00066-06.

Ali M., A.L. Lopez, Y.A. You, Y.E. Kim, B. Sah, B. Maskery, and J. Clemens (2012): The global burden of cholera. Bulletin of the World Health Organization 2012; 90:209-218A. doi: 10.2471/BLT.11.093427

Ali M., M., Emch , J.P. Donnay, M. Yunus, R.B. Sack (2002): The spatial epidemiology of cholera in an endemic area of Bangladesh. *Soc Sci Med*; 55(6):1015-24.

Aragón M.B.A., P. Tabbard, J. Chambule, C. Santos, A. Noya (1994): Epidemiologia da cólera em Moçambique no período de 1973–1992. *Revista Saúde Pública* 1994, 28(5):332-336.

Bailey, N. (1957): The mathematical theory of epidemics. Griffin, London.

Banerjee B. and J. Hazra (1974): Geoecology of cholera in West Bengal: A study in medical geography. Calcutta: Jayati Hazra Publishers

Baqui A.H., R.B. Sack, R.E. Black, K. Haider, A. Hossain, A.R. Alim, M. Yunus, H.R. Chowdhury, A.K. Siddique (1992): Enteropathogens associated with acute and persistent diarrhea in Bangladeshi children less than 5 years of age. *J Infect Dis* 1992, 166(4):792-796.

Bateman C (2002): Mozambique cholera will affect region. *S Afr Med J*, 92:104-106.

Bertuzzo, E., S. Azaele, A. Maritan, M. Gatto, I. Rodriguez-Iturbe, and A. Rinaldo (2008): On the space-time evolution of a cholera epidemic, *Water Resour. Res.*, 44, W01424, doi:10.1029/2007WR006211.

Biswas A.K., J.I. Uitto (2001): Sustainable Development of the Ganges-Brahmaputra-Meghna Basin. United Nations University Press, Tokyo

Bolloba´s, B. (1985): Random graphs. London: Academic Press.

Boots, M. and A. Sasaki, A. (1999): Small worlds' and the evolution of virulence: infection occurs locally and at a distance. *Proc. R. Soc.* B 266, 1933–1938. (doi:10.1098/rspb.1999.0869.)

Bouma, M.J. and M. Pascual (2001): Seasonal and Interannual Cycles of Endemic cholera in Bengal 1891–1940 in Relation to Climate and Geography. *Hydrobiologia* 460(1-3): 1573-5117.

Bush K., G. Luber, R. Kotha, R.S. Dhaliwal, V. Kapil, M. Pascual, D.G. Brown, H. Frumkin, R.C. Dhiman, J. Hess, M.L. Wilson, K. Balakrishnan, J. Eisenberg, T. Kaur, R. Rood, S. Batterman, A. Joseph, C.J. Gronlund, A. Agrawal, and H. Hu1 (2011): Impacts of Climate Change on Public Health in India: Future Research Directions Environmental Health Perspectives 119 (6)

Carrel, M., P. Voss, P.K. Streatfield, M. Yunus, and M. Emch (2010): Protection from annual flooding is correlated with increased cholera prevalence in Bangladesh: a zero-inflated regression analysis. *Environ. Health* 9: 13. PMCID: PMC2856547

Chowdhury, M., and M. Ward (2007): Seasonal flooding in Bangladesh— Variability and predictability, *Hydrol. Processes*; 21, 335 – 347, doi:10.1002/hyp.6236.

Chowdhury, M.R., and M.N. Ward (2004): Hydro-meteorological variability in the greater Ganges-Brahmaputra-Meghna basins, *Int. J. Climatol.*, 24, 1495–1508, doi:10.1002/joc.1076.

Chun, J., *et al.*, (2009): Comparative genomics reveals short-term and longterm clonal transitions in pandemic *Vibrio cholerae. Proc. Natl. Acad Sci. U.S.A.*; 106(36), 15,442–15,447, doi:10.1073/pnas.0907787106.

Cockburn T.A., J.G. Cassanos (1960): Epidemiology of endemic cholera. *Public Health Rep* 1960, 75:791-803.

Cohan, F.M. (2001): Bacterial Species and Speciation. *Systematic Biology* **50**:513-524.

Colwell R. R. (1996): Challenge to microbiology in Latin America: a cholera epidemics at the end of the 20th century. *Microbiologia* 12(4): 519-522.

Colwell R.R. (1996): Global climate and infectious disease: the cholera paradigm. *Science*; 274(5295): 2025-2031.

Colwell R.R. and A. Huq (2001): Marine ecosystems and cholera. *Hydrobiologia* 460:141-145.

Colwell R.R., A. Huq (1994): Vibrios in the environment: Viable but Non-Culturable *Vibrio cholerae*. In: Wachsmuth I, Blake P, Olsvik O, eds. *Vibrio cholerae* and Cholera: Molecular to Global Perspectives. *Washington, D.C. ASM Press*; 1:117–33.

Colwell R.R., A. Huq (1999): Global microbial ecology: biogeography and diversity of Vibrios as a model. *J of App Microbiol Symp Suppl*; 85:134S–7S.

Colwell R.R., R.J. Seidler, J. Kaper, *et al*., (1981): Occurrence of Vibrio cholerae serotype O1 in Maryland and Louisiana estuaries. *Appl Environ Microbiol*; 41(2):555–8.

Committee on Emerging Microbial Threats to Health in the 21st Century (2003): Microbial Threats to Health in the United States: Emergence, Detection and Response (eds Smolinski, M. S., Hamburg, M. A. & Lederberg, J.) (National Academy Press, Washington DC.

Committee on Emerging Microbial Threats to Health. Emerging Infections (1992): Microbial Threats to Health in the United States (eds Lederberg, J., Shope, R. E. & Oaks, S. C.) (National Academy Press, Washington DC,).

David M. Morens, Gregory K. Folkers & Anthony S. Fauci (2004): The challenge of emerging and re-emerging infectious diseases NATURE | VOL 430 | 8 JULY 2004 | www.nature.com/nature

David M. Morens, Gregory K. Folkers & Anthony S. Fauci (2004): The challenge of emerging and re-emerging infectious diseases Nature 430: 242-249.

de Magny, G., *et al*. (2008), Environmental signatures associated with cholera epidemics, Proc. Natl. Acad. Sci. U. S. A., 105, 17,676 – 17,681, doi:10.1073/pnas.0809654105.

de Magny, G., R. Murtugudde, P.K. Mozumder, C.J. Grim, N.A. Hasan, M. Naser, M. Alam, R.B. Sack, A. Huq, and R.R. Colwell (2011): Role of Zooplankton Diversity in Vibrio cholerae Population Dynamics and in the Incidence of Cholera in the Bangladesh Sundarbans *APPLIED AND ENVIRONMENTAL MICROBIOLOGY*, Vol. 77, No. 17, p. 6125–6132

Diekmann, O., J. A. P. Heesterbeek, and J.A.J. Metz, (1998): A deterministic epidemic model taking account of repeated contacts between the same individuals. *J. Appl. Prob*. 35, 462–468.

Dietz, H. (1967): Epidemics and Rumours: Asurvey. *J. Roy. Stat. Soc*. A 130: 505-528.

Duval P., G.C. de Ribes, J. Ranjalahy, M.L. Quilici, and J.M. Fournier (1999): Cholera in Madagascar. *Lancet* 1999; 353:2068.

Eames K.T.D. and M.J. Keeling, (2002): Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases. *Proc. Natl Acad. Sci. USA* 99, 13330–13335.

Eames K.T.D. and M.J. Keeling, (2003): Contact tracing and disease control. *Proc. R. Soc. B* 270, 2565–2571. (doi:10. 1098/rspb.2003.2554.)

Ebi K.L., R.S. Kovats, and B. Menne (2006): An approach for assessing human health vulnerability and public health interventions to adapt to climate change. Environ Health Perspect 114:1930–1934.

Echenberg M. (2011): Africa in the Time of Cholera: A History of Pandemics from 1817 to the Present. 1st ed. *New York: Cambridge University Press*, 2011:; 232 p.

Edmunds W. J., C. J. O'Callaghan, and D. J. Nokes (1997b): Who mixes with whom? A method to determine the contact patterns of adults that may lead to the spread of airborne infections. *Proc. R. Soc*. B 264, 949–957. (doi:10. 1098/rspb.1997.0131.)

Edmunds W.J., G.F. Medley and C.J. O'Callaghan, (1997a): Social ties and susceptibility to the common cold. J. *Am. Med. Assoc*. 278, 1231.

Emch M. and Ali M. (2001): Spatial and temporal patterns of diarrheal disease in Matlab, Bangladesh. *Environment and Planning* A 2001, ; 33(2):339-350.

Emch M., C. Feldacker, M. Yunus, P. K. Streatfield, V. D. Thiem, D.G. Canh, and M. Ali (2008): Local Environmental Predictors of Cholera in Bangladesh and Vietnam Am. *J. Trop. Med. Hyg.,* 78(5), pp. 823–832

Emch Michael, Caryl Feldacker, Mohammad Yunus, Peter Kim Streatfield, Vu DinhThiem, Do Gia Canh, and Mohammad Ali (2008): Local Environmental Predictors of Cholera in Bangladesh and Vietnam Am. *J. Trop. Med. Hyg.,* 78(5), pp. 823–832

Emch, J.P. Donnay, M. Yunus, and R.B. Sack (2002): The spatial epidemiology of cholera in an endemic area of Bangladesh. *Soc Sci Med* 55(6):1015-24.

Epstein P. (1993), Algal blooms in spread and persistence of cholera, *Biosystems*, 31, 209–221, doi:10.1016/0303-2647(93)90050-M.

Epstein P.R., T. E. Ford, *et al*., (1993): "Marine ecosystems." *Lancet* ; 342(8881): 1216-1219.

Eubank S., H. Guclu, V.S.A. Kumar, M.V. Marathe, A. Srinivasan, Z. Toroczkai, and N. Wang (2004): Modelling disease outbreaks in realistic urban social networks. *Nature* 429, 180–184.

Fauci, A. S. (2001) Infectious diseases: Considerations for the 21st century. *Clin. Infect. Dis*. 32, 675-685.

Folgosa E., S. Mastrandrea, P. Cappuccinelli, S. Uzzau, P. Rappelli, M.J. Brian, M.M. Colombo (2001): Molecular identification of pathogenicity genes and ERIC types in Vibrio cholerae O1 epidemic strains from Mozambique. *Epidemiol Infect*, 127(01):17-25.

Fraser C., S. Riley, R.M. Anderson, and N.M. Ferguson (2004): Factors that make an infectious disease outbreak controllable. *Proc. Natl Acad. Sci*. USA 101, 6146–6151.

Frisch H. and J.M. Hammersley (1963): Percolation processes and related topics. *J. IAM*, 11 (894918).

Gangarosa E.J. and W.H. Mosley (1974): Epidemiology and surveillance of cholera. Cholera Philadelphia: *WB Saunders*; 1974: 381-403.

Gerstl S. and K. Alberti (2007): Overall response to cholera epidemics in Angola in 2006: Evaluation of the MSF intervention. *Internal report. Paris, France: Epicentre, MSF*,; 2007: 63 p.

Ghani A.C., J. Swinton, and G.P. Garnett, (1997): The role of sexual partnership networks in the epidemiology of gonorrhea. *Sexually Transmitted Diseases*, vol. 24, no. 1, pp. 45–56,

Ghani, A.C. and G.P. Garnett (1998): Measuring sexual partner networks for transmission of sexually transmitted diseases. *J. R. Stat. Soc*. A 161, 227–238.

Gil, A.I., *et al*., (2004): Occurrence and distribution of Vibrio cholerae in the coastal environment of Peru, *Environ. Microbiol*., 6, 699–706, doi:10.1111/j.1462-2920.2004.00601.x.

Glass R.I., S. Becker, M.I. Huq, B.J. Stoll, M.U. Khan, M.H. Merson, J.V. Lee, R.E. Black (1982): Endemic cholera in rural Bangladesh, 1966–1980. *American Journal of Epidemiology* 1982, 116(6):959-970.

Gleick P. (2008): The World's Water 2008-2009: The Biennial Report on Freshwater Resources. *Island Press USA*.

Global Task Force on Cholera Control. Cholera country profile: Angola. Geneva, Switzerland: *World Health Organization*, 2009:1–2. http://www.who.int/cholera/countries/en/. Accessed 15 May 2013.

Goel A. K. and S. C. Jiang (2011): Association of Heavy Rainfall on Genotypic Diversity in Intergovernmental Panel on Climate Change. 2007. Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (Core

Writing Team, Pachauri RK, Reisinger A, eds). Geneva: Intergovernmental Panel on Climate Change.

Halloran M.E., Jr.I.M. Longini, A. Nizam, and Y. Yang (2002): Containing bioterrorist smallpox. *Science* 298, 1428–1432.

Harvell C.D., C.E. Mitchell, J.R. Ward, S. Altizer, A.P. Dobson, R.S. Ostfeld, and M.D. Samuel, (2002): Climate warming and disease risks for terrestrial and marine biota. *Science* 296 (5576):2158-2162.

Hasan NA, Choi S, Eppinger M, Clark PW, Chen A, Alam M, Haley BJ, Taviani E, Hine, H, Su Q, Tallon LJ, Prosper JB, Furth K, Hoq MM, Li H, Fraser-Liggett C, Cravioto A, Huq A, Ravel J, Cebula T, Colwell RR, 2012. Genomic diversity of 2010 Haitian cholera outbreak strains. Proc Natl Acad Sci USA 109: E2010–E2017.

Hashizume, M., B. Armstrong, S. Hajat, Y. Wagatsuma, A. Faruque, T. Hayashi, and D. Sack, (2008): The Effect of Rainfall on the Incidence of Cholera in Bangladesh. *Epidemiology*, 19(1):103-110.

Hay S. I., R.W. Snow, and D.J. Rogers (1998): Predicting malaria seasons in kenya using multitemporal meteorological satellite sensor data. *Transactions of the Royal Society of Tropical Medicine and Hygiene*;, 92(1), 12-20.

Heidelberg J.F., K.B. Heidelberg, and R.R. Colwell (2002): Seasonality of Chesapeake Bay bacterioplankton species. *Appl Environ Microbiol* 2002, 68(11):5488-5497

Hoshino, K., S. Yamasaki, A.K. Mukhopadhyay, S. Chakraborty, A. Basu, S.K. Bhattacharya, G.B. Nair, T. Shimada and Y. Takeda (1998). Development and evaluation of a multiplex PCR assay for rapid detection of toxigenic *Vibrio cholerae* O1 and O139. *FEMS Immunol Med Microbiol* 20(3): 201-207.

Huq A, Colwell RR. Vibrios in the Marine and Estuarine Environment: Tracking of Vibrio cholerae. J Ecosystem Health 1996; 2:198–214. 7.

Colwell RR, Huq A. Global microbial ecology: biogeography and diversity of Vibrios as a model. J of App Microbiol Symp Suppl 1999;85:134S–7S.

Huq A, Small E, West P, Huq M, Rahman R, Colwell R. Ecological relationship between Vibrio cholerae and planktonic copepods. *Appl Environ Microbiol* 1983;45:275–83.

Huq A, Small E, West P, Huq M, Rahman R, Colwell R. Ecological relationship between Vibrio cholerae and planktonic copepods. Appl Environ Microbiol 1983;45:275–83.

Huq A. and R.R. Colwell (1996): Vibrios in the marine and estuarine environment: Tracking *Vibrio Cholerae. Ecosystem Health* 1996, 2:198-214.

Huq A., E. Small, P. West, M. Huq, R. Rahman, R.R., Colwell (1983): Ecological relationship between *Vibrio cholerae* and planktonic copepods. *Appl Environ Microbiol* 1983; 45:275–83.

Huq A., R.B. Sack, A. Nizam, I.M. Longini, G.B.alakrish Nair, A. Ali, J.G. Morris, Jr.M.N. Huda Khan, A. K. Siddique, M. Yunus, M. J. Albert, D.A. Sack, and R.R. Colwell1 (2005): Critical Factors Influencing the Occurrence of Vibrio cholerae in the Environment of Bangladesh. *Applied and Environmental Microbiology:* Vol. 71, No. 8, p. 4645–4654

Huq Anwar, R. Bradley Sack, Azhar Nizam, Ira M. Longini, G. Balakrish Nair4 Afsar Ali, J. Glenn Morris, Jr., M. N. Huda Khan, A. Kasem Siddique, Mohammed Yunus, M. John Albert, David A. Sack, and Rita R. Colwell1 2005. Critical Factors Influencing the Occurrence of Vibrio cholerae in the Environment of Bangladesh. *Applied and Environmental Microbiology:* Vol. 71, No. 8, p. 4645–4654

Huq, A. and R.R. Colwell (1996): Vibrios in the Marine and Estuarine Environment: Tracking Vibrio

Huq, A., *et al*., (1984), Influence of water temperature, salinity, and pH on survival and growth of toxigenic Vibrio cholerae serovar O1 associated with live copepods in laboratory microcosms, Appl. Environ. *Microbiol*., 48, 420–424.

Huq, A., *et al*., (1990): Detection of *Vibrio cholerae* O1 in the aquatic environment by antibody and culture, Appl. Environ. *Microbiol*., 56, 2370–2373.

Huq, A., *et al*., (2005): Critical factors influencing the occurrence of Vibrio cholerae in the environment of Bangladesh, *Appl. Environ. Microbiol*.: 71, 4645–4654, doi:10.1128/AEM.71.8.4645-4654.2005.

International Institute for Population Sciences and Macro International (2007): National Family Health Survey (NFHS-3), 2005–06: India. Vol1, Mumbai International Institute for Population Sciences and Macro International.

Islam D., M.D. Lewis, A. Srijan, L. Bodhidatta, A. Aksomboon, M. Gettayacamin, S. Baqar, D. Scott, and C.J. Mason (2006): Establishment of a non-human primate Campylobacter disease model for the pre-clinical evaluation of Campylobacter vaccine formulations. *Vaccine*, 24: 3762-3771.

Islam M., A. Brooks, M. Kabir, G. Nair, W. Yukiko, and S. Luby (2006): Faecal contamination of drinking water sources of Dhaka city, *J. Appl. Microbiol*., 103, 80– 87, doi:10.1111/j.1365-2672.2006.03234.x.

Islam M.S. *et al*., (1994): Detection of nonculturable *V. cholerae* O1 associated with a cyanobacterium from an aquatic environment in Bangladesh. *Trans R Soc Trop Med Hyg* 88: 298-99.

Islam M.S., B.S. Drasar, and R.B. Sack (2003): The aquatic environment as a reservoir of Vibrio cholerae: a review. J Diarrhoeal Dis Res 1993, 11(4):197-206. *Annurnal of Infectious Diseases* 2003, 187(1):96-101.

Islam M.S., K.A. Talukder, N.H. Khan, Z.H. Mahmud, M.Z. Rahman, G.B. Nair, A.K. Siddique, M. Yunus, D.A. Sack, R.B. Sack, A. Huq, R.R. Colwell, (2004): Variation of toxigenic Vibrio cholerae O1 in the aquatic environment of Bangladesh and its correlation with the clinical strains. *Microbiol Immunol* 48: 773–777.

Islam S.N., and A. Gnauck (2008): Mangrove wetland ecosystems in Ganges-Brahmaputra Delta in Bangladesh, Front. *Earth Sci. China*, 2(4), 439–448, doi:10.1007/s11707-008-0049-2.

Islam, D., Lewis, M.D., Srijan, A., Bodhidatta, L., Aksomboon, A., Gettayacamin, M., Baqar, S., Scott, D., Mason, C.J. (2006): Establishment of a non-human primate Campylobacter disease model for the pre-clinical evaluation of Campylobacter vaccine formulations. *Vaccine*, 24: 3762-3771.

Islam, M., A. Brooks, M. Kabir, G. Nair, W. Yukiko, and S. Luby (2006), Faecal contamination of drinking water sources of Dhaka city, *J. Appl. Microbiol.*, 103, 80– 87, doi:10.1111/j.1365-2672.2006.03234.x.

Jutla A., E. Whitcombe, N. Hasan, B. Haley, A. Akanda, A. Huq, M. Alam, R.B. Sack, and R.R. Colwell (2013): Environmental Factors Influencing Epidemic Cholera *Am. J. Trop. Med. Hyg*.

Jutla A.S., A.S. Akanda and S. Islam, (2009): Relationship between Phytoplankton, Sea Surface Temperature and River Discharge in Bay of Bengal. *Geophysical Research Abstracts:* Vol. 11, EGU2009-1091-2, EGU General Assembly (2009), Vienna, Austria.

Jutla S., A.S. Akanda , J.K. Griffiths, R.R. Colwell, and S. Islam (2011): Warming Oceans, Phytoplankton, and River Discharge: Implications for Cholera Outbreaks *Am. J. Trop. Med. Hyg.*, 85(2), pp. 303–308

K. Rajendran, A. Sumi, M.K. Bhattachariya, B. Manna, D. Sur, N. Kobayashi, and T. Ramamurthy (2011): Influence of relative humidity in Vibrio cholerae infection: A time series model. *Indian J Med Res*; 133(2): 138–145.

Kaneko T, Colwell RR (1975): Adsorption of *Vibrio parahaemolyticus* onto chitin and copepods. Appl. Microbiol; 29(2):269–74.

Kaneko T, Colwell RR (1975): Incidence of *Vibrio parahaemolyticus* in Chesapeake Bay. Appl Microbiol; 30(2):251–7.

Kaneko T. and, R.R. Colwell (1975): Adsorption of *Vibrio parahaemolyticus* onto chitin and copepods. *Appl. Microbiol*; 29(2):269–74.

Kaper J.B., J.G. Jr. Morris, and M.M. Levine (1995): Cholera. *Clinical Microbiology Reviews*, 8(1):48-86.

Keeling M.J. and K.T.D. Eames (2005): Networks and epidemic Models. *Interface;* 2: 295- 307.

Kendall, M. (1938): A New Measure of Rank Correlation. *Biometrika* 30 (1–2): 81–89. doi:10.1093/biomet/30.1-2.81. JSTOR 2332226.

Kermack W.O. and A.G. McKendrick (1927, 1991a): Contributions to the mathematical theory of epidemics, part I, Proc. Roy. Soc. Edin. A 115, 700-721.

Khan M.U., A.R. Samadi, M.I. Huq, M. Yunus, and A Eusof (1984): Simultaneous classical and El Tor cholera in Bangladesh. *J Diarrhoeal Dis Res*, 2(1):13-18.

Kindhauser, M.K. (2003): Global defense against the infectious disease threat. World Health Organization, *Geneva*., 74- 79.

Klovdahl A.S. (1985): Social networks and the spread of infectious diseases: the AIDS example. *Soc. Sci. Med*. 21, 1203–1216.

Koelle K., X. Rodo, M. Pascual, M. Yunus, and G. Mostafa (2005): Refractory periods and climate forcing in cholera dynamics. *Nature*, 436, 696–700, doi:10.1038/nature03820.

Kretzschmar M., Y.T.H.P. van Duynhoven, and A. J. Severijnen (1996): Modeling prevention strategies for gonorrhea and chlamydia using stochastic network simulations. *Am. J. Epidem*. 144, 306–317.

Kruskal and H. William (1958): Ordinal Measures of Association *Journal of the American Statistical Association*. Retrieved 2012-11-04.

Kuperman M. and G. Abramson (2001): Small world effects in an epidemiological model. *Phys. Rev. Lett*. 86, 2909–2912.

Lam C., S. Octavia, P.R. Reeves, and R. Lan (2012): Multi-locus variable number tandem repeat analysis of 7th pandemic Vibrio cholerae. *BMC Microbiol* 2012; 12:82.

Lam C., S. Octavia, P.R. Reeves, L. Wang, and R. Lan (2010): Evolution of seventh cholera pandemic and origin of 1991 epidemic, Latin America. *Emerg Infect Dis* 2010; 16:1130–1132.

Lee JW. (2004): *Information productions on water, sanitation and health*. Geneva: World Health Organization; *catalogue water, sanitation and health*. PMCID: PMC3089044

Linthicum K. J., A. Anyamba, C.J. Tucker, P.W. Kelley, M.F. Myers, and C.J. Peters, (1999): Climate and satellite indicators to forecast rift valley fever epidemics in kenya. Science, 285(5426), 397-400.

Lipp E.K., A. Huq, and R.R. Colwell (2002): Effects of global climate on infectious disease, Clin. Microbiol. Rev., 15(4), 757–770, doi:10.1128/ CMR.15.4.757-770.2002.

Lloyd A.L. (2001): Destabilization of epidemic models with the inclusion of realistic distribution of infectious periods. *Proc. Roy. Soc. Lond, B* 268:985-993.

Lobitz, B.M., L.R. Beck, A. Huq, B. Wood, G. Fuchs, A.S.G. Faruque and R.R. Colwell, (2000), Climate and infectious disease: use of remote sensing for detection of Vibrio Cholerae by indirect measurement, PNAS, 97(4), 1438-1443.

Louis, V., *et al*. (2003), Predictability of Vibrio cholerae in Chesapeake Bay, Appl. Environ. Microbiol., 69, 2773–2785.

Magny, G., R.Murtugudde, M.Sapianob, A. Nizam, C. Brown, A. Busalacchi, M. Yunus, G.Nair, A. Gil, J. Calkins, B. Manna, K. Rajendran, M. Bhattacharya, A. Huq, R.Sack, and R.R. Colwell, 2008. Environmental signatures associated with cholera epidemics, *PNAS,* 105(46), doi: 10.1073/pnas.0809654105.

Magny, G., R.Murtugudde, Pronob K. Mozumder, Christopher J. Grim, Nur A. Hasan, M. Niamul Naser, Munirul Alam, R. Bradley Sack, Anwar Huq, and Rita R. Colwell 2011. Role of Zooplankton Diversity in Vibrio cholerae Population Dynamics and in the Incidence of Cholera in the Bangladesh Sundarbans APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Vol. 77, No. 17, p. 6125–6132

Malon, J. B., J. M. Yilma, J.C. McCarroll, B. Erko, S. Mukaratirwa, and X. Zhou, (2001): Satellite climatology and the environmental risk of schistosoma mansoni in Ethiopia and East Africa. *Acta Tropica*, 79(1), 59-

Margaret C., P. Voss, P.K. Streatfield, M. Yunus, and M. Emch (2010): Protection from annual flooding is correlated with increased cholera prevalence in Bangladesh: a zero-inflated regression analysis. *Environ Health*; 9: 13. PMCID: PMC2856547

Martin A.R., W.H. Mosely, B.B. Sau, S. Ahmed, and I. Huq (1969): Epidemiologic analysis of endemic cholera in urban East Pakistan, 1964– 1966. *American Journal of Epidemiology* 1969, 89(5):572.

Matsuda F., S. Ishimura, Y. Wagatsuma, T. Higashi, T. Hayashi, A.S.G. T., Faruque, D.A. Sack, M. Nishibuchi, (2008): Prediction of epidemic cholera due to *Vibrio cholerae* O1 in children younger than 10 years using climate data in Bangladesh. *Epidemiology and Infection*, 136 (1):73-79.

McCormack W.M., W.H. Mosley, M. Fahimuddin, and A.S. Benenson (1969): Endemic Cholera in Rural East Pakistan. *American Journal of Epidemiology* 1969, 89(4):393.

McDougald D, Rice SA, Weichart D, Kjelleberg S. Nonculturability: adaptation or debilitation? FEMS Microb Ecol 1998;25(1):1–9.

McDougald D., S.A. Rice, D. Weichart, and S. Kjelleberg (1998): Nonculturability: adaptation or debilitation? *FEMS Microb Ecol* 1998; 25(1):1–9.

Merson M.H., R.E. Black, M.U. Khan, and I. Huq (1980): Enterotoxigenic Escherichia coli diarrhea: acquired immunity and transmission in an endemic area. Cholera and related diarrheas: molecular basis of a global health problem 43rd Nobel Symposium S Karger, Basel, *Switzerland* 1980: 34-45.

Meyers L. A., B. Pourbohloul, M. E. J. B., Newman, D. M. Skowronski, and R.C. Brunham, (2005): Network theory and SARS: predicting outbreak diversity. *J. Theor. Biol*. 232, 71–81.

Mhalu F.S., W.M. Mntenga, and F.D. Mtango (1987): Seasonality of cholera in Tanzania: possible role of rainfall in disease transmission. East Afr Med J 1987, 64(6):378-387.

Miller C.J., B.S. Drasar, and R.G. Feachem, (1982): Cholera and estuarine salinity in Calcutta and London. *Lancet*; 1 (8283):1216-1218.

Mirza M., R. Warrick, N. Ericksen, and G. Kenny (2001): Are peak flows getting worse in the Ganges, Brahmaputra and Meghna basins?, *Global Environ. Change, Part B*, 3, 37–48, doi:10.1016/S1464-2867(01)00019-5.

Moore C. and M. Newman (2000): Epidemics and percolation in small-world networks. *Phys. Rev. E;* 61, 5678–5682.

Morens D., M. Morens, G.K. Folkers and A. S. Fauci (2004): The challenge of emerging and re-emerging infectious diseases Nature 430: 242-249

Muller J., M. Kretzschmar, and K. Dietz (2000): Contact tracing in stochastic and deterministic epidemic models. *Math. Biosci*. 164, 39–64.

Nielsen and T. Gissel (1994): Regulation of zooplankton biomass and production in a temperate, coastal ecosystem. Limnol. *Oceanogr*, 39(3), 508-519

Olsen J.S., T. Aarskaug, G. Skogan, E.M. Fykse, A.B. Ellingsen, and J.M. Blatny (2009): Evaluation of a highly discriminating multiplex multi-locus variable-number of tandem repeats (MLVA) analysis for Vibrio cholerae. *J Microbiol Methods* 2009; 78:271–285.

Olsson T. (1996): Malaria and cholera are traced by satellites. *Lakartidningen* 93(36):3037-40.

Pascual M., L.F. Chaves, B. Cash, X. Rodó, and M. Yunus (2008): Predicting endemic cholera: The role of climate variability and disease dynamics. *Climate Research* 36 (2), pp. 131-140.

Pascual, M., M. Bouma, and A. Dobson (2002), Cholera and climate: Revisiting the quantitative evidence, Microbes Infect., 4, 237–245, doi:10.1016/S1286-4579(01)01533-7.

Paz S. (2009): Impact of temperature variability on cholera incidence in southeastern Africa, 1971-2006. *Ecohealth*. 2009 Sep; 6(3):340-5. Epub 2009 Dec 29.

Pearson K. (1900): On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. Philosophical Magazine Series 5 50 (302): 157–175. doi:10.1080/14786440009463897.

Pollitzer R. (1959): *Cholera*. Geneva: World Health Organization.
Pollitzer, R. (1959): Cholera, O1 profile. World Of Health Organization monograph no. 43. World Health Organization, Geneva.

Pollitzer, R. (1959): Cholera, profile_O1. World of Health Organization monograph no. 43. *World Health Organization, Geneva*

Pruss A., D. Kay, L. Fewtrell, and J Bartram (2002): Estimating the burden of disease from water, sanitation, and hygiene at a global level. *Environ Health Perspect*.;110:537–42.
Rahman, M., *et al*., (2000), Environmental impact assessment on water quality deterioration caused by decreased Ganges flow, *Environ. Geol*., 40, 31–40, doi:10.1007/s002540000152.

Rajendran K., T. Ramamurthy, S.K. Bhattacharya (2008): Log-linear model to assess socioeconomic and environmental factors with childhood diarrhoea using hospital based surveillance. *J Mod Appl Stat Methods*;7:304–13.
Read J. M. and M. J. Keeling (2003): Disease evolution on networks: the role of contact structure. *Proc. R. Soc*. B 270, 699–708. (doi:10.1098/rspb.2002.2305.)

Riley, S. *et al*., (2003): Transmission dynamics of the etiological agent of SARS in Hong Kong: impact of public health interventions. *Science* 300, 1961–1966.

Rivera I.N., E.K. Lipp, A. Gil, N. Choopun, A. Huq and R.R. Colwell (2003):. "Method of DNA extraction and application of multiplex polymerase chain reaction to detect toxigenic Vibrio cholerae O1 and O139 from aquatic ecosystems." *Environ Microbiol* 5(7): 599-606.

Rogers L. (1926): The conditions influencing the incidence and spread of cholera in India. *Proc R Soc Med*.;19:59–93.

Rogers L. (1928): The incidence and spread of cholera in India; Forecasting and control of epidemics, No.9. *Indian Med Res Mem*.

Roy DK. (1959): A note on the incidence of cholera in the Chetla area for the period between November 1957- June 30, 1958. *Indian J Public Health*; 3:33–7.

Russel AJH. *Transactions of the 7 th Congress of the Far Eastern Association Tropical Medicine*. vol. 2. Calcutta: 7 th. Statistical studies in the epidemiology of cholera; p. 131.

Russell AJH. (1928): *Indian Med Res Mem* No. 12.. The epidemiology of cholera in India.

Sack R.B., A.K. Siddique, I.M. Longini, A. Nizam, M. Yunus, M.S. Islam, J.G. Morris, A. Ali, A. Huq, G.B. Nair, *et al*., (2003): A 4-year study of the epidemiology of Vibrio cholerae in four rural areas of Bangladesh. *J Infect Dis*; 187(1):96-101.

Safa A., G.B. Nair, and R.Y.C. Kong (2010): Evolution of new variants of Vibrio cholerae O1. *Trends Microbiol*. 2010; 18:46–54.

Samadi A.R., M.K. Chowdhury, M.I. Huq, and M.U. Khan (1983): Seasonality of classical and El Tor cholera in Dhaka, Bangladesh: 17-year trends. *Trans R Soc Trop Med Hyg* 1983, 77(6):853-856.

Schott, J.R. (2007): Remote sensing: the image chain approach (2nd ed.). *Oxford University Press*;. p. 1. ISBN 978-0-19-517817-3.

Schowengerdt, Robert A. (2007). Remote sensing: models and methods for image processing (3rd ed.). Academic Press. p. 2. ISBN 978-0-12-369407-2.

Schwartz, B., J. Harris, A. Khan, R. LaRocque, D. Sack, M. Malek, A. Faruque, and E. Ryan (2006):, Diarrheal epidemics in Dhaka, Bangladesh, during three consecutive floods: 1988, 1998, and 2004, *Am. J. Trop. Med. Hyg*., 74, 1067–1073.

Scott, J. (1991): Social network analysis: a handbook. London: *SAGE Publications*.

Sharma N.C., P.K. Mandal, R. Dhillion, and M. Jain (2007): Changing profile of V. Cholerae O1, O139 in Delhi & its periphery (2003-2005). Indian J Med Res 2007; 125: 525-32

Singleton, P.W., S.A. Swaify and B.B. Bahlool, (1982): Effect of salinity on Rhizobium growth survival. *Appl. Environ. Microbial*., 44: 884-890.

Stephens, D. S. *et al*., (1998): Emerging and re-emerging infectious diseases: a multidisciplinary perspective. *Am. J. Med. Sci*. 315, 64–75.

Stine OC, Alam M, Tang L, Nair GB, Siddique AK, Faruque SM, Huq A, Colwell R, Sack RB, Morris JG Jr, 2008. Seasonal cholera from multiple small outbreaks in rural Bangladesh. Emerg Infect Dis 14: 831–833.

Szendroi, B. and G. Csanyi (2004): Polynomial epidemics and clustering in contact networks. *Proc. R. Soc. B* 271, S364–S366. (doi:10.1098/rsbl.2004.0188.)

Tamplin M.L., A.L. Gauzens, A. Huq, D.A. Sack, R.R. Colwell (1990): Attachment of *Vibrio cholerae* serogroup O1 to zooplankton and phytoplankton of Bangladesh waters. *Appl Environ Microbiol*; 56(6):1977–80.

Tauxe R, Seminario L, Tapia R, Libel M. (1994): The Latin American epidemic. In *V.cholerae and Cholera*: Molecular to Global Perspectives (Wachsmuth K, Blake PA, Olsvik O, eds) American Society for Microbiology, Washington, D.C., pp321-44

Tromp SW. (1963): Medical biometeorology. 18th ed. Amsterdam: Elsevier Pub. *V. cholerae* isolates from an Outbreak in India. International Journal of Microbiology, Volume 2011, Article ID 230597, 5 pages doi:10.1155/2011/230597

Wahid S., M. Babel, and A. Bhuiyan (2007): Hydrologic monitoring and analysis in Sundarbans mangrove ecosystem, Bangladesh, *J. Hydrol*., 332, 381–395, doi:10.1016/j.jhydrol. 2006.07.016.

Waldor M. K. and J. J. Mekalanos (1996): "Lysogenic conversion by a filamentous phage encoding cholera toxin." *Science* 272: 1910-1914.

Wallinga J., W.J. Edmunds, and M. Kretzschmar (1999) Perspective: human contact patterns and the spread of airborne infectious diseases. *Trends Microbiol*. 7, 372–377.

Wardlaw T., P. Salama, C. Brocklehurst, M. chopra, and E. Mason (2010): Diarrhoea: why children are still dying and what can be done. *Lancet*. 375:870–2.

Wasserman S. and K. Faust (1994): Social network analysis. *Cambridge*: Cambridge University Press.

Watts D.J. and S. H. Strogatz, (1998): Collective dynamics of 'small-world' networks. *Nature* 393, 440–442.

WHO (1995): Cholera in 1994. *Wkly. Epidemiol. Rec*., 70: 201- 208.

WHO Regional Office for Africa (2011): Disease Prevention and Control Cluster, Epidemic and Pandemic Alert and Response Programme. Ongoing outbreaks. Lassa Fever, Yellow Fever and Cholera in Ghana. *Outbreak Bulletin* 2011; 1:3.

WHO, (2004): the global burden of diseases: update 2008. Available from: http://www.who.int/healthinfo/global_burden_isease/GBD_report_2004update_full.pdf, accessed on February 2, 2011.

WHO. Cholera, (2006). *Wkly Epidemiol Rec* 2007; 82:273–84.

WHO. Cholera, 2007. Wkly Epidemiol Rec 2008; 83:269–83. Aidara A, Koblavi S, Boye CS, *et al*., Phenotypic and genotypic characterization of Vibrio cholerae isolates

from a recent cholera outbreak in Senegal: comparison with isolates from Guinea-Bissau. *Am J Trop Med Hyg* 1998; 58:163–7.

WHO: Cholera in 1997. *Wkly Epidemiol Rec* 1998, 73(27):201-208.

Worden, A.Z., M. Seidel, A. Wick, F. Malfatti, D. Bartlett, and F. Azam (2006): Trophic Regulation of Vibrio Cholerae in Coastal Marine Waters. *Environ. Microbiol.* 8(1):21-29.

Wylie J. L., L. Shah, and A. Jolly, (2007): Incorporating geographic settings into a social network analysis of injection drug use and blood borne pathogen prevalence. *Health and Place,* vol. 13, no. 3, pp. 617–628.

Wylie, J. L. and A. Jolly (2001): Patterns of chlamydia and gonorrhea infection in sexual networks in Manitoba, Canada. *Sex. Transm. Dis*. 28, 14–24.

You D., T. Wardlaw, P. Salama, and G. Jones (2010): Levels and trends in under-5 mortality, 1990-2008. *Lancet*;375:100–3.
Yu C., A.M. Lee, B.L. Bassler, and S. Roseman (1991): S. Chitin Utilization by Marine-Bacteria – a Physiological-Function for Bacterial Adhesion to Immobilized Carbohydrates. J Biol Chem 1991; 266(36):24260–7.

Zuckerman J.N., L. Rombo, and A. Fisch (2007): The true burden and risk of cholera: implications for prevention and control. *Lancet Infect Dis* 7:521–530.