



Examination of models for cholera: insights into model comparison methods

Olcay Akman^a, Marina Romadan Corby^b and Elsa Schaefer^c

^aDepartment of Mathematics, Illinois State University, Normal, IL, USA; ^bDepartment of Mathematics, Computer Science and Statistics, University of Illinois at Chicago, Chicago, IL, USA; ^cDepartment of Mathematics, Marymount University, Arlington, TX, USA

ABSTRACT

This article provides an overview of the Akaike and Bayesian Information Criteria as applied to the setting of deterministic modelling, with the perspective that this may be a useful tool for biomathematics researchers whose primary interests lie in the analysis of compartmental models. We additionally examine a wide range mechanistic and parameter assumptions in the cholera literature through the unifying lens of model selection criteria. Five models for cholera are considered using multiple model selection formulations, and implications for cholera modelling and for model selection criteria are discussed.

ARTICLE HISTORY

Received 13 March 2016
Accepted 16 June 2016

KEYWORDS

Model comparison; AIC; cholera

1. Introduction

We have two primary goals in this work: to share with the biomathematics community a statistically valid framework for the comparison of deterministic models and, as a case study, to express some particular mechanistic and parameter-related concerns that we have gleaned from many years of cholera study. To this end, we put forward a method for model comparison and contrast this technique for model selection with others by measuring them against data from the first year of the cholera outbreak that followed the 2010 Haitian earthquake. First, in Section 2 we introduce and discuss the Akaike Information Criterion (AIC) as a basis for model comparison. This section is written for a diverse group of researchers who are consumers of model selection theory, as opposed to those whose sole research focus may be model assessment. Additionally, we provide a brief overview of the Bayesian Information Criteria (BIC). Section 3 begins with a thorough review of how the spread of cholera might be modelled from a mechanistic viewpoint, and highlights disagreements in the literature regarding key parameter assumptions. Here we also describe five models to be compared using the AIC and BIC criteria against data from the recent outbreak in Haiti. In Section 4, we share some methods and results from prior model selection investigations for cholera, as well as an investigation not intended for cholera that provides a different point of view for model selection methodology.

CONTACT Elsa Schaefer  elsa@marymount.edu

Finally, we discuss the methods and results from our model selection efforts for cholera in Section 5, and provide a brief conclusion in Section 6.

2. An introduction to maximum likelihood estimators and AIC

Model formulation involves judgement, experience, trial and error. The validity of a model should (i) be consistent with knowledge of the system under study, (ii) extrapolate to related sets of data and (iii) have reasonable mathematical and statistical properties. Often subject-matter considerations suggest a stochastic argument for a range of suitable models. At other times, however, an analysis can only be realistically completed for one of several competing models, and hence a basis for the selection of a single model is needed.

A precept that is usually implied but is often unstated is the *principle of parsimony*: 'it is vain to do with more what can be done with fewer'. That is, we should prefer simpler models over more complex ones that fit our data about equally well. But what exactly does this mean? If we have models with one, two and three parameters and respective errors of, say, 100, 2 and 1 in fitting the data, then the second clearly improves on the first, but do the second and the third perform 'about equally well'? Model selection criteria are intended to balance the increased model complexity with any improvement in fit to make this determination.

Before we scrutinize various criteria for model selection from a statistical point of view, we must first define the word 'model'. Do two different parameter values applied to the same mechanistic structure define two different models, or are two models viewed as different because they have different mathematical structures, regardless of the particular values of the parameters within each model? Usually, when the term 'model' is used in a statistical article about model selection, it refers to a probability distribution (population) from which data can be randomly sampled, and the question is, 'From which true probability distribution were the data that I observed sampled?' However, those of us who work with mechanistic models in systems biology typically employ a different meaning for the term 'model' – mostly, we consider a system of differential equations paired with an appropriate parameter set to be a 'model'. In this article we will use 'model' to refer to a mathematical structure paired with a particular set of parameters, and we will refer to a 'family' of models as a mathematical structure whose family members are equipped with differing feasible parameter choices.

A second point to be considered when assessing the relative values of two families of models is that the underlying mechanistic modelling assumptions may be wildly different between models, making the two options not equally suited for a particular application. A third important point is that we may not wish to treat the models as the 'most true' from an application viewpoint after all. The best example of this would be neural network models, where the fitted model is a 'black box', whose contents have no intrinsic interest or meaning but are merely used for predictions.

In this introduction, we seek to reconcile our systems biology point-of-view with the statistical model comparison point-of-view, and to define a statistically accurate process under which several mechanistic models for systems may be compared. Given a particular data sample and asked to find the best approximating model, we encounter two types of errors: (i) an error caused by the choice of the family of models, and (ii) an error due to the estimated parametrization of the model family. This concept can be formulated using

the concept of *Overall Risk*, where

$$\text{overall risk} = \text{risk of modelling} + \text{risk of estimation}.$$

Often, during the course of the analysis of data, we may discover that the particular form of the specified model is not appropriate for the data at hand. In this case, the *risk of modelling* will be a larger portion of the *overall risk* due to the incorrect specification of the model. In fact, the correct specification of the model is a necessary but by no means a sufficient condition for it to be selected. For a given model family, the *risk of estimation* is encountered while we estimate the true parameter(s). When the true parameter values are not included in the search space, then a *bias* is caused. Our goal in model selection is to minimize the overall risk, and this is where model assessment measures are used.

2.1. Formulation of the AIC for deterministic model comparison

In the following paragraphs, we will introduce the AIC for model selection. Before we explore the details, let us view the numerical value assigned to the AIC intuitively:

$$\text{AIC} = \text{lack of fit} + \text{penalty for model complexity}. \quad (1)$$

The *lack of fit* term, as the name implies, decreases when the model is more likely to describe the data. In comparison, the *penalty for complexity*, determined by the number of free parameters within the model, increases for more complicated models. The number of free parameters is considered to be a measure of complexity, and is a compensation for the bias in the lack of fit when the maximum likelihood estimates are used. Note that a poor model choice (*risk of modelling*) will increase the lack of fit if it describes the data poorly, or may instead/additionally have a high penalty for complexity. If a model's parameters are poorly estimated (*risk of estimation*), this would cause the lack of fit to increase as well. Thus, AIC demands a compromise between the maximized likelihood (the lack of fit component) and the number of free parameters estimated within the model (the penalty component). Hence, when AIC is applied to models from the same family, the penalty for complexity has no variability. On the other hand, the number of parameters become an important factor when comparing models with varying numbers of parameters.

When we are comparing models by the simple expedient of fitting them to data, we must first accept that there are no true models. Indeed, models only approximate reality. The problem is then to determine which model would best approximate reality, given the data we have recorded. In other words, the best we can do is to minimize the loss of information. In itself, the value of the AIC for a given data-set has no meaning. It becomes useful when it is compared to the AIC of a collection of predetermined models, the model with the lowest AIC being the *best* model among all models specified for the data at hand. If only poor models are considered, then AIC will select the best of the poor models. For family of non-deterministic models, as the sample size gets large, the ability to choose 'good' parameters improves, and first term in the AIC formulation – the likelihood – increases but the penalty term does not, since the penalty only depends on the (fixed) number of parameters. This means that the penalty term has little effect if n , the number of points in the data-set, is large. On the one hand, it is intuitively clear that a complex model cannot be justified

to describe a small amount of data; however, as the size of the data increases, objections have been raised that minimizing AIC may not point to the correct model (Bhansali & Downham, 1977; Schwarz, 1978; Woodroffe, 1982). Therefore, critics say, the importance of AIC as a model selection criterion should not be exaggerated (Forster, 1999; Hannan, 1986).

We now wish to investigate the AIC methodology from a statistician's point of view so that we may derive how AIC may properly be applied in a deterministic setting. Suppose that we have a random sample Y_1, \dots, Y_n from the true unknown model $g(y)$. Given this random sample of size n from the true distribution defined by g and a specific family of candidate models $f(y; \theta)$, we may calculate the AIC. The work begins by making statistical inferences about the parameters for the model family f . Hence, we proceed by fitting the candidate model $f(y; \theta)$ by maximizing the conditional probability that f describes the data, given that the parameter is θ ; i.e. we maximize the likelihood $\mathcal{L}(\theta) = \prod f(y_i; \theta)$. Now, let $\hat{\theta}$ denote the corresponding value of θ that maximizes the likelihood $\mathcal{L}(\theta)$. With the optimal parameter $\hat{\theta}$ for the family f selected, an assessment of the quality of the constructed model needs to be made. The form that the AIC described first in Equation (1) takes in this context is

$$\text{AIC} = -2 \ln \mathcal{L}(\hat{\theta}) + 2p, \quad (2)$$

where $\mathcal{L}(\hat{\theta})$ is the maximized likelihood function and p is the number of free parameters in the model. We note that in this formulation, the coefficients in front of the terms in the summation simply weight the terms equally; however, in the following formulation (3), equal weighting would result in the selection of an overly-simple model. At this point of the modelling process, we have simply considered a single family of models, and, as we described previously, AIC is not yet a valuable measure. When the AIC value is calculated for multiple model families, the chosen model is the model-parameter combination with the minimum AIC. The negative sign in Equation (2) ensures that higher likelihood contributes to a smaller AIC, while an increased penalty contributes to a higher AIC and the factor 2 puts the differences for AIC on the same scale as 'likelihood ratio statistics', another method for model comparison.

In biomathematics settings, where compartmental models such as *Susceptible-Infected-Removed* (SIR) are prevalent, the implementation of AIC as a model 'quality' measure fundamentally differs from its implementation in describing statistical models, where the variation is partially due to the randomness inherent in data being drawn from a population (distribution). In the latter setting, the probability density function (pdf), and hence the likelihood function, specifies the probability of observing the true data vector given a particular parameter vector. Given a set of parameter values, the corresponding pdf will show that some data are more probable than other data. In reality, however, we have already observed the data, and hence we are faced with the inverse problem: given the observed data and a model of interest, find one pdf, among all the probability densities that the model describes, that is most likely to have produced the data. To solve this inverse problem, we define the likelihood function by reversing the roles of the data vector and the parameter vector. The principle of maximum likelihood estimation states that the desired probability distribution is the one that makes the observed data 'most likely'.

However, when fitting a deterministic compartmental model, the very idea of drawing a random sample from a prespecified distribution is no longer applicable. The basic foundation of a compartmental model selection process is the determination of the parameters of

a model based on how well they estimate the response variable of interest, and not on the likelihood of observing the parameters. In the light of our earlier discussion on how AIC works, consider a given family of deterministic models, such as the epidemiological SIR family, and a given data-set which the family is to describe when paired with appropriate parameter choices. Given the data-set, and a particular set of parameter estimates, model fitness then corresponds to the prediction error of the response variable (the output variable of interest). Clearly, different sets of parameter estimates will result in different prediction errors.

Thus, referring to our intuitive AIC definition in Equation (1), the *lack of fit* for a deterministic model can be measured by the residual sum of squares (RSS) between the model prediction and the observed data. Our goal in parameter estimation is to minimize the error of fit, and this is certainly different from the traditional likelihood approach, where we try to maximize the joint probability of a sample being drawn from a particular statistical population.

Now, let us assume that we are comparing several deterministic model families, such as SIR versus SIRB (containing an environmental reservoir) from Section 3 below, so that each model may potentially have a different number of parameters. We then add the *penalty for model complexity* to the AIC formulation similar to the original formulation given in Equation (2). When measuring the error of fit from a deterministic model, the AIC can be computed by Burnham and Anderson (2002)

$$\text{AIC} = n \ln \left(\frac{\text{RSS}}{n} \right) + 2K. \quad (3)$$

Again, RSS is the sum of the model prediction squared errors for a particular candidate model that was the winning member of its respective family, and K is the total number of estimated parameters p plus 1, because the variance of the observed error is considered to be an additional parameter for the penalty term. The division by n is by convention and clearly the deletion of this term would simply translate all results equally. As we mentioned following Equation (2), a different coefficient is needed for the *lack of fit* term in Equation (3) versus in the likelihood formulation given in Equation (2). First, since the minimum AIC should select the preferred model and a decreased RSS already indicates an improved model, a negative sign would not be appropriate. Second, the ‘balancing term’ must now be the number of observations, and not 2 as in the previous AIC equation. As we will discuss in Sections 4 and 5, while there is a mathematically equivalent ‘likelihood’ derivation to the measure of *lack of fit* in Equation (3) above, the interpretation of likelihood in a traditional manner is not supported.

In Section 5, we will compare several implementations of the formulation given in Equation (3) with varying approaches to measuring the *lack of fit*. In the same manner, we will consider a likelihood formulation similar to that in Equation (2). Finally, we will examine an altogether different comparison measure that is briefly described in the following section.

2.2. AIC versus BIC for biological models

In the previous section we explained the framework for the AIC. Another tool too may prove useful in model comparison is the BIC. BIC was derived to provide a consistent estimator of the order/dimension of the true model. While the AIC seeks to measure the amount of information lost when a function f is used to approximate the true data generating function g , BIC assumes that one, true, fixed model exists within the set of candidate models, and that as the sample size n increases, the probability of selecting the true model approaches 1. However, oftentimes the sample size required for the consistent criteria to point to the true model with a reasonably high probability must be very large (in the thousands) (Burnham & Anderson, 2002). Thus, Burnham and Anderson explain that the primary assumptions needed for BIC comparisons do not apply in biological sciences or any other ‘noisy’ science (Burnham & Anderson, 2002). Nonetheless, some in mathematical ecology do consider the BIC in evaluating models. We will discuss a contrasting viewpoint to the utility of BIC in Section 4.4 by presenting equations for BIC in several contexts, and investigate the value of BIC results in Section 5.

3. Case study: multiple models for a cholera outbreak

3.1. A mechanistic consideration of the spread of cholera

The inspiration for the current investigation has its origins in a desire to improve and validate several mechanistically-driven models for cholera. With recent significant outbreaks in Zimbabwe and later Haiti, there has been much progress in the development of useful cholera models. The ultimate goal of these modelling efforts is, of course, to understand the dynamics of a cholera outbreak better so that we can greatly reduce the morbidity to be incurred in future outbreaks.

Much is known about the complex microbiological and ecological details that underlie a typical cholera outbreak (Anh et al., 2011; Durham et al., 1998; Glass et al., 1982; Kaper, Morris, & Levine, 1995; Longini et al., 2002; Merrell et al., 2002; Naficy et al., 1998; Nelson, Harris, Morris, Calderwood, & Camilli, 2009; Pollitzer, 1959; Qureshi et al., 2006; Reyburn et al., 2011; Shultz et al., 2009; World Health Organization, 2010b). At the same time, these outbreaks keep occurring in developing nations, where careful disease surveillance is not possible. Possibly due to the limited data available to modellers, it is difficult to justify the mathematical use of a complex mechanistically-driven model, because accurate and unique parameterizations for these models cannot be obtained. The modelling literature contains a spectrum of deterministic models from very simple to more complex structures (Andrews & Basu, 2011; Eisenberg, Robertson, & Tien, 2013; Hartley, Morris, Jr., & Smith, 2005; King, Ionides, Pascual, & Bouma, 2008; Mukandavire et al., 2011; Miller Neilan, Schaefer, Gaff, Fister, & Lenhart, 2010; Rinaldo et al., 2012; Tien & Earn, 2010; Tuite et al., 2011). We note that some researchers complement a relatively simple deterministic model with complex spatial considerations and migration.

Thus, due to our limitations in obtaining accurate data, there is some disconnect between our mechanistic understanding of how cholera spreads and the models we use to describe the spread. There is tremendous disagreement throughout both kinds of research, biological and mathematical, about what parameter values are realistic and what

modelling assumptions are important. We will explore some of these disagreements, and then introduce several candidate models for cholera.

Generally, cholera enters a human population through the ingestion of food or water that has been contaminated by the bacterium *V. cholerae*. The vibrios colonize in the small intestine for 12–72 h, and those humans who experience symptomatic infections then shed an enormous number of vibrios through diarrhoea which can last 1–2 weeks. The infection leads to rapid death without hydration therapy, which is inexpensive and easily applied in all but the most under-developed settings (Nelson et al., 2009). Initial cases of cholera in a population can be caused by the chance ingestion of vibrios that lurk indefinitely in the environment of many affected regions, or that are imported from asymptomatic visitors, as may well be the case in Haiti with the Nepalese peacekeepers (Piarroux et al., 2011). After the vibrios enter the human population, there is a human-environmental amplification of the disease in regions without adequate sanitation, as vibrios that have been shed from humans in turn contaminate the water that is used for drinking and cleaning (Nelson et al., 2009). This amplification is additionally powered by a 700-fold increase in the virulence of the bacteria in its first five or so hours after leaving the human intestine (Merrell et al., 2002). It is thus compelling to consider a model for cholera that attempts to describe the human-environmental disease amplification, as well as one that addresses the importance of freshly shed bacteria in the spread of the disease.

As we mentioned earlier in this section, only some humans will have symptomatic infections, and the proportion of humans with severe infections depends on the bacterial biotype (Kaper et al., 1995) and perhaps other possible factors. It is known that the degree of susceptibility depends on genetic factors (Nelson et al., 2009) and probably on nutrition as well (Glass et al., 1989; Nelson et al., 2009). Additionally, it is believed that prior exposure to the disease causes short-term complete or partial immunity (Longini et al., 2002; Nelson et al., 2009), and Glass et al. (1982) in particular suggest that recovery from a symptomatic cholera infection should protect an individual from a subsequent *severe* infection. It is noted that in areas with endemic disease, there are a smaller number of symptomatic cases than in areas that are immunologically naïve (Nelson et al., 2009), and from this we might infer that in countries with frequent cholera outbreaks, residents may enjoy some partial immunity if they re-encounter the bacteria. Even if one had accurate data representing the fraction of infections that are asymptomatic, then the extent to which asymptomatic infections result from a physiological predisposition versus prior exposure to cholera would not be clear, except, of course, in an outbreak for which the existing populations have no prior exposure. This latter point makes data from Haiti, whose residents had not experienced cholera prior to 2010, especially important for a model study.

In addition to the lack of quantification available as to the cause of asymptomatic infections, additionally there is wide disagreement regarding the proportion and importance of asymptomatic infections in any given cholera outbreak. There are excellent studies that suggest the per cent of symptomatic individuals might be quite small (Centers for Disease Control, xxx; Glass et al., 1982; King et al., 2008), about 20% (Rinaldo et al., 2012; World Health Organization, 2010a), or possibly 50% (Nelson et al., 2009). Thus, the proportion of asymptomatic infections in an outbreak is not well understood, and additionally the contribution of asymptomatic individuals to disease amplification is a subject of dispute. Some researchers believe that the ‘silent shedders’ who are typically undocumented in data collected during an outbreak are driving the disease spread (King et al., 2008;

Table 1. Summary of all model variables used in one or more of the five models.

$S(t)$	Susceptible humans
$\hat{S}(t)$	Susceptible with partial immunity
$I(t)$	Cholera-infected humans
$I_S(t)$	Symptomatic infected humans
$I_A(t)$	Asymptomatic infected humans
$R(t)$	Recovered humans
$R_S(t)$	Humans recovered from symptomatic infection
$R_A(t)$	Humans recovered from asymptomatic infection
$N(t)$	Total population in all human compartments
$B(t)$	Concentration of cholera vibrios in environment
$B_H(t)$	Concentration of high-infectious vibrios in environment
$B_L(t)$	Concentration of low-infectious vibrios in environment

Rinaldo et al., 2012), while others think that these shedders are of less consequence (Nelson et al., 2009). For those with asymptomatic infections, does the lack of disease awareness cause an increased contribution to the environmental and human-to-human spread of the bacteria during the first one to two days of asymptomatic illness? Are those with symptomatic infections effectively quarantined due to the evident need to observe strict sanitation? Models may thus need to consider a range of bacterial contact rates for each type of individual (symptomatic and asymptomatic) in order to account for the full range of possible outcomes.

Another parameter which requires attention is the rate of bacterial non-viability. We have observed that a majority of cholera models assume that the bacteria will not be viable after 30 days in the environment; however, a recent commentary on cholera modelling suggests that in fact persistence in the water may last anywhere from 3 to 40 days, and that this change in assumption can cause an enormous difference in model predictions (Grad, Miller, & Lipsitch, 2012).

With the discussion of disagreements within the literature in hand, we now explore several model options.

3.2. Five models for cholera

3.2.1. Model 1: SIR

We begin by suggesting a simple SIR model, depicted in Figure 1. The model assumes that disease spread is density-dependent, and hence the transmission term includes division by the total population size N . The SIR model is given by the system

$$\begin{aligned}\frac{dS}{dt} &= bN - dS - \beta_I \frac{SI}{N} + \omega R \\ \frac{dI}{dt} &= -dI + \beta_I \frac{SI}{N} - \gamma I \\ \frac{dR}{dt} &= -dR + \gamma I - \omega R,\end{aligned}\tag{4}$$

where the meanings of the variables and parameters are given in Tables 1 and 2. We would claim that this model has little mechanistic merit in its description of cholera transmission, and we include it as a check of our model comparison technique. In Section 5, we will show that in some cases, the fit of this quite possibly inappropriate model to the data we have

Table 2. Model parameters and range assumptions used in at least one of the five models described in Sections 3.2.1–3.2.5. Section 3.3 provides some discussion about the values considered.

Parameters with fixed values		
b	Natural birth rate of humans	.000072
d	Natural death rate of humans	.000044
κ	Half-saturation constant of vibrios	10^6
κ_L	Half-saturation constant of low-infectious vibrios	10^6
κ_H	Half-saturation constant of hyperinfectious vibrios	$\frac{10^6}{700}$
Parameters chosen using Genetic Algorithms approach (Akman & Schaefer, 2015)		
h	Proportion of (symptomatic) infected hospitalized	0 – 1
ρ	Prob. of new infected from S to be asymptomatic	.001 – 1
β_B	Ingestion rate of vibrio from environment	.001 – 0.5
β_I	person-to-person infectivity	.001 – 0.5
β_{I_S}	infectivity from symptomatic infected contact	.001 – 0.5
β_{I_A}	infectivity from asymptomatic infected contact	.001 – 0.5
γ	Cholera recovery rate for infecteds	$\frac{1}{14}$ – $\frac{1}{7}$
γ_{I_S}	Cholera recovery rate for symptomatic infecteds	$\frac{1}{14}$ – $\frac{1}{7}$
γ_{I_A}	Cholera recovery rate for asymptomatic infecteds	$\frac{1}{2}$ – 1
ω	Rate of waning immunity from R to S	.00001 – $\frac{1}{190}$
ω_{R_A}	Rate of waning immunity from R_A to \hat{S}	.00001 – $\frac{1}{90}$
ω_{R_S}	Rate of waning immunity from R_S to \hat{S}	.00001 – $\frac{1}{360}$
$\omega_{\hat{S}}$	Rate of waning immunity from \hat{S} to S	.00001 – $\frac{1}{1000}$
η	Rate of infected contribution to environmental vibrio concentration	.0001 – 10
η_{I_A}	Rate of contribution by asymptomatic infecteds	.0001 – 10
η_{I_S}	Rate of contribution by symptomatic infecteds	.0001 – 10
δ	Death rate of vibrios from HI to non-HI state	$\frac{1}{40}$ – $\frac{1}{3}$
init	Initial proportion of κ in the environment	.001 – 1

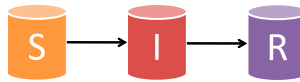


Figure 1. The possibly too-simple SIR model is included in our analysis as a control and is described in Section 3.2.1.

is good, particularly from the model selection viewpoint that favours models with few parameters.

3.2.2. Model 2: SIRB

The SIRB model is pictured in Figure 2 and consists of the differential equations

$$\begin{aligned}
 \frac{dS}{dt} &= bN - dS - \beta_B \frac{BS}{\kappa + B} - \beta_I \frac{SI}{N} + \omega R \\
 \frac{dI}{dt} &= -dI + \beta_B \frac{BS}{\kappa + B} + \beta_I \frac{SI}{N} - \gamma I \\
 \frac{dR}{dt} &= -dR + \gamma I - \omega R \\
 \frac{dB}{dt} &= \eta I - \delta B,
 \end{aligned}
 \tag{5}$$

with variable and parameter meanings given in Tables 1 and 2. This model includes an environmental class for bacteria, which can be interpreted as the concentration of infectious cholera bacteria in the drinking water. Varying formulations of this simple model, with

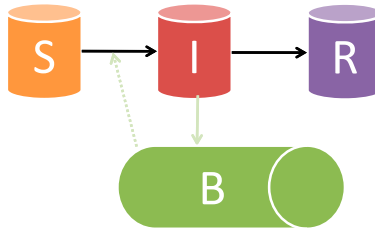


Figure 2. The SIRB model is used by many researchers to model the spread of cholera, though the precise mechanisms included, in particular two pathways of infection and the inclusion of waning immunity, differ within the references. See Section 3.2.2 for the model equations.

or without waning immunity or person-to-person infectivity as included in Equations (5), have been used for decades (Capasso & Serio, 1978; Codeço, 2001). Recently, the SIRB model was applied to cholera quite effectively, especially when used in combination with spatial considerations, to model the outbreaks in Zimbabwe (Mukandavire et al., 2011) and Haiti (Bertuzzo, Casagrandi, Gatto, Rodriguez-Iturbe, & Rinaldo, 2010; Rinaldo et al., 2012; Tuite et al., 2011). In other model comparison efforts for cholera, a variation of the SIRB formulation is often selected as the most appropriate to model the available data.

3.2.3. Model 3: SIIRBB

The SIRB model considered in Section 3.2.2 ignores the hyperinfectious state for bacteria as well as the role of asymptomatic infected individuals that we discussed in Section 3.1. First published in 2010 (Miller Neilan et al., 2010), the SIIRBB model

$$\begin{aligned}
 \frac{dS}{dt} &= bN - dS - \beta_{BL} \frac{B_L S}{\kappa_L + B_L} - \beta_{BH} \frac{B_H S}{\kappa_H + B_H} + \omega R \\
 \frac{dI_S}{dt} &= -dI_S + (1-p) \left[\beta_{BL} \frac{B_L S}{\kappa_L + B_L} + \beta_{BH} \frac{B_H S}{\kappa_H + B_H} \right] - \gamma_{I_S} I_S \\
 \frac{dI_A}{dt} &= -dI_A + p \left[\beta_{BL} \frac{B_L S}{\kappa_L + B_L} + \beta_{BH} \frac{B_H S}{\kappa_H + B_H} \right] - \gamma_{I_A} I_A \\
 \frac{dR}{dt} &= -dR + \gamma_{I_S} I_S + \gamma_{I_A} I_A - \omega R \\
 \frac{dB_H}{dt} &= \eta_{I_A} I_A + \eta_{I_S} I_S - \chi B_H \\
 \frac{dB_L}{dt} &= \chi B_H - \delta B_L
 \end{aligned} \tag{6}$$

adds a hyperinfectious bacterial class as well as an asymptomatic infected human class to the SIRB model. The SIIRBB structure is pictured in Figure 3. This model expanded a 2006 model that first introduced the hyperinfectious class to cholera modelling (Hartley et al., 2005), and similarly did not include the person-to-person infectivity term that we see in the SIRB model in Section 3.2.2. A model similar to the one in Equations (6) has also been used Andrews and Basu (2011) to consider control measures for the outbreak in Haiti.

3.2.4. Model 4: SIIRB

Shortly after the publication of the first cholera model to include hyperinfectivity (Hartley et al., 2005), a short note was published (Pascual, Koelle, & Dobson, 2006),

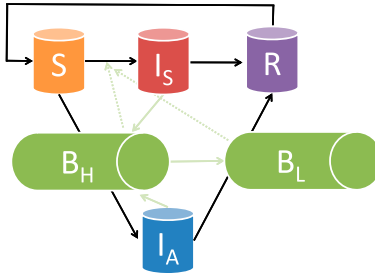


Figure 3. The diagram is an illustration of the SIIRBB model, which includes a hyperinfectious class of bacteria and also differentiates between the dynamics for symptomatic and asymptomatic infected individuals. See a full description in Section 3.2.3.

suggesting that the inclusion of a hyperinfectious class of bacteria is only a different way to show a faster mode of transmission, and that epidemic data can therefore be modelled equally well by inclusion of person-to-person transmission rather than through a hyperinfectious class of bacteria. In the context of the SIIRBB model of the previous section, we consider replacing the hyperinfectious bacterial class with person-to-person transmission. The resulting model is

$$\begin{aligned}
 \frac{dS}{dt} &= bN - dS - \beta_{I_S} \frac{I_S S}{N} - \beta_{I_A} \frac{I_A S}{N} - \beta_B \frac{BS}{\kappa + B} + \omega R \\
 \frac{dI_S}{dt} &= -dI_S + (1-p) \left[\beta_{I_S} \frac{I_S S}{N} + \beta_{I_A} \frac{I_A S}{N} + \beta_B \frac{BS}{\kappa + B} \right] - \gamma_{I_S} I_S \\
 \frac{dI_A}{dt} &= -dI_A + p \left[\beta_{I_S} \frac{I_S S}{N} + \beta_{I_A} \frac{I_A S}{N} + \beta_B \frac{BS}{\kappa + B} \right] - \gamma_{I_A} I_A \\
 \frac{dR}{dt} &= -dR + \gamma_{I_S} I_S + \gamma_{I_A} I_A - \omega R \\
 \frac{dB}{dt} &= \eta_{I_A} I_A + \eta_{I_S} I_S - \delta B,
 \end{aligned} \tag{7}$$

where again variable and parameter values are provided in Tables 1 and 2. We add the above SIIRB model to our comparisons to check just this idea: can the data be described equally well by a model that includes person-to-person transmission in lieu of a hyperinfectious class of bacteria? Our model comparison results, if we jump ahead a little to Section 5, suggest ‘not quite’. The model is diagrammed in Figure 4.

3.2.5. Model 5: SSIIRRB

Finally, we consider a so-called SSIIRRB model that is expanded to allow the possibility that the capacity to obtain partial immunity may drive cholera dynamics. A similar model appears in an age-structured partial differential equations model for cholera with several

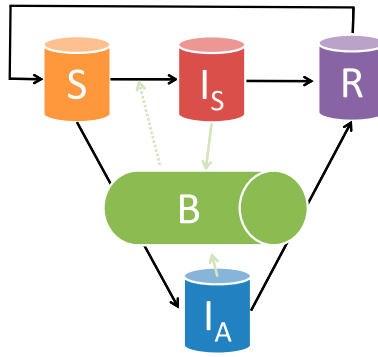


Figure 4. Above is an illustration of the SIIRB model that is described in Section 3.2.4. Note: This model is motivated by a desire to see if the hyperinfectious class of the SIIRBB model described in Section 3.2.3 can be replaced by person-to-person transmission.

collaborators (Fister et al., 2016). The model equations are given by

$$\begin{aligned}
 \frac{dS}{dt} &= bN - dS - \beta_B \frac{BS}{\kappa + B} - \beta_{I_S} \frac{I_S S}{N} - \beta_{I_A} \frac{I_A S}{N} + \omega_{\hat{S}} \hat{S} \\
 \frac{d\hat{S}}{dt} &= -d\hat{S} - \beta_B \frac{B\hat{S}}{\kappa + B} - \beta_{I_S} \frac{I_S \hat{S}}{N} - \beta_{I_A} \frac{I_A \hat{S}}{N} \\
 &\quad + \omega_{R_A} R_A + \omega_{R_S} R_S - \omega_{\hat{S}} \hat{S} \\
 \frac{dI_S}{dt} &= -dI_S + (1-p) \left[\beta_B \frac{BS}{\kappa + B} + \beta_{I_S} \frac{I_S S}{N} + \beta_{I_A} \frac{I_A S}{N} \right] - \gamma_{I_S} I_S \\
 \frac{dI_A}{dt} &= -dI_A + \beta_B \frac{B\hat{S}}{\kappa + B} + \beta_{I_S} \frac{I_S \hat{S}}{N} + \beta_{I_A} \frac{I_A \hat{S}}{N} \\
 &\quad + p \left[\beta_B \frac{BS}{\kappa + B} + \beta_{I_S} \frac{I_S S}{N} + \beta_{I_A} \frac{I_A S}{N} \right] - \gamma_{I_A} I_A \\
 \frac{dR_S}{dt} &= -dR_S + \gamma_{I_S} I_S - \omega_{R_S} R_S \\
 \frac{dR_A}{dt} &= -dR_A + \gamma_{I_A} I_A - \omega_{R_A} R_A \\
 \frac{dB}{dt} &= \eta_{I_A} I_A + \eta_{I_S} I_S - \delta B
 \end{aligned} \tag{8}$$

with variable and parameter values given in Tables 1 and 2. The inclusion of a partially immune class may lead to interesting results for long-term dynamics in Haiti, where initially the class with partial immunity would have been empty since the Haitian population had no prior exposure to cholera. Although this more complex model may be justifiable from a mechanistic point of view, and is certainly relevant in age-based considerations, the number of parameters needed for this model is extensive. Are there enough data to justify the use of this model? See Figure 5 for graphics describing the model.

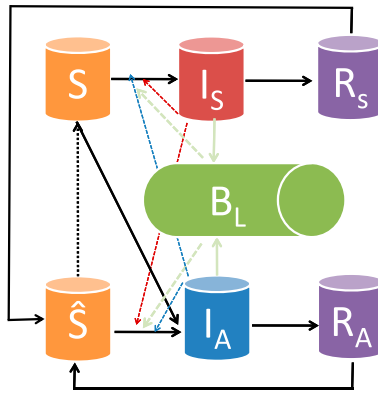


Figure 5. The SSIIRRB model is described in Section 3.2.5.

Note: The model does not include a hyperinfectious class and, thus, is an extension of the SIIRB model described in Section 3.2.4.

3.3. Further parameter discussion

One reason that model choice is tricky for the study of cholera is the disagreement surrounding parameter choice. We will highlight some of the discordance within the literature for the parameter values, and use the parameter discussions from the literature to choose biologically-reasonable ranges (see Table 2) for our parameter search. As we will discuss in more detail in Sections 3.4 and 5 following, we intend for the models to describe surveillance data from Haiti which provides the number of hospitalizations from the beginning of the cholera epidemic there. With data that only suggests the number of reported infections, it is quite difficult to discover the transmission route (Eisenberg et al., 2013). Thus, it is not only difficult to recover specific parameter values related to person-to-person or bacterial infectivity (the β_* values in Table 2); as we parametrize the five models considered using the hospitalization data, it is also unlikely that we could determine the relative importance of β_{I_A} versus β_{I_S} . Similarly, as discussed earlier, there is disagreement between the CDC and WHO websites (Centers for Disease Control, xxxx; World Health Organization, 2010a), a *Nature Reviews Microbiology* overview (Nelson et al., 2009), and several modelling articles (King et al., 2008; Rinaldo et al., 2012) regarding what proportion of infections may be asymptomatic, which suggests that we should allow a wide range of values for our parameter p . Recall also that for the SSIIRRB model (Section 3.2.5), which allows for partial immunity, the parameter p assumes a slightly different meaning than in the SIIRBB and SIIRB models (Sections 3.2.3 and 3.2.4), in which all susceptible individuals are considered equivalent.

Finally, we observe that bacterial parameter values may be quite difficult to determine. In models that include the human-environment amplification, the contributions by humans to the bacterial load in the environment (η_*) cannot be estimated outside of numerical model simulations, because of the impossible-to-classify volume and location of water in the environment, as well as the reality that much of the spread of the disease is due to bacterial contamination of household objects (Nelson et al., 2009).

3.4. Initial conditions for all models

The model parameters and assumed ranges for the five models described here are given in Table 2. Each model is parametrized using surveillance data from the Haitian Ministries of Public Health for approximately the first fifteen months following the 2010 earthquake (Republique d’Haiti, xxxx). We have considered this data separately for each of the ten Haitian departments as well as its capital city, as we will elaborate in Section 5. For each of the five models, the initial conditions depend on the demographics of the specific region considered.

For all models and all Haitian data-sets, we considered those living in poverty to be the only individuals who would be susceptible to cholera. We used poverty (Verner, 2007) and population (Wikipedia, xxxx) data to obtain initial values for the number of susceptible humans in each department. Where relevant, we assumed that initially all susceptible humans would carry no partial immunity (since Haiti had not been exposed to cholera prior to the epidemic in this study). For each model/data-set combination, we assumed that Day Zero was the first day with hospitalized infected humans as recorded by the Haitian Ministries of Public Health (Republique d’Haiti, xxxx). In all models, we assumed that only a proportion h of infected humans are hospitalized (see Table 2). Thus,

$$h \cdot I(0) = I^* \equiv \text{number of hospitalizations on Day Zero.}$$

We assumed for the models with symptomatic and asymptomatic populations that only symptomatic individuals would be hospitalized, so that

$$I_S(0) = \frac{I^*}{h}.$$

If $I(0)$ were to represent the true number of infected inclusive of asymptomatic infections, then $I_S(0) = pI(0)$, and $I_A(0) = (1 - p)I(0)$. It follows that

$$I_A(0) = \frac{1 - p}{p} I_S(0).$$

The precise numerical values chosen for the initial conditions above are summarized in Table 3.

We assumed that there were no recovered individuals at the start of the epidemic, and (where relevant) no hyperinfectious bacteria on Day Zero. We did assume an environmental concentration of low-infectious bacteria that is some multiple of the Michaelis constant κ (or κ_L for the models that consider hyperinfectious bacteria). Depending on the chosen model, this amounts to one of the following (see Table 2):

$$B(0) = \text{init} \cdot \kappa, \quad B_L(0) = \text{init} \cdot \kappa_L.$$

4. Prior results in model selection applied to cholera

As we will highlight in the following discussion, there have been a number of interesting studies considering model selection for cholera in recent years. No two research groups seem to be considering model comparison through the same lens.

Table 3. Initial assumptions.

Department/city	Population size	Poverty rate	Init. Inf. in data
Artibonite	1,168,000	.59	460
Centre	564,200	.56	6
Grand'Anse	337,516	.63	7
Nippes	263,000	.63	5
Nord	972,200	.63	1
Nord-Est	300,500	.80	2
Nord-Ouest	445,080	.65	16
Ouest	2,943,200	.29	3
Sud-Est	518,200	.57	2
Sud	745,000	.63	1
Port au Prince	942,000	.57	7

See Section 3.4 for references. Note that initial infections are estimated from data given in graphical form.

4.1. Model comparison applied to cholera models using data from Haiti

A thoughtful look at multiple models for cholera is shared by [Rinaldo et al. \(2012\)](#). The authors consider four structural models to describe the spread of cholera through Haiti, which include similar terms to those in our study but also include detailed geographic information and rainfall intensity as well as human mobility patterns. The models consider the data throughout Haiti as a whole, connected by multiple systems of differential equations that are linked through migration and that are parametrized with local information. This is truly an impressive undertaking in its realistic parametrization of several key details.

On the other hand, the model comparisons do not include all structures that interest us, and additionally, some assumptions of constant parameter values given in their Table S2 are based on assumptions used by or derived within previous cholera modelling articles. As an example, the length of immunity after acquiring cholera is based on a simulation outcome in one researcher's work ([Koelle, Rodó, Pascual, Yunus, & Mostafa, 2005](#)). The length of immunity was based on data from Matlab, Bangladesh. We might hypothesize that this population with partial immunity might be quite different from that in Haiti, since Haitians had not received prior exposure to cholera. Another example of different parameter assumptions is in the consistent use of the contact rate with infected waters to be $\beta_B = 1$. Our previous discussion in Section 3.3 explains our concern with this assumption.

Thus, our work differs from [Rinaldo et al. \(2012\)](#) in that we assume that fewer parameters are known, and in that we desire to explore a greater array of mechanistic possibilities in various models. We are limited by our resources to consider only individual departments in contrast to a model with spatially-connected departments ([Rinaldo et al., 2012](#)). At the same time, we find the question of model selection when the departments are separated interesting. If the data from different departments show support for differing models, then how might one choose an overall model for the region? This question is of interest to other researchers who lack the resources to consider fine GIS-based detail, and hence, a comparison between the viewpoints is valuable.

The form of model comparison in the work by [Rinaldo et al. \(2012\)](#) is in agreement with the form that we discussed in Section 1 and provided in Equations (3). For this article, the authors used one data point for each of 49 weeks spread across 10 Haitian departments and thus considered $n = 490$ data points. Referring to Equation (3), this [Rinaldo et al.](#) also use $K = p + 1$ parameter values; that is, the number of model parameters p plus a

residual variance parameter. The references cited for this approach to model comparison are a 1998 version of [Burnham and Anderson \(2002\)](#) that is discussed in the Introduction, and an article by [Corani and Gatto \(2007\)](#), who cite the work of [Burnham and Anderson \(2002\)](#) and extend the idea to demographic ecological (regression) models.

4.2. Model selection applied to cholera models in Bengal

[King et al. \(2008\)](#) engage in a highly stochastic study comparing compartmental models for cholera against data from Bengal that is rich in time and space. They draw likelihood-based inferences and refer the reader to certain textbooks ([Casella & Berger, 2002](#); [Cox & Barndorff-Nielsen, 1994](#); [Rice, 2006](#)) for a background for inference. They provide a detailed algorithm explaining how they find the maximum likelihood via iterated filtering. For model comparison, the authors refer to Akaike's article ([Akaike, 1974](#)) as well as the Burnham and Anderson text ([Burnham & Anderson, 2002](#)) to explain their application of AIC, as well as AICc which is intended for small sample sizes and is introduced in the following section ([Burnham & Anderson, 2002](#); [Richards, 2005](#)). The Akaike article considers the comparison of autoregressive and moving average models, and since King et al. do not specifically write the form of AIC used it is not clear how the AIC references are applied.

4.3. Model selection applied to cholera data from Angola

[Eisenberg et al. \(2013\)](#) perform an interesting experiment of testing whether parameters can be identified using perfect data in an SIRB model for cholera (this discussion replaces the article's notation with our own for the reader's convenience). The authors then ask, using a small data-set from a cholera outbreak in Angola, whether it is possible to determine the mode of transmission for the disease (person-to-person, water-borne, or both). For this task, they apply a corrected AIC (AIC_C) method for small sample sizes, and they cite [Hurvich and Tsai \(1991\)](#). Hurvich and Tsai explain that the AIC_C method is for a 'normal linear regression, or autoregressive, model with p regression or autoregressive, parameters', and is given by

$$\begin{aligned} \text{AIC}_C &= n \log (2\pi \hat{\sigma}^2) + n \frac{1 + \frac{p}{n}}{1 - \frac{p+2}{n}} \\ &= n \log (2\pi \hat{\sigma}^2) + n \frac{n+p}{n-p-2}, \end{aligned}$$

where σ^2 is given in terms of the derivation for variance in a least squares estimate for linear regression. We note that as described in Section 1, the AIC_C would be defined by

$$\begin{aligned} \text{AIC}_C &= \text{AIC} + \frac{2K(K+1)}{n-K-1} \\ &= n \ln \left(\frac{\text{RSS}}{n} \right) + 2K + \frac{2K(K+1)}{n-K-1} \\ &= n \ln \left(\frac{\text{RSS}}{n} \right) + \frac{2Kn}{n-K-1}, \end{aligned}$$

where K is the number of model parameters plus one for the variance of the error in the fit and n is the number of observations from the data in Angola. The authors do not share the details of the model comparison, and we do not know the precise form of AIC_C that they use.

4.4. A new approach for model comparison involving seasonal epidemics

A different approach for model comparison can be derived if one assumes a certain probability distribution for the number of hospitalized cases in an epidemic. [Ponciano and Capistrán \(2011\)](#) show in detail how the problem of model selection and parameter estimation can be approached if one assumes that the number of new hospitalizations follows a Poisson distribution. With this viewpoint, the AIC can be determined based on a likelihood function that parameter choices allow the new infections to follow a scaled Poisson distribution. We appreciated this derivation and viewpoint, but also felt that in our case, where the epidemic data shows epidemic waves, a Poisson distribution approach would not be a good basis for parameter fitting or model comparison.

The authors also explain a derivation for AIC in the case of parameterizing the model using weather data. In either case (the Poisson-derived likelihood for a compartment-based model for the number of infections or a more statistically-motivated model for the weather), the authors use the formula for AIC given by

$$AIC = -2 \ln \hat{\mathcal{L}} + 2p, \quad (9)$$

where $\hat{\mathcal{L}}$ is the maximum likelihood estimate for the given model and p is the number of parameters, and as we read the article p is not increased by the parameter for the variance of the fit errors. The likelihood function derived for the error in the weather data in the article is

$$\mathcal{L}(\theta_w) = \prod_{j=0}^q \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(w_j - \omega_{t_j})^2}{2\sigma^2} \right\}, \quad (10)$$

where w_j is the j th weather observation and ω_{t_j} is the weather model prediction, which they explain is equivalent to minimizing the expression

$$SSQ(\theta)_\omega = \sum_{j=0}^q (w_j - \omega_{t_j})^2.$$

Although the model for weather in this article is not a compartmental model, the equivalence of maximizing likelihood in Equation (10) to our goal of minimizing the RSS values caused us to consider the output of the form of AIC given in Equation (9) in our analysis in the following section.

The authors do not stop with an AIC exploration of models considered; rather, they state that one should also check the BIC values. They also state, in contrast to the discussion in Section 2.2, that if there is disagreement between AIC and BIC recommendations, then this is a signal that there is not sufficient data to choose the best model ([Ponciano & Capistrán, 2011](#)). We examine this claim for our case study in Section 5.

5. Methods and results

We applied several formulations of model comparison criteria to the five models for cholera that were described in Section 3. Thus, we sought the model that best described the data giving hospitalizations in each Haitian department and its capital approximately during the first fifteen months of the recent outbreak following the 2010 earthquake. This data is available in chart form from the Haitian Ministries for Public Health ([Haiti Institute of Statistics, 2011](#)), and we used the freely available software Plot Digitizer to convert that data to numeric form ([SourceForge, xxxx](#)). Hence, in addition to the expected surveillance errors given the setting, there were surely additional errors in representing the data numerically. We believe that these errors are typical, and that the data will have noise in any large-scale cholera study. As a result, if there were a ‘true’ model that could describe the surveillance data in such an epidemiological setting, then that ‘true’ model would not be a ‘true’ representation of the actual disease dynamics. This is simply unavoidable, and we assume that the model which best describes the noisy data would be the best model to describe the actual data if it were possible to discover that data.

The first step in performing model comparison is to choose the best parameter set for each of the five models listed in Section 3, so that the data is optimally described using each model. We chose the parameter sets independently for each of the ten Haitian departments and for the capital (eleven data-sets) for each of the five models. There are many approaches to parameter selection. In this work, we first found twenty best-fit parameter sets for each of the 55 data-set?-model combination using genetic algorithms (GA) ([Akman & Schaefer, 2015](#)). As explained in the reference, each single run is the result of the evolution over many generations of 5000 randomly chosen parameter sets from the biologically feasible parameter space. Due to the stochasticity of the parameter fitting method, we observed variation in the goodness of fit and in the selected parameter values through multiple GA runs, and, as we explain further below, this variation is the reason for considering twenty parameter sets for each combination considered. We also used Matlab’s `lsqcurvefit` ([MATLAB and optimization toolbox release, 2012a](#)) as a tool for parameter estimation; we note that other researchers are using a Markov Chain Monte Carlo (MCMC) approach which uses a random walk through the parameter space and a method for evaluating proposed moves to find optimal parametrization ([Rinaldo et al., 2012](#)). In a future article, we will explore the use of particle swarm optimization as an efficient and effective tool for parameter estimation.

Since the model’s fit to the data is a cornerstone of an assigned AIC or BIC value, we worried about the impact that stochasticity might play in model selection. For each of the $5 \times 11 = 55$ model–data-set combinations, we found the lowest RSS/ n value (where, again, n is the number of data observations for the given data-set). However, our work in exploring methods for parameter selection have convinced us that it is unlikely that differing parameter estimation approaches, or even repeated approaches, are likely to retrieve the same parameter values or goodness of fit (this was explored further in a previous article of ours ([Akman & Schaefer, 2015](#))). Thus, we observe that if we might get ‘lucky’ in finding a strongly performing data-set for model A, but ‘unlucky’ for model B. Should we be cautious about the limitations of *any* parameter fitting method for its ability in a single run to truly find the *best* parameter set? As a result of this concern, we considered two values in each model comparison effort – one for the minimum value observed over

the 20 parameter sets chosen for each of the 55 model–data-set combinations and another for the average value observed over the 20 parameter sets chosen. The former makes a comparison based on the best result obtained for each parameter set, and the latter makes a comparison based on the average over all parameter fitting efforts.

In the light of the prior results explored in Section 4, we studied several formulations for model comparison. We first considered the AIC formulation given in Equation (3) in Section 2, which is also the formulation suggested by [Burnham and Anderson \(2002\)](#) and by [Rinaldo et al. \(2012\)](#) (Section 4.1). However, to explore our concerns about the consistency and quality of parameter fitting efforts, we calculated this AIC value using both the minimum and average RSS values found for each model–data-set combination. Thus, for each of the 55 model–data-set combinations, we calculated

$$\begin{aligned} \text{AIC}_{1,\min} &= n \ln \left(\frac{\text{RSS}_{\min}}{n} \right) + 2(p + 1), \\ \text{AIC}_{1,\text{avg}} &= n \ln \left(\frac{\text{RSS}_{\text{avg}}}{n} \right) + 2(p + 1), \end{aligned}$$

where n is the number of data points for a given department and p is the number of parameters in a given model.

We also considered the AIC development used for weather that was described in Section 4.4 above, but with the use of epidemic data. We believe a likelihood formulation should provide consistent results, and thus as an illustration we also considered the likelihood function \mathcal{L} given in Equation (10) for each of the 55 model–data-set combinations, and calculated, following Equation (9), the values

$$\begin{aligned} \text{AIC}_{2,\min} &= -2 \ln (\mathcal{L}_{\min}) + 2p, \\ \text{AIC}_{2,\text{avg}} &= -2 \ln (\mathcal{L}_{\text{avg}}) + 2p, \end{aligned}$$

where n is the number of data points for a given department and p is the number of parameters in a given model. To be precise, for each model–data-set combination we calculated the RSS between our scaled data ($h \cdot I(j)$ or $h \cdot I_S(j)$ for $j = 0, 1, \dots, n - 1$) and the surveillance data ($I^*(j)$ for $j = 0, 1, \dots, n - 1$), and we defined

$$\sigma^2 = \frac{\text{RRS}}{n}.$$

Then we define the likelihood function as

$$\mathcal{L} = \prod_{j=0}^{n-1} \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(hI(j) - I^*(j))^2}{2\sigma^2} \right\}.$$

We now note that

$$\ln (\mathcal{L}) = \sum_{j=0}^{n-1} \ln \left(\frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(hI(j) - I^*(j))^2}{2\sigma^2} \right\} \right). \quad (11)$$

Finally, despite our comments in Section 4 above, we follow the recommendation of [Ponciano and Capistrán \(2011\)](#), Section 4.4 and we consider the measure for BIC in both the least squares and likelihood cases. This allows us to consider the perhaps controversial point ([Ponciano & Capistrán, 2011](#)) that if AIC and BIC do not agree, then further study is warranted. For the least squares case, we suggest the form given by [Hansen \(2007\)](#) modified for the framework that we described in Section 2 to be

$$\text{BIC} = n \ln \left(\frac{\text{RSS}_{\min}}{n} \right) + p \cdot \ln(n),$$

where n and p are the number of data points and parameters, respectively, considered for the given model. The likelihood formulation given for BIC is

$$\text{BIC} = -2 \ln(\mathcal{L}) + p \cdot \ln n,$$

where \mathcal{L} is the likelihood function, p is the number of parameters, and n is the number observations ([Burnham & Anderson, 2002](#)). Thus, we also consider the formulas

$$\begin{aligned} \text{BIC}_{1,\min} &= n \ln \left(\frac{\text{RSS}_{\min}}{n} \right) + p \ln n, \\ \text{BIC}_{1,\text{avg}} &= n \ln \left(\frac{\text{RSS}_{\text{avg}}}{n} \right) + p \ln n, \end{aligned}$$

as well as

$$\begin{aligned} \text{BIC}_{2,\min} &= -2 \ln(\mathcal{L}) + p \ln n, \\ \text{BIC}_{2,\text{avg}} &= -2 \ln(\mathcal{L}) + p \ln n, \end{aligned}$$

as we make our comparisons.

Thus, we have defined eight model comparison methods. For each of these, we are seeking the minimum AIC and BIC values, and we can only declare a clear victor if there are no close runners-up. Following [Burnham and Anderson \(2002\)](#), [Rinaldo et al. \(2012\)](#), we measure the difference between each model's AIC/BIC value and the minimum value; these values are recorded by columns in Figure 6 and labelled Δ . The standard approach to interpreting the differences Δ assumes criteria that we cannot meet (independent observations, large sample sizes and nested models) ([Burnham & Anderson, 2002](#); [Richards, 2005](#)); however, we presume, as did [Rinaldo et al., \(2012\)](#), that the guidelines of the levels of support for each model vary similarly to the levels suggested by [Burnham and Anderson \(2002\)](#). Thus, referring to our results in Figure 6, the model selected has $\Delta = 0$ (dark shading), and we presume that there is substantial to moderate support for a runner-up model when $0 < \Delta \leq 5$ (moderate shading). Finally, there is weak support for a runner-up model when $5 < \Delta \leq 10$ (light shading). (The criteria in [Burnham and Anderson \(2002\)](#) are slightly different, but do not include all possible differences.)

Consider first our 'inadequate' SIR model from Section 3.2.1, for which results are shown in the top-left block in Figure 6. We note that the AIC_1 and BIC_1 criteria never select this model, but in Nippes, which experienced far fewer infections than in other departments ([Date et al., 1994](#)), the likelihood formulations AIC_2 and BIC_2 select the SIR model. We do

Summary of Comparison Measure Differences

	Data Set	1) SIR Model								2) SIRB Model							
		$\Delta AIC(1,min)$	$\Delta AIC(1,avg)$	$\Delta AIC(2,min)$	$\Delta AIC(2,avg)$	$\Delta BIC(1,min)$	$\Delta BIC(1,avg)$	$\Delta BIC(2,min)$	$\Delta BIC(2,avg)$	$\Delta AIC(1,min)$	$\Delta AIC(1,avg)$	$\Delta AIC(2,min)$	$\Delta AIC(2,avg)$	$\Delta BIC(1,min)$	$\Delta BIC(1,avg)$	$\Delta BIC(2,min)$	$\Delta BIC(2,avg)$
1) SIR Model	1: Artibonite	260	260	260	260	232	231	231	231	55	55	55	55	43	43	43	43
	2: Centre	55	143	143	143	129	129	129	129	0	0	0	0	0	0	0	0
	3: Grand Anse	153	153	152	153	139	138	138	138	0	0	0	0	0	0	0	0
	4: Nippes	74	73	0	0	60	60	0	0	0	0	8	8	0	0	22	22
	5: Nord	564	563	563	563	547	547	546	547	0	0	0	0	0	0	0	0
	6: Nord-Est	82	79	78	79	62	62	61	62	3	0	0	0	0	0	0	0
	7: Nord-Ouest	106	101	100	101	76	75	74	75	15	10	9	10	0	0	0	0
	8: Ouest	185	185	185	185	171	171	171	171	0	0	0	0	0	0	0	0
	9: Sud-Est	113	113	112	113	98	98	98	98	0	0	0	0	0	0	0	0
	10: Sud	275	273	272	273	258	257	256	257	0	0	0	0	0	0	0	0
	11: Port au Prince	172	172	172	172	157	157	157	157	0	0	0	0	0	0	0	0
	Total	2704	2692	2060	2071	2660	2649	2019	2027	108	99	16	23	90	81	0	4
3) SIIRBB Model	1: Artibonite	0	0	0	0	0	0	0	0	56	56	56	56	60	60	62	60
	2: Centre	5	5	5	5	15	16	16	16	7	8	8	8	21	21	22	21
	3: Grand Anse	15	14	14	14	25	25	24	25	8	7	7	7	22	22	21	22
	4: Nippes	12	12	14	14	22	22	39	39	7	8	16	16	21	22	44	43
	5: Nord	20	20	21	20	32	33	34	33	15	21	22	21	31	37	38	37
	6: Nord-Est	14	15	22	15	23	28	34	28	0	1	4	1	13	17	20	17
	7: Nord-Ouest	3	0	0	0	1	3	3	3	12	9	9	9	14	16	16	16
	8: Ouest	6	6	6	6	16	17	16	17	11	11	11	11	25	25	25	25
	9: Sud-Est	7	8	8	8	18	19	19	19	6	7	9	7	21	22	24	22
	10: Sud	18	20	19	20	30	32	31	32	6	4	4	4	22	21	20	21
	11: Port au Prince	6	7	7	7	17	18	18	18	8	8	8	8	23	23	23	23
	Total	0	0	0	0	0	0	3	0	104	105	19	18	111	111	28	25
5) SSIIRBB Model	1: Artibonite	25	32	39	32	37	44	51	44								
	2: Centre	11	11	11	11	31	32	32	32								
	3: Grand Anse	71	74	76	74	93	96	97	96								
	4: Nippes	20	33	20	20	41	54	55	58								
	5: Nord	51	56	65	56	76	81	89	81								
	6: Nord-Est	143	151	160	151	165	175	184	175								
	7: Nord-Ouest	0	2	7	2	10	17	22	17								
	8: Ouest	11	12	13	12	33	34	35	34								
	9: Sud-Est	25	26	26	26	47	48	48	48								
	10: Sud	31	30	29	30	55	54	54	54								
	11: Port au Prince	13	13	14	13	35	35	36	35								
	Total	124	145	283	258	143	164	304	277								
4) SIRB Model	1: Artibonite	0	0	0	0	0	0	0	0	56	56	56	56	60	60	62	60
	2: Centre	5	5	5	5	15	16	16	16	7	8	8	8	21	21	22	21
	3: Grand Anse	15	14	14	14	25	25	24	25	8	7	7	7	22	22	21	22
	4: Nippes	12	12	14	14	22	22	39	39	7	8	16	16	21	22	44	43
	5: Nord	20	20	21	20	32	33	34	33	15	21	22	21	31	37	38	37
	6: Nord-Est	14	15	22	15	23	28	34	28	0	1	4	1	13	17	20	17
	7: Nord-Ouest	3	0	0	0	1	3	3	3	12	9	9	9	14	16	16	16
	8: Ouest	6	6	6	6	16	17	16	17	11	11	11	11	25	25	25	25
	9: Sud-Est	7	8	8	8	18	19	19	19	6	7	9	7	21	22	24	22
	10: Sud	18	20	19	20	30	32	31	32	6	4	4	4	22	21	20	21
	11: Port au Prince	6	7	7	7	17	18	18	18	8	8	8	8	23	23	23	23
	Total	0	0	0	0	0	0	3	0	104	105	19	18	111	111	28	25

Figure 6. The results of our model comparison formulation investigations are summarized in this chart and discussed in Section 5.

Notes: Each model is considered separately, with individual data-sets given as rows and model comparison technique given as columns. Dark shading represents the model is selected by the criteria in the corresponding column. Light to moderate shading indicates support for the runner-up model as an alternative to the selected model.

note that while the RSS and likelihood formulations are in disagreement, the AIC versus BIC values are consistent between the formulations.

We now consider the SIRB model described in Section 3.2.2 and view the results in the top-right block within Figure 6. We observe that the SIRB model is almost always selected, and this result is in agreement with most conclusions drawn in Section 4. There are four exceptions, and a small ‘blip’ (Nord-Est). The first exception is Artibonite, which was hit very early with cholera and whose cases in the period were considered dwarfed by those in other departments. In Artibonite, all selection criteria chose the SIIRBB model (see the next paragraph). Second, Nippes, which was discussed in the previous paragraph and had

very few infections, selected the SIR model. The third region which did not select the SIRB model is Nord-Ouest, and here only BIC formulations choose the SIRB model. While the number of hospitalized in Nord-Ouest does not stand out as did the number for Nippes, the particular shape of the surveillance data (Republique d’Haiti, xxxx) is rather unusual for Nord-Ouest in that there is a rather sharp, brief peak in hospitalizations several months into the epidemic. Finally, while the individual data-sets generally justify the choice of the SIRB model, when the data are considered in total, most formulations do not select this model. The ‘Total’ rows in Figure 6 shows the result when the RSS values or the terms in Equation (11) are summed over all points in all data-sets, and n is taken to be the total points in all data-sets (similar to Rinaldo et al. (2012) but without their complex geographical setup). Thus, there is a disconnect between our intuitive idea that ‘clearly the SIRB model is the appropriate model to select given the criteria’ and the model that is actually selected when the data are considered as a whole. Where Rinaldo et al., considering a different subset of models, connected the Haitian departments rather than consider them separately, they did select an SIRB formulation (Rinaldo et al., 2012). The question that remains in this study is, ‘When the data are considered as a whole, is a complex model truly able to describe the data better, or is our method of considering the data as a whole allowing too much weight for a ‘deviant’ data set?’

The SIIRBB model, described in Section 3.2.3, has results pictured in the middle-left block of Figure 6. Several Haitian departments’ data were best described by the SIIRBB model. In particular, the hard-hit Artibonite selects this model across all criteria. Three of the four AIC criteria selected the ‘spiky’ data from Nord-Ouest, and the Δ values suggested that the SIIRBB model has support from all criteria. With the exception of one BIC criteria, the SIIRBB model is selected as the best model to describe the aggregate data. However, outside of these departments, there is weak or, more likely, no support for this model. We note that where there is a difference between the AIC and BIC choices, the Δ values generally suggest that there is evidence for choice between the two differing models.

We included the SIIRB model to check the hypothesis that the variation of the SIIRBB model could describe the data equally well (see Section 3.2.4). Results are summarized in the right-centre block of Figure 6. It was surprising to us that the SIIRB showed consistently weaker performance than SIIRBB in all but isolated cases. On the other hand, for some departments, and particularly those with fewer hospitalizations (Date et al., 1994), we see that the choice of person-to-person infectivity over the use of a hyperinfectious bacterial class has some merit. Our multiple comparison approaches suggest that the two models cannot be interchanged so easily, thus suggesting that models which include a hyperinfectious class of bacteria may be more appropriate than those substituting person-to-person infectivity in regions with high morbidity from cholera.

Finally, the SSIIRRB model – see Section 3.2.5 and the lower-left block of Figure 6 – only enjoyed weak support in a single department. We wonder if the SSIIRRB model would perform more strongly in a longer timeframe. The complexity in the model is intended to uncover long-time dynamics. We did not check this hypothesis. Certainly for the data-sets considered, the addition of multiple unknown parameter values would cause this model to perform poorly in for the comparison metrics, but in reality, the complicated SSIIRRB rarely outperformed other models in describing the data-sets even without penalty from additional parameters.

6. Conclusions and remarks

The initial motivation for this study was to explore the settings in which mechanistically-complex models might be needed or justified to explain the dynamics of a cholera outbreak. We find that within the majority of the individual Haitian departments, for the time-frame of about 15 months, the simple SIRB model enjoys the most support. However, in describing areas that experience larger numbers of cholera infections, or in describing the region as a whole, there is evidence for a model that includes the complexities of asymptomatic infections and a hyperinfectious bacterial state. Finally, we observe that the modelling choices of a hyperinfectious bacterial state versus person-to-person transmission are not interchangeable.

Modelling is an art. Comparison methods select between models by balancing the complexity and performance of each model; however, the more subjective step of choosing appropriate models and parameter spaces cannot be under-valued in this process. The decisions of how many data-sets and which time steps should be used to inform the model selection process are critical to the outcome of a model selection process, while choice across correct formulations of the comparison approaches is less important.

In our study, we focused on the commonly used model assessment metrics RSS and likelihood AIC, both of which yielded similar results, and for the most part these comparison metrics were consistent whether we used the minimum or average RSS values for the data-set in question. Further, we observe that while some authors discount the value of BIC in this setting (Burnham & Anderson, 2002), the BIC results do seem similar to the AIC results (Ponciano & Capistrán, 2011), and where there were differences there were other disagreements that also suggested that more information would be helpful in selecting an appropriate model. Finally, we believe that there are hurdles in interpreting composite model recommendations obtained by combining multiple data-sets. In particular, we observe that more complex models seem to enjoy support as the number of observed infections (not the number of data points) increases.

In conclusion, we would urge that increased transparency about the parameter selection process and the particular format of AIC/BIC chosen by modellers would be a service to our community, with the understanding that often these longer discussions may need to be provided in supplementary materials so that they do not distract from the researchers' immediate goals.

Acknowledgements

We thank Schaefer's prior co-authors for their ideas which spurred the model comparisons herein, especially Renee Fister, Holly Gaff and Suzanne Lenhart. We are grateful for the support provided by the IBA. We additionally thank the reviewers for their helpful suggestions and references.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716–723.

- Akman, O., & Schaefer, E. (2015). An evolutionary computing approach for parameter estimation investigation of a model for cholera. *Journal of Biological Dynamics*, 9, 147–158.
- Andrews, J. R., & Basu, S. (2011). Transmission dynamics and control of cholera in Haiti: An epidemic model. *The Lancet*, 377, 1248–1255.
- Anh, D. D., Lopez, A. L., Thiem, V. D., Grahek, S. L., Duong, T. N., Park, J. K., ... Clemens, J. D. (2011). Use of oral cholera vaccines in an outbreak in vietnam: A case control study. *PLoS Neglected Tropical Diseases*, 5, e1006.
- Bertuzzo, E., Casagrandi, R., Gatto, M., Rodriguez-Iturbe, I., & Rinaldo, A. (2010). On spatially explicit models of cholera epidemics. *Journal of the Royal Society Interface*, 7, 321–333.
- Bhansali, R. J., & Downham, D. Y. (1977). Some properties of the order of an autoregressive model selected by a generalization of Akaike's EPF criterion. *Biometrika*, 64, 547–551.
- Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach*. New York, NY: Springer Science & Business Media.
- Capasso, V., & Serio, G. (1978). A generalization of the Kermack-Mckendrick deterministic epidemic model. *Mathematical Biosciences*, 42, 43–61.
- Casella, G., & Berger, R. L. (2002). *Statistical inference*, (Vol. 2). Pacific Grove, CA: Duxbury.
- Centers for Disease Control. (xxxx) *Cholera general information*. Retrieved August 10, 2015, from <http://www.cdc.gov/cholera/general/>
- Codeço, C. T. (2001). Endemic and epidemic dynamics of cholera: the role of the aquatic reservoir. *BMC Infectious Diseases*, 1 (1), 1–14.
- Corani, G., & Gatto, M. (2007). Structural risk minimization: A robust method for density-dependence detection and model selection. *Ecography*, 30, 400–416.
- Cox, D., & Barndorff-Nielsen, O. (1994). *Inference and asymptotics* (Vol. 52). Boca Raton, FL: CRC Press.
- Date, K. A., Vicari, A., Hyde, T. B., Mintz, E. D., Danovaro-Holliday, M. C., .. Dietz, V. 2010-2011. *Considerations for oral cholera vaccine use during outbreak after earthquake in Haiti*. Retrieved September 12, 2015, from <http://wwwnc.cdc.gov/eid/article/17/11/11-0822-f1>
- Durham, L. K., Longini, I. M., Halloran, M. E., Clemens, J. D., Azhar, N., & Rao, M. (1998). Estimation of vaccine efficacy in the presence of waning: Application to cholera vaccines. *American Journal of Epidemiology*, 147, 948–959.
- Eisenberg, M. C., Robertson, S. L., & Tien, J. H. (2013). Identifiability and estimation of multiple transmission pathways in cholera and waterborne disease. *Journal of Theoretical Biology*, 324, 84–102.
- Fister, K., Gaff, H., Numfor, E., Schaefer, E., Wang, J., & Lenart, S. (2016). Optimal control of vaccination in an age-structured cholera model. In C. Gerardo, & J. M. Hyman, (Eds.), *Mathematical modeling for emerging and re-emerging infectious diseases*. Switzerland: Springer.
- Forster, M. R. (1999). Model selection in science: The problem of language variance. *The British Journal for the Philosophy of Science*, 50, 83–102.
- Glass, R. I., Becker, S., Huq, M. I., Stoll, B. J., Khan, M. U., Merson, M. H., ... Black, R. E. (1982, December). Endemic cholera in rural bangladesh, 1966–1980. *American Journal of Epidemiology*, 116, 959–970.
- Glass, R. I., Svennerholm, A. M., Stoll, B. J., Khan, M. R., Huda, S., Huq, M. I., & Holmgren, J. (1989, February). Effects of undernutrition on infection with vibrio cholerae o1 and on response to oral cholera vaccine. *The Pediatric Infectious Disease Journal*, 8, 105–109.
- Grad, Y. H., Miller, J. C., & Lipsitch, M. (2012). Cholera modeling: Challenges to quantitative analysis and predicting the impact of interventions. *Epidemiology (Cambridge, Mass.)*, 23, 523–530.
- Haiti Institute of Statistics. 2011, June 7. *Haiti MSPP daily case reports*. Retrieved June 17, 2011, from <http://www.mspp.gouv.ht/site/download.php?file=downloads/Rapport/journalier/0MSPP/du/05/juin/2011.pdf>
- Hannan, E. J. (1986). Remembrance of things past. In *The craft of probabilistic modelling* (pp. 31–42). Cambridge, MA: Springer.
- Hansen, B. E. (2007). Least squares model averaging. *Econometrica*, 75, 1175–1189.
- Hartley, D. M., Morris, J. G., Jr., & Smith, D. L. (2005, December). Hyperinfectivity: A critical element in the ability of *v. cholerae* to cause epidemics? *PLoS Medicine*, 3, e7.

- Hurvich, C. M., & Tsai, C.-L. (1991). Bias of the corrected AIC criterion for underfitted regression and time series models. *Biometrika*, 78, 499–509.
- Kaper, Jr, J. B., Glenn Morris, J., & Levine, M. M. (1995, January). Cholera. *Clinical Microbiology Reviews*, 8, 48–86.
- King, A. A., Ionides, E. L., Pascual, M., & Bouma, M. J. (2008, August 4). Inapparent infections and cholera dynamics. *Nature*, 454, 877–880.
- Koelle, K., Rodó, X., Pascual, M., Yunus, M., & Mostafa, G. (2005). Refractory periods and climate forcing in cholera dynamics. *Nature*, 436, 696–700.
- Longini, I. M., Yunus, M., Zaman, K., Siddique, A., Sack, R. B., & Nizam, A. (2002). Epidemic and endemic cholera trends over a 33-year period in Bangladesh. *Journal of Infectious Diseases*, 186, 246–251.
- MATLAB and optimization toolbox release (2012). Natick, MA: The MathWorks, Inc.
- Merrell, D., Butler, S., Qadri, F., Dolganov, N., Alam, A., Cohen, M., ... Camilli, A. (2002, June). Host-induced epidemic spread of the cholera bacterium. *Nature*, 417, 642–645.
- Miller Neilan, R. L., Schaefer, E., Gaff, H., Fister, K. R., & Lenhart, S. (2010, November). Modeling optimal intervention strategies for cholera. *Bulletin of Mathematical Biology*, 72, 2004–2018.
- Mukandavire, Z., Liao, S., Wang, J., Gaff, H., Smith, D. L., & Morris, J. G. (2011). Estimating the reproductive numbers for the 2008–2009 cholera outbreaks in Zimbabwe. *Proceedings of the National Academy of Sciences*, 108, 8767–8772.
- Naficy, A., Rao, M. R., Paquet, C., Antona, D., Sorkin, A., & Clemens, J. D. (1998). Treatment and vaccination strategies to control cholera in sub-Saharan refugee settings. *JAMA: The Journal of the American Medical Association*, 279, 521–525.
- Nelson, E. J., Harris, J. B., Morris, J. G., Calderwood, S. B., & Camilli, A. (2009, October). Cholera transmission: The host, pathogen and bacteriophage dynamic. *Nature Reviews Microbiology*, 7, 693–702.
- Pascual, M., Koelle, K., & Dobson, A. P. (2006, June). Hyperinfectivity in cholera: A new mechanism for an old epidemiological model? *PLoS Medicine*, 3, e280.
- Piarroux, R., Barraï, R., Faucher, B., Haus, R., Piarroux, M., Gaudart, J., ... Raoult, D. (2011). Understanding the cholera epidemic, Haiti. *Emerging Infectious Diseases*, 17, 1–8.
- Pollitzer, R. (1959). *Cholera, Monograph series, no. 43*. Geneva: World Health Organization. (pp. 1019).
- Ponciano, J. M., & Capistrán, M. A. (2011). First principles modeling of nonlinear incidence rates in seasonal epidemics. *PLoS Computational Biology*, 7, e1001079.
- Qureshi, K., Mølbak, K., Sandström, A., Kofoed, P.-E., Rodrigues, A., Dias, F., ... Sevnnerholm, A.-M. (2006). Breast milk reduces the risk of illness in children of mothers with cholera: Observations from an epidemic of cholera in Guinea-Bissau. *The Pediatric Infectious Disease Journal*, 25, 1163–1166.
- Republique d’Haiti [Republic of Haiti]. *Republique d’Haiti ministère de la santé publique et de la population: Rapports journaliers du mspp sur l’évolution du choléra en Haiti* [Republic of Haiti Ministry of Public Health and Population : Reports daily MSPP on the evolution of cholera in Haiti]. Retrieved June 9, 2014, from <http://mspp.gouv.ht/site/downloads/>
- Reyburn, R., Deen, J. L., Grais, R. F., Bhattacharya, S. K., Sur, D., Lopez, A. L., ... von Seidlein, L. (2011). The case for reactive mass oral cholera vaccinations. *PLoS Neglected Tropical Diseases*, 5, e952.
- Rice, J. (2006). *Mathematical statistics and data analysis*. Cengage Learning.
- Richards, S. A. (2005). Testing ecological theory using the information-theoretic approach: Examples and cautionary results. *Ecology*, 86, 2805–2814.
- Rinaldo, A., Bertuzzo, E., Mari, L., Righetto, L., Blokesch, M., Gatto, M., ... Rodriguez-Iturbe, I. (2012). Reassessment of the 2010–2011 Haiti cholera outbreak and rainfall-driven multiseason projections. *Proceedings of the National Academy of Sciences*, 109, 6602–6607.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6, 461–464.
- Shultz, A., Omollo, J. O., Burke, H., Qassim, M., Ochieng, J. B., Weinberg, M., ... Breiman, R. F. (2009). Cholera outbreak in Kenyan refugee camp: Risk factors for illness and importance of sanitation. *The American Journal of Tropical Medicine and Hygiene*, 80, 640–645.

- SourceForge. (2001, June) *Plot digitizer*. Retrieved August 10, 2015, from <http://http://plotdigitizer.sourceforge.net/>
- Tien, J., & Earn, D. (2010). Multiple transmission pathways and disease dynamics in a waterborne pathogen model. *Bulletin of Mathematical Biology*, 72, 1506–1533.
- Tuite, A. R., Tien, J., Eisenberg, M., Earn, D. J., Ma, J., & Fisman, D. N. (2011). Cholera epidemic in Haiti, 2010: Using a transmission model to explain spatial spread of disease and identify optimal control interventions. *Annals of Internal Medicine*, 154, 593–601.
- Verner, D. (Ed.). (2007). *Social resilience and state fragility in Haiti*. Washington, DC: A World Bank Country Study, World Bank Publications.
- Wikipedia. *Departments of Haiti*. Retrieved September 10, 2015, from https://en.wikipedia.org/wiki/Departments_of_Haiti
- Woodroffe, M. (1982). On model selection and the arc sine laws. *The Annals of Statistics*, 10, 1182–1194.
- World Health Organization. 2010a. *Cholera, fact sheet N° 107, updated July 2015*. Retrieved August 10, 2015, from <http://www.who.int/mediacentre/factsheets/fs107/en/>
- World Health Organization. (2010b, March). Cholera vaccines: Who position paper. *Weekly Epidemiological Record*, 85, 117–128.