



Review

Cite this article: Wu J, Dhingra R, Gambhir M, Remais JV. 2013 Sensitivity analysis of infectious disease models: methods, advances and their application. *J R Soc Interface* 10: 20121018.
<http://dx.doi.org/10.1098/rsif.2012.1018>

Received: 11 December 2012

Accepted: 21 June 2013

Subject Areas:

systems biology

Keywords:

sensitivity analysis, infectious disease modelling, Sobol' method, Morris method, partial rank correlation coefficient, sensitivity heat map

Author for correspondence:

Justin V. Remais

e-mail: justin.remais@emory.edu

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rsif.2013.1018> or via <http://rsif.royalsocietypublishing.org>.

Sensitivity analysis of infectious disease models: methods, advances and their application

Jianyong Wu¹, Radhika Dhingra¹, Manoj Gambhir² and Justin V. Remais^{1,3}

¹Department of Environmental Health, Rollins School of Public Health, Emory University, 1518 Clifton Road NE, Atlanta, GA 30322 USA

²MRC Centre for Outbreak Analysis and Modeling, Department of Infectious Disease Epidemiology, Imperial College London, London, UK

³Program in Population Biology, Ecology and Evolution, Graduate Division of Biological and Biomedical Sciences, Emory University, 1510 Clifton Road, Atlanta, GA 30322, USA

Sensitivity analysis (SA) can aid in identifying influential model parameters and optimizing model structure, yet infectious disease modelling has yet to adopt advanced SA techniques that are capable of providing considerable insights over traditional methods. We investigate five global SA methods—scatter plots, the Morris and Sobol' methods, Latin hypercube sampling—partial rank correlation coefficient and the sensitivity heat map method—and detail their relative merits and pitfalls when applied to a micro-parasite (cholera) and macroparasite (schistosomiasis) transmission model. The methods investigated yielded similar results with respect to identifying influential parameters, but offered specific insights that vary by method. The classical methods differed in their ability to provide information on the quantitative relationship between parameters and model output, particularly over time. The heat map approach provides information about the group sensitivity of all model state variables, and the parameter sensitivity spectrum obtained using this method reveals the sensitivity of all state variables to each parameter over the course of the simulation period, especially valuable for expressing the dynamic sensitivity of a microparasite epidemic model to its parameters. A summary comparison is presented to aid infectious disease modellers in selecting appropriate methods, with the goal of improving model performance and design.

1. Introduction

Mathematical models are widely used to examine, explain and predict the dynamics of infectious disease transmission, and models of specific diseases of global import have played a vital role in developing public health strategies for control and prevention [1–3]. Infectious disease models provide a mathematical representation of the dynamic transmission cycle, involving interactions between infected and susceptible hosts that are generally expressed as a set of coupled ordinary differential equations (ODEs) [3]. Model outputs, the ODE solutions over a simulation interval, provide a dynamic representation of the transmission process. Where possible, model parameters are estimated based on experimental or observational data, and where parameter values have gone unestimated, they are often set to plausible values or ranges based on analogous systems, statistical inference or expert opinion [3,4]. Recent advances in computational science and information technology have made widespread applications of mathematical modelling possible [2], yet model outputs often have complex, nonlinear relationships with model parameters, and inappropriate choices of parameter values can lead to bias in model outputs [5,6].

Sensitivity analysis (SA) characterizes the response of model outputs to parameter variation [7], helping to allocate resources to follow-up experimentation and field study; to isolate major sources of parametric uncertainty; to identify parameters that can be shed to yield a simpler model; to elucidate the plausible range of system outcomes, for forecasting purposes, when data are not available; and to

determine the robustness of a modelling study's qualitative conclusions [7–10]. Renewed attention to SA is warranted, particularly as advanced techniques have emerged that are capable of providing considerable insights over traditional SA methods. Modern parameter estimation techniques for complex models, often based upon Bayesian methods using, for example, Markov Chain Monte Carlo algorithms, cover complementary territory to SA; following data collection, parameter estimation ideally results in a joint posterior distribution over all of the model parameters given the available data. This posterior information could then be used to determine the sensitivity of the model fit—generally a log likelihood—to each of the parameters, by marginalizing the joint distribution. Much would remain to be known, however, about the parameter sensitivity of other model outputs of interest, leaving ample room for useful SA. Performing SA prior to data collection can also benefit our comprehension of a modelled complex system; indeed, SA can help to determine what data need be collected in order to most informatively narrow parameter and output uncertainty [11]. Similar approaches exist in many disciplines and guises, e.g. elasticity theory in economics and ecology [12,13], response surface methodology [14] and design of experiments [15] each used in multiple fields.

It should also be noted that SA techniques such as these may be used to explore the parameter sensitivity of the log-likelihood surface itself, with the Fisher Information Matrix describing gradients on this surface and identifying the most profitable parameters to measure experimentally [11]. There is, therefore, both a strong direct and indirect link between SA and parameter estimation. Here, we focus primarily on global sensitivity analysis (GSA) methods, which examine the response of model output variables to parameter variation within a selected parameter space, at a time when those working in epidemiological modelling may not be aware of recent advances in SA. GSA methods include variance-based methods (e.g. the Sobol' method and Fourier Amplitude Sensitivity Test) [16,17], global screening methods (e.g. the Morris method, also called the elementary effect method) [18,19], sampling-based methods (e.g. Monte Carlo filtering; Latin hypercube sampling with partial rank correlation coefficient index, LHS-PRCC) [20] and others. The recently developed sensitivity heat map (SHM) method, while not strictly a global method by the above definition, is capable of exploring the sensitivity of complex models to many parameters simultaneously and over time [11] and provides comparable results to those above.

GSA techniques have been broadly applied in systems biology [11,21], environmental modelling [22,23] and infectious disease modelling. Examples of recent published infectious disease models that describe and carry out an SA are shown in table 1, and, among these, LHS-PRCC [26–28,32] and other simple sensitivity methods [24,25,31] are frequently used. GSA methods popular in other mathematical modelling applications, such as the Sobol' and Morris methods, have sparsely been applied to infectious disease models. Observing the shortage of well-structured SA in such modelling studies, Bailey & Duppenthaler [30] summarized the range of SA tools available in the early 1980s. More recently, Okais *et al.* [5] emphasized the importance of SA in infectious disease modelling and used a varicella-zoster model to demonstrate how to conduct univariate and multivariate SA. As new GSA methods emerge that are capable of providing considerable insight over traditional methods, adoption by the infectious disease modelling community will be crucial to the advancement of the field.

Here, we assess and compare the performance of five GSA techniques—scatter plots, LHS-PRCC, the Morris method, the Sobol' method and the SHM method—by applying them to the previously published microparasite and macroparasite models and discuss their relative merits and pitfalls for application to infectious disease models generally.

2. Overview of sensitivity analysis methods for infectious disease models

2.1. Scatter plots

Scatter plots are occasionally used to visually examine the correlation between a model output variable and parameters. An output variable that is sensitive to the selected parameter will yield an obvious correlation pattern in the scatter plot. Generally, a Monte Carlo algorithm is used to sample the parameter space, and multiple scatter plots are drawn illustrating the relationship between each parameter and each output variable of interest [7]. Visual recognition of the correlation between parameter and model output values can be contingent on the choice of axis scales.

2.2. The Morris method

The Morris method, also called the elementary effects method, is based on the ratio of the change in an output variable to the change in an input parameter [18]. Given the general relationship between a model's output Y and input parameters X , $Y = f(X)$, the elementary effect of x_i can be expressed as

$$EE_i(\mathbf{X}) = \frac{y_i(x_1, x_2, \dots, x_i + \Delta, x_{i+1}, \dots, x_k) - y_j(\mathbf{X})}{\Delta}, \quad (2.1)$$

where $X \in [0, 1]^k$ is a scaled vector of k input parameters, y_i is the state variable of interest, Δ is a value in the set $\{1/(p-1), \dots, 1 - 1/(p-1)\}$ and p is the number of levels into which each dimension of the parameter space is divided. The distribution of $EE_i(\mathbf{X})$, denoted F_i , is obtained by repeated random sampling of X from its k -dimensional, p -level parameter space.

The mean of F_i (denoted μ), the mean of $|F_i|$ (denoted μ^* ; [19]), and the standard deviation of F_i (denoted σ) are the resulting sensitivity measures of an output variable to a parameter. A large μ suggests that the parameter has strong influence on the output, while a large σ suggests either that the relationship between the parameter and output is non-linear or the parameter interacts with other parameters [33]. Importantly, when the F_i distribution contains both positive and negative values, elementary effects may cancel each other out to produce a small μ , but the parameter may still be influential; thus, the use of the absolute elementary effect, μ^* , as a remedy, has been recommended. As the elementary effect of x_i as estimated by the Morris method bears more resemblance to $\Delta y_j / \Delta x_i|_X$ rather than $\partial y_j / \partial x_i|_X$ [18], the accuracy of the Morris index depends on the smoothness of y over the parameter domain.

The Morris method takes $r(k+1)$ model runs, where r is the number of trajectories (a sampling path which begins at a random point and each subsequent point varies only one randomly selected parameter, holding all others constant at the values from the last point) through the input parameter space [18] and $EE(x_i)$ is calculated at $k+1$ points along each trajectory.

Table 1. Select infectious disease modelling literature providing quantitative SA results. SA, sensitivity analysis; n.a., not applicable; PRCC, partial rank correlation coefficient; LHS, Latin hypercube sampling.

pathogen/ disease	no. state variables	no. parameters	SA method	objectives and findings of the SA	reference
cholera	7	19	one-way SA	objective: to determine which parameters had the greatest influence on model projections. <i>Finding:</i> reservoir size did not substantially affect projected cases	[24]
rabies virus	8	18	simple graph comparison	objectives: to examine whether initial conditions affected the number of rabies cases; to find which control strategies were most effective. <i>Finding:</i> initial conditions of dogs not only affected case numbers, but also affected the timing of peak cases; reducing the annual crop of newborn puppies and increasing the dog immunization rate reduced the rabies incidence rate	[25]
bluetongue virus	8	18	LHS-PRCC	objective: to find which parameters impact the basic reproduction number, R_0 . <i>Finding:</i> temperature, the probability of transmission from host to vector, and the vector to host ratio were the most important factors	[26]
malaria	7	17	normalized forward sensitivity index	objective: to determine the sensitivity of the basic reproduction number, R_0 , to model parameters. <i>Finding:</i> R_0 was most sensitive to the mosquito biting rate, suggesting that strategies targeting the mosquito biting rate may be effective in malaria control and improved estimates of the mosquito biting rate may improve model precision	[9]
HIV	34	20	LHS-PRCC	objective: to determine which parameters were important in contributing uncertainty to predictions of an HIV model. <i>Finding:</i> three key parameters were identified as critical for precisely predicting the number of HIV cases	[27]
tuberculosis	5	11	PRCC	objective: to understand the role of key parameters in a model of tuberculosis transmission in the absence of treatment. <i>Finding:</i> uncertainty in the effective contact rate was an important contributor to uncertainty in prevalence of the disease	[28]
dengue	n.a.	n.a.	regression analysis	objective: to explore the response of dengue model outputs to changes in parameters. <i>Finding:</i> survival in each life stage of the mosquito, mosquito biting behaviour and the duration of the infectious period in humans were the most important parameters	[29]
typhoid fever	9	24	partial differentiation	objective: to demonstrate how to conduct SA on infectious disease models. <i>Finding:</i> partial differentiation was a useful method when applied to infectious disease models	[30]
AIDS/HIV	n.a.	15	one-way SA	objective: to illustrate the effect of small changes in parameter values on model outputs. <i>Finding:</i> the model was most sensitive to the migration rate, the number of stages in the progression to AIDS and the average number of new sexual partners	[31]
tuberculosis	10	17	LHS-PRCC	objective: to quantify the impact of variation in parameters on TB prevalence. <i>Finding:</i> transmission rate, proportion of fast progressors, initial immunity afforded by previous infection and death rate of untreated TB were the top influential factors	[32]

Although the Morris method is easy to understand, does not depend on assumptions about the model (e.g. monotonicity) and is computationally inexpensive, it cannot quantify the contribution of a parameter to the variability of the output.

2.3. Latin hypercube sampling-partial rank correlation coefficient

A measure of the nonlinear, but monotonic, relationship between two variables, the PRCC is an efficient and reliable sampling-based SA method that provides a measure of monotonicity between parameters and model output after removing the linear effects of all parameters except the parameter of interest [21]. A standard correlation coefficient, ρ , for two variables, x and y , is calculated as follows:

$$\rho = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}}, \quad (2.2)$$

where $\{(x_i, y_i) | x_i \in x, y_i \in y\}$ are the set of paired, sampled data, \bar{x} is the sample mean of x , and \bar{y} is the sample mean of y . The PRCC determines the sensitivity of an output state variable to an input parameter as the linear correlation, ρ , between the residuals, $(X_j - \hat{X}_j)$ and $(Y - \hat{Y})$, where X_j is the rank transformed, sampled j th input parameter, and Y is the rank transformed output state variable, while keeping all other parameter values fixed [34]; \hat{X}_j and \hat{Y} are determined for k samples by the linear regression models

$$\hat{X}_j = c_0 + \sum_{\substack{p=1 \\ p \neq j}}^k c_p X_p \quad (2.3)$$

and

$$\hat{Y} = b_0 + \sum_{\substack{p=1 \\ p \neq j}}^k b_p X_p. \quad (2.4)$$

PRCC is often combined with LHS in the conduct of an SA [21,27]. LHS is a stratified Monte Carlo sampling method [21,35] that divides each parameter's range into N equal intervals and randomly draws one sample from each interval. The method explores the entire range of each parameter and each interval for each parameter is sampled only once [21,27]. Because of its dense stratification over the input parameter space and the rapid convergence of the sample mean to the true population mean as the number of samples increases, LHS is more efficient—requiring fewer simulations—than a general Monte Carlo sampling approach [36]. The combined LHS-PRCC procedure is described fully elsewhere [21,27], but generally involves: (i) LHS of the parameter space, (ii) obtaining model output for each set of sampled parameters, (iii) ranking parameter and output values and replacing their original values with their ranks, and (iv) calculating the PRCC for each input parameter.

2.4. The Sobol' method

The Sobol' method is a variance-based GSA technique capable of estimating the influence of individual parameters, or a group of parameters, on the output variables of a nonlinear model [16,37]. Given a model of the relationship between output variables and parameters, $Y = f(\mathbf{X}) = f(x_1,$

$x_2, \dots, x_k)$ that is square integrable over its unit hypercube parameter space, the model function a single state variable, $y = f(\mathbf{X})$, can be decomposed into summands of increasing dimensionality, known as the high-dimensional model representation (HDMR; [16])

$$y = f(\mathbf{X}) = f_0 + \sum_{i=1}^k f_i(x_i) + \sum_{j>i}^k f_{ij}(x_i, x_j) + \dots + f_{1,2,\dots,k}(x_1, x_2, \dots, x_k). \quad (2.5)$$

Sobol' demonstrated that if each term in this expansion has a zero mean, then the total variance of an output variable can be decomposed into the HDMR ANOVA, represented as

$$V(y) = \int f(\mathbf{X})^2 d\mathbf{X} - f_0 = \sum_{i=1}^k V_i + \sum_i \sum_{j>i}^k V_{ij} + \sum_i \sum_{j>i}^k \sum_{h>j}^k V_{ijh} + \dots + V_{1,2,\dots,k}, \quad (2.6)$$

where $V(y)$ is the variance of the model output y , k is the number of parameters and $V_{i_1, i_2, \dots, i_s} = \int f_{i_1, i_2, \dots, i_s}^2 dx_{i_1} dx_{i_2}, \dots, dx_{i_s}$ for a given set of indices, i_1, \dots, i_s . The Sobol' sensitivity indices are the ratios of the partial variance given an individual parameter or the interactions of a parameter subset to the total variance. Two Sobol' indices are often calculated; the main effects—also called the 'first-order index'—and the total effects of a parameter, are expressed, respectively, as [16]

$$S_i = \frac{V[E(y|x_i)]}{V(y)} \quad (2.7)$$

and

$$S_{T_i} = 1 - \frac{V[E(y|\mathbf{X}_{\sim i})]}{V(y)}, \quad (2.8)$$

where $\mathbf{X}_{\sim i}$ denotes all elements of \mathbf{X} except x_i . These indices have the property that $S_i \leq S_{T_i} \leq 1$ and when $S_i = S_{T_i} = 0$, it can be concluded that $f(\mathbf{X})$ does not depend on x_i , while $S_i = S_{T_i} = 1$ indicates that $f(\mathbf{X})$ depends solely on x_i [16].

Considered versatile and effective [38], the Sobol' method has several key advantages. The method is independent from model structure (e.g. linearity and monotonicity); it captures the effects of individual parameters as well as their interactions and it provides quantitative information on the contribution of each parameter to model sensitivity. A major disadvantage is the high computational cost: given a model with k parameters and n samples from each parameter, the model must run at least $k*(n+2)$ times [7] to generate sufficient output in order to calculate partial variances for individual parameters.

2.5. Sensitivity heat map method

Two new graphical techniques have been developed to analyse the sensitivity of complex models: the SHM and the parameter sensitivity spectrum (PSS) [11]. An SHM depicts the sensitivity of each output variable to all (or a subset of) model parameters, while the PSS shows the sensitivity of all (or a subset of) output variables to each individual parameter. The methods were developed to take advantage of the fundamental property of dynamic systems, where a change in high-dimensional

parameter spaces will lead to a low-dimensional change in output variables [11]. Given a set of n ODEs:

$$\frac{d\mathbf{X}}{dt} = f(t, \mathbf{X}, \mathbf{K}), \quad (2.9)$$

where $\mathbf{X} = (x_1, \dots, x_n)$ is the vector of state variables, t is time and $\mathbf{K} = (k_1, \dots, k_s)$ is a vector of parameters, there is a solution or set of solutions $\mathbf{X} = \mathbf{g}(t, \mathbf{K})$ for $0 \leq t \leq T$. A change in parameter, $\delta\mathbf{K}$ results in a change in the solution, $\delta\mathbf{g}$ such that $\delta\mathbf{g}(t) = M\delta\mathbf{K} + \mathcal{O}(\|\delta\mathbf{K}\|^2)$, where M is the linear map from parameter space \mathbb{R}^s to a Hilbert space \mathcal{H} (as defined in [39]). Using the notation of Rand [11], a time-interval normalized (i.e. by $\sqrt{(t_{m+1} - t_m)/t_N}$) matrix, M_{t_i} , composed of the partial derivatives $\partial g_j / \partial k_i$ at all simulation time points $t = \{t_1, \dots, t_N\}$ is subjected to singular value decomposition to arrive at values for \mathbf{U} , \mathbf{W} and $\{\sigma_i\}$, defined below, which express the sensitivity of output variables to parameters [11]. Rand shows that, at a minimum of average error, there exists a unique sequence of positive numbers $\sigma_1 \geq \dots \geq \sigma_s$, a unique set of orthonormal vectors $\{\mathbf{V}_i\}$ in \mathbb{R}^s and a unique n -dimensional orthogonal time-series basis set $\{\mathbf{U}_1(t), \dots, \mathbf{U}_n(t)\}$ in \mathcal{H} , such that a small change in $\delta\mathbf{g}(t, \mathbf{K})$ is given by $\delta\mathbf{g}(t) = \sum_i \sigma_i [\sum_j W_{ij} \delta k_j] + \mathcal{O}(\|\delta\mathbf{K}\|^2)$, where W_{ij} are the elements of $\mathbf{W} = \mathbf{V}^{-1}$, and $\mathbf{V} = [\mathbf{V}_1, \dots, \mathbf{V}_n]$ [11]. Thus, the effect of a single parameter on the system, the PSS, is expressed by $\partial \mathbf{g}(t) / \partial k_j = \sum_i \sigma_i W_{ij} \mathbf{U}_i(t)$, while the SHM for state variable $g_i(t)$ is depicted as $\sigma_i (\max_j |W_{ij}|) \mathbf{U}_i(t)$. Compared with other SA methods, this highly computationally efficient approach allows the modeller to observe the sensitivity characteristics of the set of all output variables, instead of a single output variable.

3. Material and methods

3.1. Infectious disease models

A previously described a microparasite (cholera) model [24] and a macroparasite (schistosomiasis) model [40] were used to examine the merits of the five SA methods. Microparasites (e.g. bacteria and viruses) generally replicate within hosts at a rapid rate and are able to generate an epidemic in a short period of time. Characteristics of the microparasite epidemic curve (the time course of the infected population) are often studied using ODE models. Macroparasites are generally larger and longer lived than microparasites; they include parasitic helminths and certain vector-borne diseases. Macroparasites generally cannot carry out their full life cycle within a single host, and their transmission is often sustained over time, achieving steady-state, endemic disease dynamics capable of rising from low levels rapidly, akin to a microparasite epidemic curve, but then rising or declining from the endemic equilibrium very slowly [1].

3.1.1. Cholera model

A microparasite model describing the transmission dynamics of cholera in Haiti in 2010 was previously used to examine the potential effect of interventions on cases of cholera and mortality [24]. This model was slightly modified for the present analysis by excluding the effects of vaccination, leading to a final model with six ODEs and 16 parameters

$$\begin{aligned} \frac{dS(t)}{dt} &= \mu N(t) + \omega R(t) - \alpha S(t) \frac{B_L(t)}{K_L + B_L(t)} \\ &\quad - \alpha S(t) \frac{B_H(t)}{K_H + B_H(t)} - \mu S(t), \end{aligned} \quad (3.1)$$

$$\begin{aligned} \frac{dI(t)}{dt} &= (1-p)\alpha S(t) \frac{B_L(t)}{K_L + B_L(t)} + (1-p)\alpha S(t) \\ &\quad \times \frac{B_H(t)}{K_H + B_H(t)} - (\mu_c + \mu + (1-\theta)\gamma + \theta\gamma\lambda)I(t), \end{aligned} \quad (3.2)$$

$$\begin{aligned} \frac{dA(t)}{dt} &= p\alpha S(t) \frac{B_L(t)}{K_L + B_L(t)} + p\alpha S(t) \frac{B_H(t)}{K_H + B_H(t)} \\ &\quad - (\mu + \gamma)A(t), \end{aligned} \quad (3.3)$$

$$\frac{dR(t)}{dt} = ((1-\theta)\gamma + \theta\gamma\lambda)I(t) + \gamma A(t) - (\mu + \omega)R(t), \quad (3.4)$$

$$\frac{dB_H(t)}{dt} = (\psi\theta + (1-\theta)) \frac{\xi_S}{W} I(t) + \frac{\xi_A}{W} A(t) - \chi B_H(t), \quad (3.5)$$

$$\text{and } \frac{dB_L(t)}{dt} = \chi B_H(t) - \delta B_L(t), \quad (3.6)$$

where $N(t)$ is the total population, $S(t)$ the susceptible population, $A(t)$ the asymptomatic population, $I(t)$ the infected population, $R(t)$ the recovered population and $B_H(t)$ and $B_L(t)$ are the concentration of high-infectious cholera and low-infectious cholera, respectively, in environmental reservoirs. Model parameters are summarized in table 2 and are fully described elsewhere [24]. The model was run in Matlab (R 2011b) for 60 time steps with each parameter set, as described later.

3.1.2. Schistosomiasis model

A macroparasite model describing the transmission of schistosomiasis was previously used to simulate the worm burden in human hosts and the density of infected snails in China [40,41]. For the present analysis, the model was simplified to two state variables and 14 parameters as follows:

$$\frac{dW(t)}{dt} = \beta c - \mu W(t) - \pi W(t), \quad (3.7)$$

$$\frac{dZ(t)}{dt} = \rho m x - \varepsilon Z(t), \quad (3.8)$$

$$e = 0.5 h g n W(t) \phi, \quad (3.9)$$

$$m = \frac{\alpha e}{b} \quad (3.10)$$

$$\text{and } c = \frac{\sigma Z(t) a}{b}, \quad (3.11)$$

where $W(t)$ is the average population worm burden, $Z(t)$ is the density of infected snails and e , m and c are parasite eggs released into environment, miracidia density and cercaria density, respectively. Model parameters are summarized in table 2 and fully described elsewhere [40,41]. The model was run in Matlab (R 2011b) for 300 time steps with each parameter set, as described later.

3.2. Set-up of parameter space and distribution

Parameter values and ranges were adopted from previous work [24,40,41], and unspecified parameter ranges were assigned by dividing (lower bound) and multiplying (upper bound) the parameter point estimate by 2. A uniform distribution was assigned to parameters where the distribution was unspecified [21]. All sensitivity analyses were run in Matlab (R 2011b) unless otherwise indicated, and model outputs were sampled at every time step (when examining sensitivity over time) and near the peak level of infection at the 15th simulation step (when examining sensitivity at a single point in time).

3.3. Sensitivity analysis

The primary model outputs of interest for the sensitivity analyses were the infected population, $I(t)$, in the microparasite model and the worm burden, $W(t)$, in the macroparasite model.

Table 2. Parameter definitions and ranges for the cholera and schistosomiasis models.

model	parameter (units)	symbol	value	range	code
cholera model	rate of contaminated water consumption	α	0.1	0.05–0.2	<i>p1</i>
	Low-infectious (LI) <i>V. cholera</i> infectious dose (cells)	κ_L	10^5	5×10^4 – 2×10^5	<i>p2</i>
	Hyperinfectious (HI) <i>V. cholera</i> infectious dose (cells)	κ_H	2000	1000–4000	<i>p3</i>
	rate of decay of HI to LI <i>V. cholera</i> (per day)	χ	1	0.5–2	<i>p4</i>
	death rate of <i>V. cholera</i> in the environment (per day)	δ	0.033	0.0167–0.0667	<i>p5</i>
	natural birth/death rate (per day)	μ	4.49×10^{-5}	2.25×10^{-5} – 8.98×10^{-5}	<i>p6</i>
	mortality rate, symptomatic cholera	μ_c	0.0194	0.0097–0.0388	<i>p7</i>
	recovery rate, cholera (per day)	γ	0.2	0.1–0.4	<i>p8</i>
	proportion of cases asymptomatic patient	p	0.79	0.6–0.9	<i>p9</i>
	rate of excretion of <i>V. cholera</i> , symptomatic patient (cell/day)	ξ_S	1.3×10^{11}	6.5×10^{10} – 2.6×10^{11}	<i>p10</i>
	rate of excretion of <i>V. cholera</i> , asymptomatic patient (cell/day)	ξ_A	1.3×10^8	6.5×10^7 – 2.6×10^8	<i>p11</i>
	rate of waning of natural immunity (per day)	ω	0.00137	0.00055–0.0055	<i>p12</i>
	proportion symptomatic individuals receiving antibiotics	Θ	0.08	0.04–0.16	<i>p13</i>
	relative rate of shedding, receiving antibiotics	ψ	0.52	0.26–1.04	<i>p14</i>
	relative rate of recovery, receiving antibiotics	λ	2.3	1.05–4.6	<i>p15</i>
	size of water reservoir (L^* day)	W	5.3×10^{10}	5.3×10^9 – 5.3×10^{11}	<i>p16</i>
schistosomiasis model	snail density in a village (snails per square metre)	x	20	10–40	<i>q1</i>
	human population in a village	n	200	100–400	<i>q2</i>
	snail habitat area in a village (m^2)	a	20 000	10 000–40 000	<i>q3</i>
	surface water area in a village (m^2)	b	20 000	10 000–40 000	<i>q4</i>
	snail infection	ρ	0.000005	0.0000025–0.00001	<i>q5</i>
	per capita snail mortality (per day)	ε	0.003	0.0015–0.006	<i>q6</i>
	cercarial production (per infected snail)	σ	50	25–100	<i>q7</i>
	human infection in a village (per cercaria per day)	β	0.01	0.0005–0.002	<i>q8</i>
	per capita worm mortality (per day)	μ	0.00091	0.00046–0.00182	<i>q9</i>
	egg production (per worm pair per gram stool)	h	5	2.5–10	<i>q10</i>
	stool production (g)	g	100	50–200	<i>q11</i>
	worm mating probability	ϕ	0.4	0.2–0.8	<i>q12</i>
	miracidial production (per egg)	α	0.01	0.005–0.02	<i>q13</i>
	chemotherapy intervention in a village	π	0.05	0.025–0.1	<i>q14</i>

Recall, however, that the SHM method evaluates all output variables simultaneously.

3.3.1. Scatter plots

For both models, 1000 samples from a uniform distribution of each parameter range were generated and used as simulation inputs to obtain the output values of interest. Scatter plot matrices were drawn showing the relationship between each parameter and the output variable of interest at a single time point (the 15th time step).

3.3.2. The Morris methods

For both models, each parameter range was divided into six levels ($p = 6$, $\Delta = 0.6$). A total of 100 trajectories of 17 points and 15 points, respectively, were built for the cholera model and schistosomiasis model, as described elsewhere [18]. For each parameter, 1700 and 1500 samples were generated for the cholera and schistosomiasis models, respectively. Model outputs were then obtained for each parameter set, and Morris indices, including μ , μ^* and σ , were calculated as described earlier.

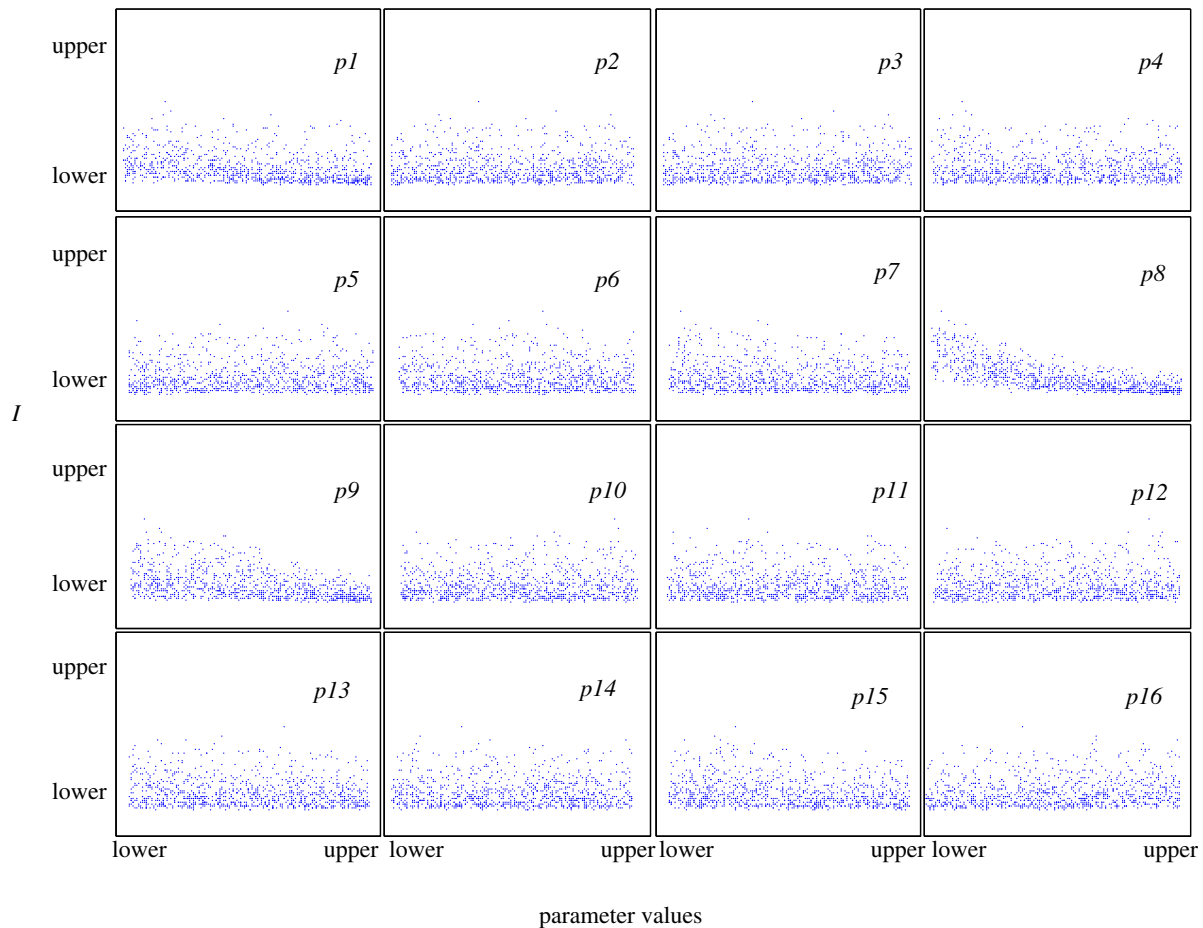


Figure 1. Scatter plots illustrating the relationship between the cholera model infected population size, I , examined at the peak of the epidemic ($t = 15$) and parameters.

3.3.3. Latin hypercube sampling with partial rank correlation coefficient index

Using Latin hypercube sampling, 100 samples from a uniform distribution of the parameter ranges of both models were taken. Model outputs were obtained and the parameter and output values were transformed into their ranks. PRCCs were calculated following the procedures described previously [21]. PRCC falls between -1 and $+1$, with an absolute value of PRCC close to 1 indicating the parameter has a strong impact on the model output.

3.3.4. The Sobol' method

Using Latin hypercube sampling, 100 samples were drawn from a uniform distribution of each parameter range for both models. Samples were permuted randomly 100 times yielding 10 000 parameter sets that were used in models runs. The first order of the Sobol' index was calculated as described earlier.

3.3.5. The sensitivity heat map method

The software package SASSY [42] was used to analyse the ODE models and to generate SHMs and PSS within Matlab (R2011b). After carrying out a limit cycle simulation using the Matlab ODE45 solver, model output was obtained for 60 time steps for the cholera model and 300 time steps for the schistosomiasis model, and the derivatives of model variables with respect to model parameters for all simulation time points were entered into a partial derivative matrix. Following the previous work [11], singular value decomposition of the partial derivative matrix was used to generate SHM and PSS, respectively. For SHMs, the sensitivity strip of each output variable was sorted from high to low by its maximum value and scaled by a ratio of its maximum value to

the maximum value of the colour bar to produce the sensitivity strip scale.

4. Results

A typical epidemic curve of the infected population size in the cholera model and the endemic curve of the worm burden in the human host in the schistosomiasis model are shown in the electronic supplementary material, figure S1, generated using the parameter values listed in table 2. The epidemic (cholera) curve rises sharply, quickly peaking at day 6 and dropping back to baseline after day 30. By contrast, the endemic (schistosomiasis) curve has a blunt peak at day 50 followed by a slow decline, illustrating the sustained burden of long-lived macroparasites in human hosts.

Cholera model scatter plots (figure 1) show the variation in the infected population size with changes in parameters when examined at the peak of the epidemic (time step 15). Model sensitivity to $p8$ and $p9$ is evident, but sensitivity to other parameters is unclear. The Morris index (figure 2) concurs, finding that $p8$ and $p9$ exert the strongest influence on the infected population size. However, the Morris index also identifies parameters $p1$, $p15$, $p7$, $p13$ and $p16$ as influential. The PRCC index provides a measure of the relative influence of these parameters (as well as the direction of the relationship), agreeing in the ranking of important parameters with the Morris index. Finally, the Sobol' sensitivity index indicates that parameters $p8$, $p9$ and $p1$ most strongly influence infected population size when evaluated near the peak population.

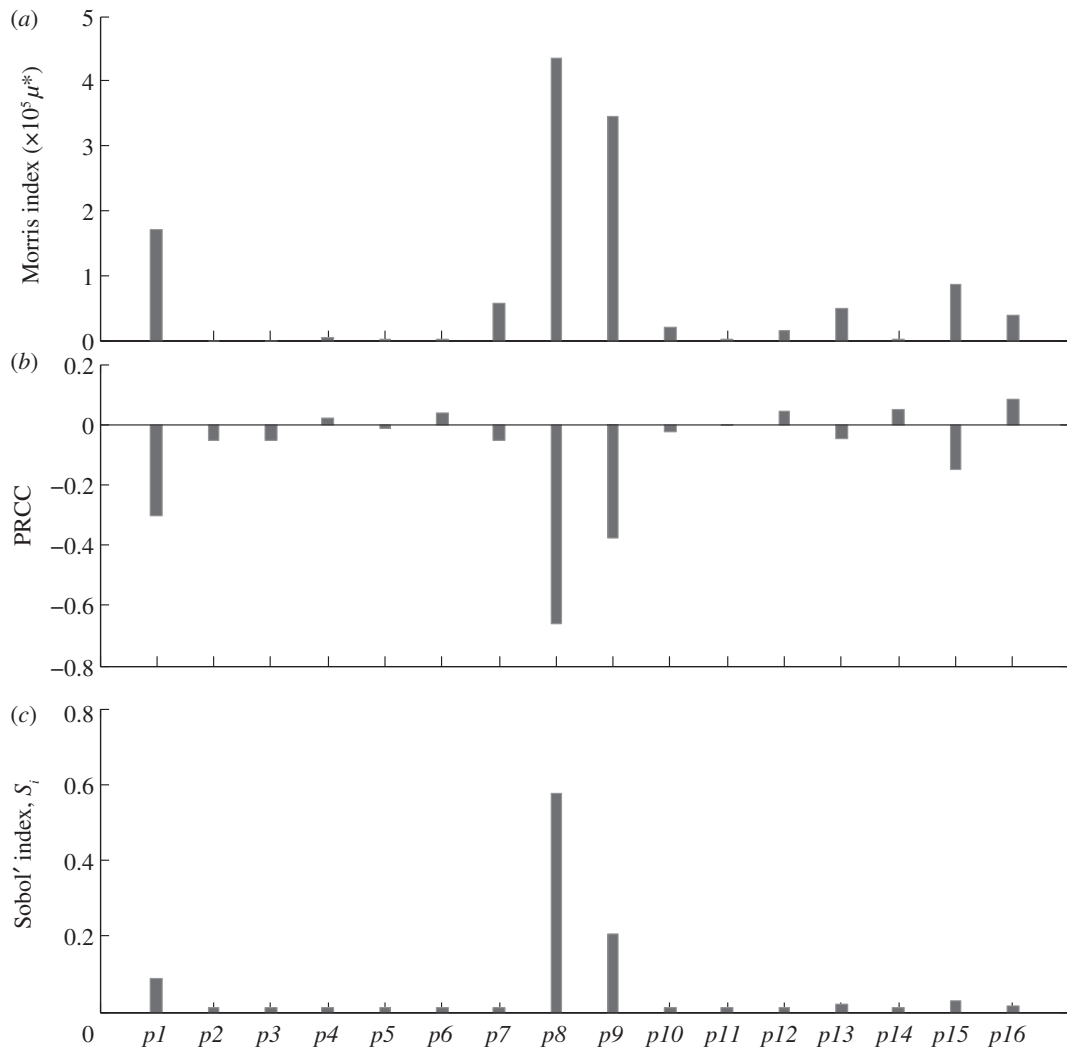


Figure 2. Sensitivity of the infected population size to changes in parameters in the cholera model as indicated by (a) the Morris index, μ^* , (b) the PRCC index and (c) the Sobol' index, S_i .

When the sensitivity index for an output variable is examined over the time course of the simulation, the relative importance of parameters is shown to vary widely depending on the time point of inspection (figure 3). For the cholera model, the Morris index shows p_9 , p_8 and p_1 to be the most influential parameters on the infected population size, but their relative importance changes during the simulation, where for instance p_8 overtakes p_1 and p_9 in importance after day 10. However, the index values of these parameters all decline after day 10 and are close to 0 at the end of simulation. The PRCC and the Sobol' indices also show variation in parameter sensitivity; however, these indices cannot meaningfully be interpreted over time. Both PRCC and Sobol', for instance, suggest that p_{12} has no influence on the infected population size state variable at the beginning of the simulation, but then has influence later in the simulation period (figure 3). Yet, neither measure is indexed to some unchanging value (e.g. a perturbation of the parameter of interest as in the Morris method or unit change in infected population). Thus, without accounting for the overall variation of the infected population size over the simulation period (see the electronic supplementary material, figure S1a), these indices cannot be interpreted as indicating that p_{12} is more important later in the simulation than at early time points. If this variation is taken into account—i.e. that the state variable is

small near the end of the outbreak—the influence of p_{12} might not in fact be high.

The sensitivity of each cholera model output variable to all parameters as a whole is shown by the SHM in figure 4a. The colour strip indicates the product of the first singular value (σ_1) and the first principle component ($pc1$) of the derivative of the solution (the state variable) to the parameter (UI_j). Thus, by comparison of the scale ratio for each sensitivity strip, it is clear that state variables B_H , B_L and I are much more sensitive to parameters as a whole than are A , S and R . Looking at individual state variables, sensitivity differs widely across the simulation time period. For B_H , A , I and S , sensitivity is highest early in the simulation interval, during the epidemic peak. High sensitivity of B_L is achieved in the middle of the simulation, versus late in the period for R . This is because B_L gradually peaks in the middle period and then decreases, while R reaches and remains at its highest value in the late period. Figure 4b shows the sensitivity of output variables as a whole to each individual parameter, confirming parameter p_9 as most influential when evaluated simultaneously across all state variables. Interestingly, parameter p_1 , individually important for determining the infected population size (figure 2) is of limited importance when sensitivity is examined together for all state variables (figure 4b).

Applying the five methods to the schistosomiasis model yields several interesting contrasts. The Morris, PRCC and

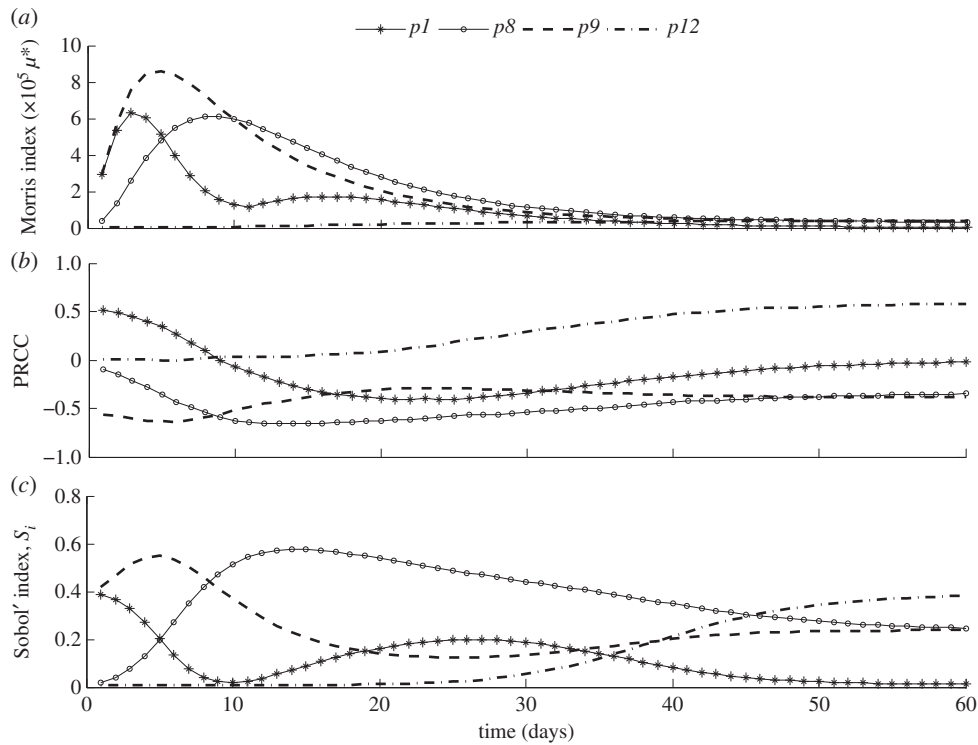


Figure 3. The temporal variation of the sensitivity of the infected population size to key parameters in the cholera model indicated by (a) the Morris index μ^* , (b) the PRCC index and (c) the Sobol' index S_i . Sensitivity to all model parameters is shown in the electronic supplementary material, figure S2.

Sobol' indices all indicate the importance of parameters $q3$, $q4$, $q6$, $q7$ and $q14$. Though both models show temporally varying sensitivity indices, the sensitivity indices of the most important parameters in the schistosomiasis model are, unlike the cholera model (figure 3), approximately constant later in the simulation (figure 5). The SHM for the schistosomiasis model (figure 6a) shows that the worm burden state variable is considerably more sensitive to model parameters, when taken as a whole, than the infected snail density variable. That the worm burden is highly sensitive to model parameters from days 40 to 120, and is subsequently slightly less sensitive afterward, is consistent with the results of traditional sensitivity methods as shown in figure 5 and corresponds to the endemic curve of the worm burden which reaches its peak at day 40 and is stable through the remainder of the simulation (see the electronic supplementary material, figure S1b). Finally, the PSS (figure 6b) reveals that both state variables, considered together, are sensitive to $q3$, $q4$, $q6$, $q7$, $q8$ and $q14$. These PSS-determined sensitive parameters correspond, with the exception of $q8$, with the results of traditional methods applied only to worm burden (figure 5).

5. Discussion

We described five GSA methods and examined their relative merit in identifying important parameters and characterizing their influence in two infectious disease mathematical models. The methods yielded similar results with respect to identifying influential parameters, but offered different insights as to the importance of parameters over the simulation period, as well as different information when applied to microparasite versus macroparasite models.

Scatter plots offer a straightforward way to visualize the relationship between parameters and output variables, providing a natural starting point for SA [7,21]. The Morris, LHS-PRCC and Sobol' methods can provide more definitive information, effectively distinguishing between parameters that are of high versus negligible importance. The Morris method is mainly used as a screening technique for identifying important parameters [7,35], while the PRCC index, ranging from -1 to $+1$, provides specific qualitative information on the relationship between an output variable and a parameter—as well as indicates whether the relationship is concordant or discordant—after controlling for the influence of other parameters [21]. Based on the Sobol' first-order sensitivity index, S_i , four parameters were found to be the primary contributors to the variance of the infected population size in the cholera model, and five were found to greatly influence worm burden in the schistosomiasis model. The Sobol' index, a variance decomposition approach, also captures the influence of the interaction among parameters on a model output [7]. The first-order sensitivity index, S_i , reflects the contribution of each parameter to the variance of the output variable, while the total effect index, S_T , takes both individual contributions and interaction effects into account [7,38]. An exploration of the interaction effects was outside the scope of our application of the Sobol' index, and thus the total effect index was not calculated.

The temporal variation of state variable sensitivity to parameters is rarely addressed in infectious disease modelling studies. Modellers should explore this temporal effect, as their conclusions about the importance of particular parameters, or the robustness of their findings, may be mistaken if sensitivity is examined only at one time point. If, for instance, sensitivity is evaluated at a point following the epidemic peak, parameters that were instrumental to the rising epidemic curve may appear irrelevant. The

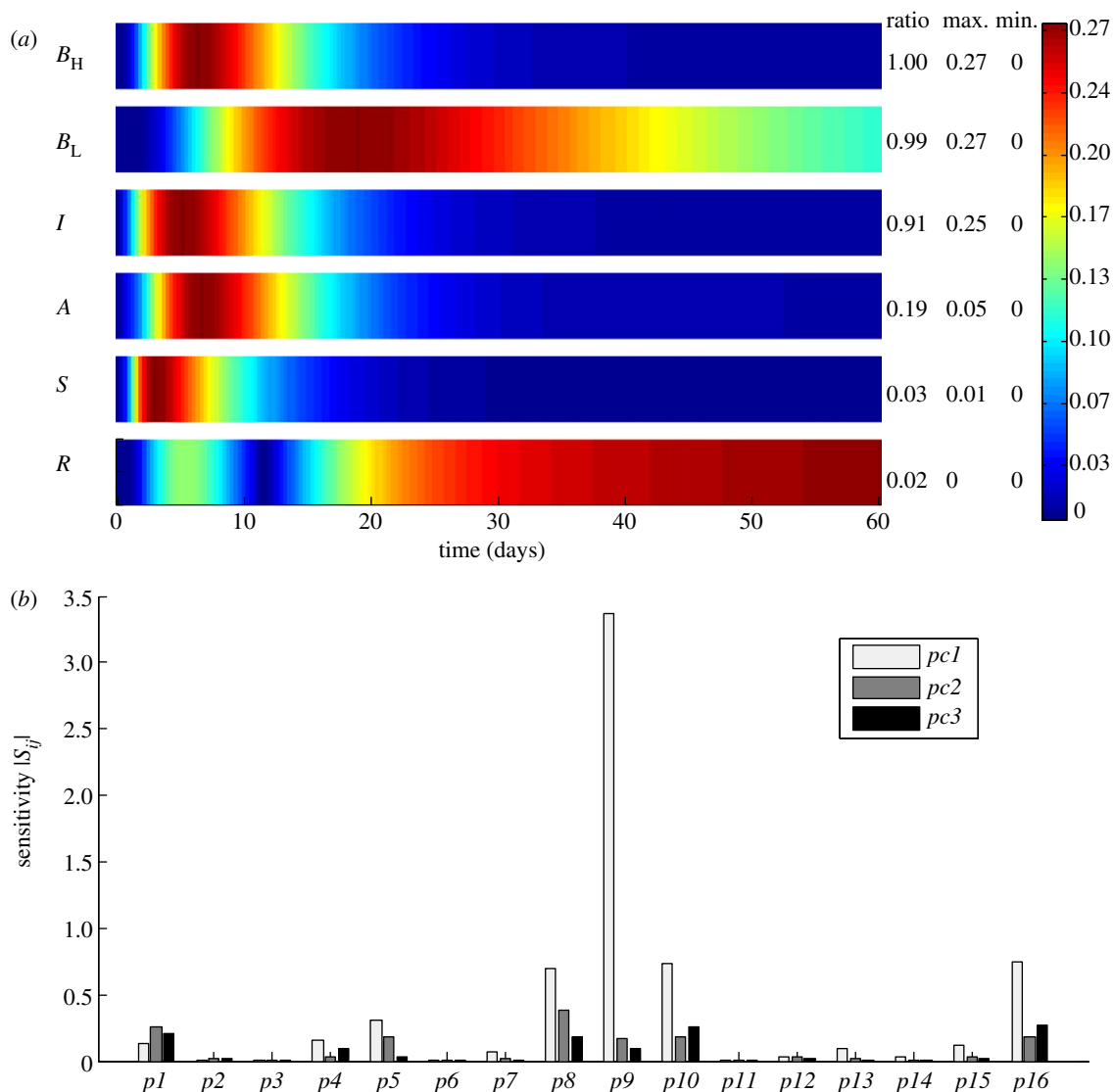


Figure 4. (a) SHM and (b) PSS for the cholera model, following previous work [11].

Morris method can provide some information about temporal variation in a single parameter's sensitivity over the simulation period, being as the index is based on the ratio of the change of an output to the change of an input parameter, the latter of which can be kept constant over the period calculated. However, the index is sensitive to the magnitude of the state variables over the course of the simulation, and thus must be interpreted with care when examined over time.

By contrast, the Sobol' method cannot provide a meaningful comparison of the change in sensitivity of an output variable to a parameter over the course of simulation. The Sobol' sensitivity index for a particular parameter represents its contribution to the total variance of the system at the time point evaluated. If examined over time, a change in the index may be related to a change in the importance of the parameter or a change in the output variance, and these two possibilities cannot be distinguished. Likewise, the PRCC method does not allow for quantitative comparison of a single parameter's influence at two time points, because the calculation ranks parameters and state variables at a point in time; the rank, residual of a rank and the resulting PRCC of a particular parameter at one time point has no direct relationship to those at another time point. Modellers should thus be cautious about applying variance decomposition-

based methods to quantitatively compare model sensitivity at multiple time points.

The SHM method, and the associated graphical objects SHM and PSS, offer considerable benefits over traditional techniques. This method efficiently handles models with high-dimensional parameter spaces and many state variables. SHM offers a powerful view of state variable sensitivity over time to many parameters taken together [11]. The sensitivity strips clearly show the time period when a state variable is sensitive to model parameters as a whole, information that traditional methods do not provide. Such an examination of sensitivity over time is of particular value to the modeller interested in, say, the change in timing of particular events occurring over the course of the epidemic. PSS, on the other hand, summarizes the importance of individual parameters when considering all, or a set of, model state variables. This view may be of particular use in agent-based or network transmission models, where modellers are most interested in how parameters influence the model system as a whole, rather than the sensitivity of a specific state variable as defined for a particular network node or agent [21,43].

Alternatively, the simultaneous sensitivity of several state variables (e.g. infectious state variables of the cholera model, $I(t)$, $B_H(t)$ and $B_L(t)$), to a single parameter may also be of

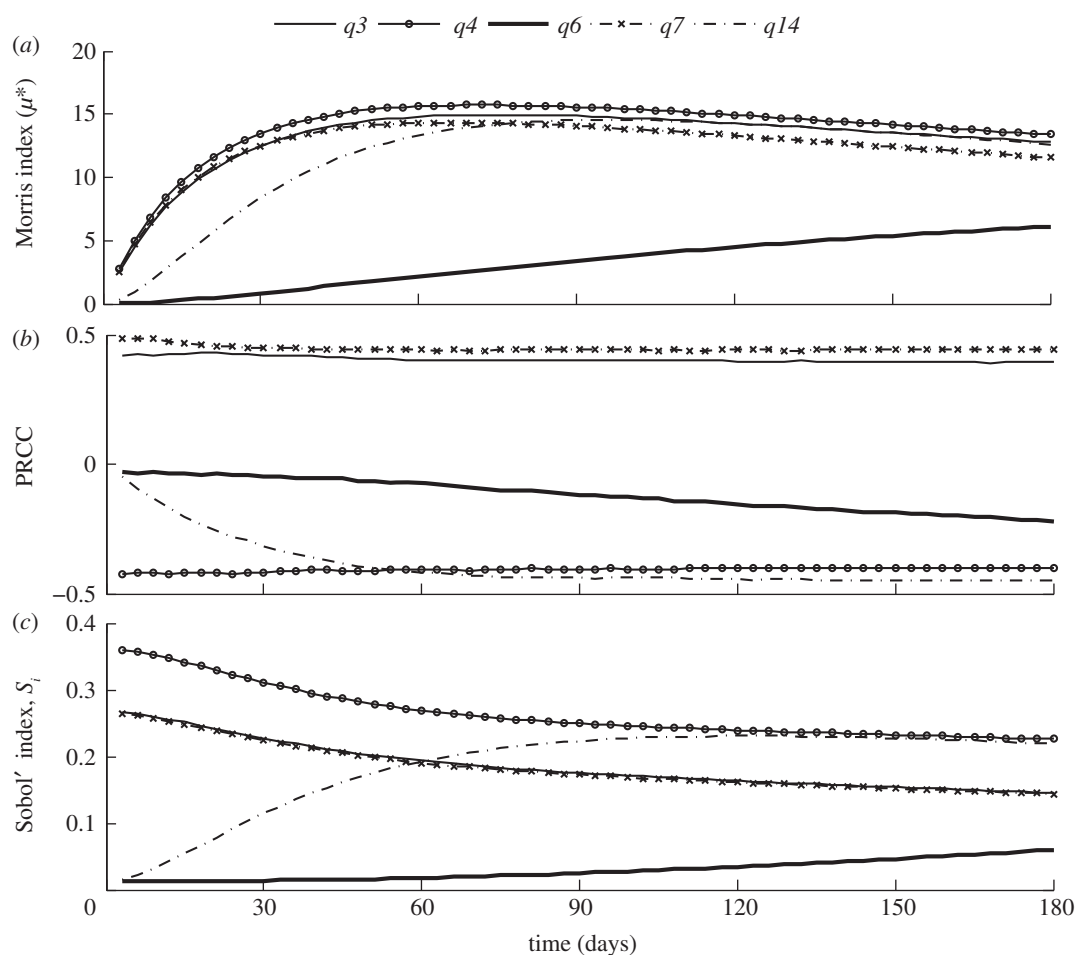


Figure 5. Temporal variation of the sensitivity of the worm burden (W) to key parameters in the schistosomiasis model as indicated by (a) the Morris index μ^* , (b) the PRCC index and (c) the Sobol' index, S_i . Sensitivity to all model parameters is shown in the electronic supplementary material, figure S3.

interest. Since the index for each parameter is constant, PSS avoids the complicated application of traditional SA methods at a single time point. When a model has only a few state variables and parameters, the strengths of SHM methods still hold as shown by SA of the schistosomiasis model with only two state variables. By comparison of figures 5 and 6, the same parameters were generally identified as important in the schistosomiasis model by both SHM methods and the traditional methods. However, the former more flexibly examines the sensitivity associated with different parameter or output subsets and provides more information about sensitivity than the latter. To our knowledge, the SHM method has not yet been applied to infectious disease models, and future work should consider the insights these powerful techniques can provide on both autonomous and non-autonomous systems [44].

In the development and evaluation of infectious disease models, our results can inform the selection of SA methods for aims such as model refinement, understanding model sensitivity over time, apportioning sensitivity across complex parameter spaces and intervention evaluation. If model refinement through improved parameter estimation is the research objective, PSS can be used to show the relative importance of each parameter over the period of simulation. The approach is particularly valuable if such model refinement is aimed at a time-specific goal (e.g. improved estimation of the peak population time). After the most critical parameters have been identified through these means, the feasibility of their refinement through additional data

collection or experimentation can subsequently be determined. A modeller interested in examining model sensitivity at one point in time to parameters and their interactions may choose the Sobol' method, whose chief advantage is the ability to tease out the model's response to a single parameter when all interactions are accounted for [45]. This apportioning of output sensitivity to a specific parameter may be useful when, for instance, multiple interventions are compared by examining the output's response to each altered parameter as separate first-order effects, and as total effects including the parameter's interactions [46]. On the other hand, a modeller interested in identifying when a parameter achieves maximum influence may benefit from the Morris method, which can provide information about the variation of sensitivity over the simulation interval. To our knowledge, the Morris method has not been applied in this fashion, and we hope to prompt its application in such infectious disease modelling contexts.

If a modeller aims to achieve specific goals defined as criteria on output variables, such as limiting the size of the infected population; the targeting of specific parameters that impact goal achievement are of interest, and PRCC can be particularly valuable [21] and has been applied to several infectious disease models in this fashion [26–28,32]. Other analyses have highlighted, through SA methods such as PRCC, the most effective times or parameters for efficient intervention [9,10], or the need for additional experimental data to aid research (e.g. estimates for host mixing rates [8]). Since PRCC indicates the degree of monotonicity between an output

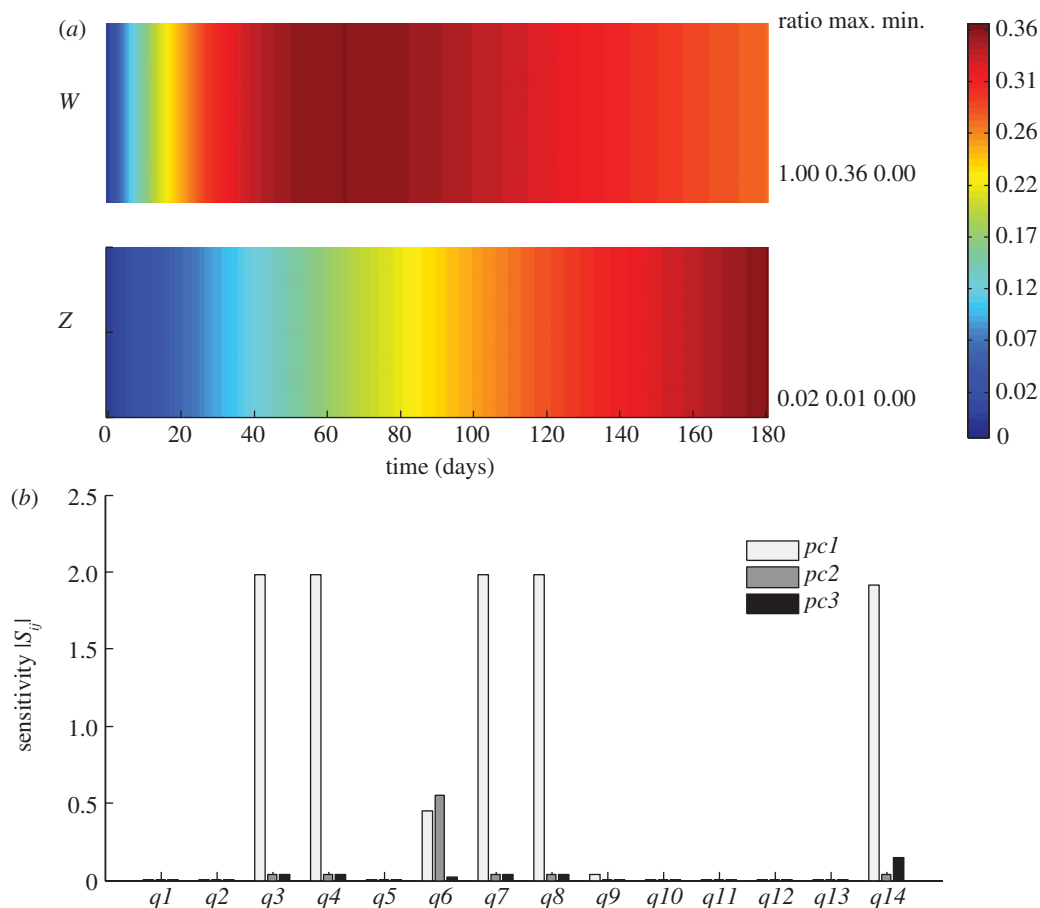


Figure 6. (a) SHM and (b) PSS of worm burden (W) and snail density (Z) for the schistosomiasis model, following previous work [11].

variable and a specific parameter in a model, preliminary examination the behaviour of the model is necessary [27]. Traditional scatter plots roughly screen the relationships between model outputs and parameters and may be used for a preliminary assessment. Otherwise, more powerful techniques, such as flow-based scatter plots, may be more suitable for the complexity typical of infectious disease models [47].

In resource poor settings, models, coupled with SA, may be used as a first step towards understanding control options. Whether by means of modelling behavioural interventions, biomedical interventions or through evaluation of social conditions and their impact on the disease system, infectious disease models have been used by decision-making bodies to explore and evaluate various policy options [48]. Where data are meager, SA can be used to identify key parameters upon which to focus limited data collection resources. Alternatively, an identified sensitive parameter may highlight a mechanism in the modelled system most suitable for cost-effective intervention. Prioritized parameters can also inform more efficient uncertainty analysis; assumed variation in only the most important parameters could be investigated to simulate the range of probable outcomes. GSA has also been used formally to aid in experimental design, to determine which parameters to focus upon when resources are scarce, and this has been achieved by analysing gradients of the model log-likelihood surface with respect to previously collected data [11]. This approach may lead to a fertile partnering of parameter estimation and SA methods, for model/data pairs with complex likelihoods.

Computation time is a final important point of comparison between SA methods. The SHM and PSS methods exhibit

highly optimized and efficient computation [11], while the computational requirements of traditional methods depend on the number of model runs, which itself is contingent on the number of parameters and how exhaustive the parameter space is sampled. The computational cost of the Sobol' method is the highest among the four traditional SA methods, and the more model runs undertaken, the more accurate the Sobol' index is [7,38].

We conclude by pointing out interesting insights that emerge from a comparison of the terms that appear in R_0 for the cholera model, and the relative importance of these and other parameters based on the SA findings. Examination of R_0 (equation (5.1)) produced by the next generation method [49] indicates that K_L , δ , γ , p , ξ_S , ξ_A , w and λ , (p_3 , p_5 , p_8 , p_{10} , p_{11} , p_{12} , p_{15}) are potentially influential terms.

$$R_0 = \left(\frac{1}{w}\right) \left[\frac{p\xi_A}{\mu + \gamma} + \frac{(\psi\theta + (1-\theta))\xi_S}{\mu_c + \mu + (1-\theta)\gamma + \theta\gamma\lambda} \right] \times \left[\frac{1}{\kappa_H\chi} + \frac{1}{\kappa_L\delta} \right]. \quad (5.1)$$

The SA results indicate that p_1 , p_8 , p_9 and p_{12} are important parameters and the remaining parameters have limited influence. Thus, SA has effectively narrowed the set of parameters worthy of additional scrutiny when compared with the qualitative conclusions a modeller might have reached based on a simple examination of R_0 . Of course, at times, R_0 is the model output of interest and is itself subjected to SA [9].

By comparing the application of five GSA methods to two infectious disease models, we have explored the strengths,

Table 3. Comparison of five GSA methods.

	scatter plots	the Morris method	LHS-PRCC	the Sobol' method	the SHM method
How easily understood is the concept behind the method?	easily	easily	medium	medium	difficult
Can the method rank the importance of parameters?	no	yes	yes	yes	yes
Can the method measure the variation of sensitivity over time?	no	yes	no	no	yes
Can the method show the sensitivity of the full model system?	no	no	no	no	yes
Can the method show the sensitivity to all parameters simultaneously?	no	no	no	no	yes
Can the method attribute the variance of a model output to each parameter?	no	no	no	yes	no
Can the method show the interaction effects on the sensitivity between parameters?	no	no	no	yes	no
Can the method show the direction of sensitivity? ^a	yes	no	yes	no	no
Is the method easy to carry out? (e.g. difficulty to calculate and present results)	easy	medium	slightly difficult	slightly difficult	easy ^b
What is the relative computation cost?	low	medium	medium	high	low

^aThe direction of sensitivity refers to whether the change in parameters and the change in model outputs are concordant or discordant.

^bExecution is greatly facilitated by the SASSY Matlab package.

weaknesses and appropriate uses of GSA methods in the context of infectious disease models, and our assessments are summarized in table 3. While we found the LHS-PRCC and the SHM method to be the most widely applicable to infectious disease models, all five SA methods explored here can support improvements in model performance and design, if the method selected is consonant with the modeller's goals. As the acquisition of data on dynamic disease transmission phenomena expands, and the demand for models to analyse and interpret these data increases, modellers need to quickly understand what the main 'levers' of influence are in models that potentially contain dozens or even hundreds of parameters. Through comparison of the research questions being addressed, the model's structure and the characteristics

of the SA methods investigated here, modellers can make informed SA choices that support their research aims and advance the productive application of mathematical models to infectious disease systems.

Acknowledgements. This manuscript was considerably strengthened by the critical remarks of four anonymous referees. Thanks, also, to Chloe Robbins and the rest of the Remais Lab. This work was supported in part by the Ecology of Infectious Disease program of the National Science Foundation under grant no. 0622743, by the Chemical, Bioengineering, Environmental, and Transport Systems Division of the National Science Foundation under grant no. 1249250, by the National Institute for Allergy and Infectious Disease (K01AI091864) and by the Global Health Institute at Emory University. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

References

1. Anderson RM, May RM. 1991 *Infectious diseases of humans: dynamics and control*. Oxford, UK: Oxford University Press.
2. Grassly NC, Fraser C. 2008 Mathematical models of infectious disease transmission. *Nat. Rev. Microbiol.* **6**, 477–487.
3. Keeling M, Rohani P. 2007 *Modeling infectious diseases in human and animals*. Princeton, NJ: Princeton University Press.
4. Hooker G, Ellner SP, Roditi LD, Earn DJD. 2011 Parameterizing state-space models for infectious disease dynamics by generalized profiling: measles in Ontario. *J. R. Soc. Interface* **8**, 961–974. (doi:10.1098/rsif.2010.0412)
5. Okais C, Roche S, Kurzinger ML, Riche B, Bricout H, Derrough T, Simondon F, Ecochard R. 2010 Methodology of the sensitivity analysis used for modeling an infectious disease. *Vaccine* **28**, 8132–8140. (doi:10.1016/j.vaccine.2010.09.099)
6. Moore JL, Liang S, Akullian A, Remais JV. 2012 Cautioning the use of degree-day models for climate change projections in the presence of parametric uncertainty. *Ecol. Appl.* **22**, 2237–2247. (doi:10.1890/12-0127.1)
7. Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, Saisana M, Tarantola S. 2008 *Global sensitivity analysis. The primer*. Chichester, UK: John Wiley & Sons.
8. Buhnerkempe MG, Eisen RJ, Goodell B, Gage KL, Antolin MF, Webb CT. 2011 Transmission shifts underlie variability in population responses to *Yersinia pestis* infection. *PLoS ONE* **6**, e22498. (doi:10.1371/journal.pone.0022498)
9. Chitnis N, Hyman JM, Cushing JM. 2008 Determining important parameters in the spread of malaria through the sensitivity analysis of a

- mathematical model. *Bull. Math. Biol.* **70**, 1272–1296. (doi:10.1007/s11538-008-9299-0)
10. McLeod RG, Brewster JF, Gumel AB, Slonowsky DA. 2006 Sensitivity and uncertainty analyses for a SARS model with time-varying inputs and outputs. *Math. Biosci. Eng.* **3**, 527–544. (doi:10.3934/mbe.2006.3.527)
 11. Rand DA. 2009 Mapping global sensitivity of cellular network dynamics: sensitivity heat maps and a global summation law. *J. R. Soc. Interface* **5**, S59–S69. (doi:10.1098/rsif.2008.0084.focus)
 12. Benton TG, Grant A. 1999 Elasticity analysis as an important tool in evolutionary and population ecology. *Trends Ecol. Evol.* **14**, 467–471. (doi:10.1016/S0169-5347(99)01724-3)
 13. Samuelson PA, Nordhaus WD. 2010 *Microeconomics*, 19th edn. Boston, MA: McGraw-Hill Irwin.
 14. Myers RH, Montgomery DC, Anderson-Cook CM. 2009 *Response surface methodology: process and product optimization using designed experiments*, 3rd edn. Hoboken, NJ: Wiley.
 15. Montgomery DC. 2013 *Design and analysis of experiments*, 8th edn. Hoboken, NJ: John Wiley & Sons, Inc.
 16. Sobol IM. 1990 On sensitivity estimates for nonlinear mathematical models. *Matematicheskoe Modelirovanie* **2**, 112–118. [In Russian.]
 17. Saltelli A, Tarantola S, Chan KPS. 1999 A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics* **41**, 39–56. (doi:10.1080/00401706.1999.10485594)
 18. Morris MD. 1991 Factorial sampling plans for preliminary computational experiments. *Technometrics* **33**, 161–174. (doi:10.1080/00401706.1991.10484804)
 19. Campolongo F, Cariboni J, Saltelli A. 2007 An effective screening design for sensitivity analysis of large models. *Environ. Model. Softw.* **22**, 1509–1518. (doi:10.1016/j.envsoft.2006.10.004)
 20. Helton JC, Davis FJ. 2002 Illustration of sampling-based methods for uncertainty and sensitivity analysis. *Risk Anal.* **22**, 591–622. (doi:10.1111/0272-4332.00041)
 21. Marino S, Hogue IB, Ray CJ, Kirschner DE. 2008 A methodology for performing global uncertainty and sensitivity analysis in systems biology. *J. Theor. Biol.* **254**, 178–196. (doi:10.1016/j.jtbi.2008.04.011)
 22. Hamby DM. 1994 A review of techniques for parameter sensitivity analysis of environmental-models. *Environ. Monit. Assess.* **32**, 135–154. (doi:10.1007/BF00547132)
 23. Yang J. 2011 Convergence and uncertainty analyses in Monte-Carlo based sensitivity analysis. *Environ. Model. Softw.* **26**, 444–457. (doi:10.1016/j.envsoft.2010.10.007)
 24. Andrews JR, Basu S. 2011 Transmission dynamics and control of cholera in Haiti: an epidemic model. *Lancet* **377**, 1248–1255. (doi:10.1016/S0140-6736(11)60273-0)
 25. Zhang J, Jin Z, Sun G, Zhou T, Ruan S. 2011 Analysis of Rabies in China: transmission dynamics and control. *PLoS ONE* **6**, e20891. (doi:10.1371/journal.pone.0020891)
 26. Gubbins S, Carpenter S, Baylis M, Wood JLN, Mellor PS. 2008 Assessing the risk of bluetongue to UK livestock: uncertainty and sensitivity analyses of a temperature-dependent model for the basic reproduction number. *J. R. Soc. Interface* **5**, 363–371. (doi:10.1098/rsif.2007.1110)
 27. Blower SM, Dowlatabadi H. 1994 Sensitivity and uncertainty analysis of complex-models of disease transmission—an HIV model, as an example. *Int. Stat. Rev.* **62**, 229–243. (doi:10.2307/1403510)
 28. Porco TC, Blower SM. 1998 Quantifying the intrinsic transmission dynamics of tuberculosis. *Theor. Popul. Biol.* **54**, 117–132. (doi:10.1006/tpbi.1998.1366)
 29. Ellis AM, Garcia AJ, Focks DA, Morrison AC, Scott TW. 2011 Parameterization and sensitivity analysis of a complex simulation model for mosquito population dynamics, dengue transmission, and their control. *Am. J. Trop. Med. Hyg.* **85**, 257–264. (doi:10.4269/ajtmh.2011.10-0516)
 30. Bailey NTJ, Duppenhaller J. 1980 Sensitivity analysis in the modeling of infectious-disease dynamics. *J. Math. Biol.* **10**, 113–131. (doi:10.1007/BF00275837)
 31. Hethcote HW, Vanark JW, Karon JM. 1991 A simulation-model of AIDS in San Francisco. 2. Simulations, therapy, and sensitivity analysis. *Math. Biosci.* **106**, 223–247. (doi:10.1016/0025-5564(91)90078-W)
 32. Legrand J, Sanchez A, Le Pont F, Camacho L, Larouze B. 2008 Modeling the impact of tuberculosis control strategies in highly endemic overcrowded prisons. *PLoS ONE* **3**, e2100. (doi:10.1371/journal.pone.0002100)
 33. Saltelli A, Tarantola S, Campolongo F, Ratto M. 2004 *Sensitivity analysis in practice: a guide to assessing scientific models*. Chichester, UK: John Wiley & Sons.
 34. Hamby DM. 1995 A comparison of sensitivity analysis techniques. *Health Phys.* **68**, 195–204. (doi:10.1097/00004032-199502000-00005)
 35. McKay MD, Beckman RJ, Conover WJ. 1979 A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* **21**, 239–245.
 36. Helton JC, Davis FJ. 2003 Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliab. Eng. Syst. Safety* **81**, 23–69. (doi:10.1016/S0951-8320(03)00058-9)
 37. Sobol IM. 2001 Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math. Comp. Simul.* **55**, 271–280. (doi:10.1016/S0378-4754(00)00270-6)
 38. Saltelli A, Annoni P, Azzini I, Campolongo F, Ratto M, Tarantola S. 2010 Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Comp. Phys. Commun.* **181**, 259–270. (doi:10.1016/j.cpc.2009.09.018)
 39. Domijan M, Rand DA. 2011 Balance equations can buffer noisy and sustained environmental perturbations of circadian clocks. *Interface Focus* **1**, 177–186. (doi:10.1098/rsfs.2010.0007)
 40. Liang S, Maszle D, Spear RC. 2002 A quantitative framework for a multi-group model of *Schistosomiasis japonicum* transmission dynamics and control in Sichuan, China. *Acta Trop.* **82**, 263–277. (doi:10.1016/S0001-706X(02)00018-9)
 41. Liang S *et al.* 2007 Environmental effects on parasitic disease transmission exemplified by schistosomiasis in western China. *Proc. Natl Acad. Sci. USA* **104**, 7110–7115. (doi:10.1073/pnas.0701878104)
 42. Brown P, Rand DA. 2012 Using the Global Sensitivity GUIs. See <http://wsbcwarwick.acuk/software/indexhtml>.
 43. Segovia-Juarez JL, Ganguli S, Kirschner D. 2004 Identifying control mechanisms of granuloma formation during M-tuberculosis infection using an agent-based model. *J. Theor. Biol.* **231**, 357–376. (doi:10.1016/j.jtbi.2004.06.031)
 44. Remais J, Zhong B, Carlton EJ, Spear RC. 2009 Model approaches for estimating the influence of time-varying socio-environmental factors on macroparasite transmission in two endemic regions. *Epidemics* **1**, 213–220. (doi:10.1016/j.epidem.2009.10.001)
 45. Mokhtari A, Frey HC. 2005 Sensitivity analysis of a two-dimensional probabilistic risk assessment model using analysis of variance. *Risk Anal.* **25**, 1511–1529. (doi:10.1111/j.1539-6924.2005.00679.x)
 46. Busschaert P, Geeraerd AH, Uyttendaele M, Van Impe JF. 2011 Sensitivity analysis of a two-dimensional quantitative microbiological risk assessment: keeping variability and uncertainty separated. *Risk Anal.* **31**, 1295–1307. (doi:10.1111/j.1539-6924.2011.01592.x)
 47. Chan YH, Correa CD, Ma KL (eds) 2010 Flow-based Scatterplots for Sensitivity Analysis. In *IEEE Symp. on Visual Analytics Science and Technology Salt Lake City, UT, 25–26 October 2010*. New York, NY: IEEE.
 48. Griffin JT *et al.* 2010 Reducing *Plasmodium falciparum* malaria transmission in Africa: a model-based evaluation of intervention strategies. *PLoS Med.* **7**, e1000324. (doi:10.1371/journal.pmed.1000324)
 49. van den Driessche P, Watmough J. 2002 Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math. Biosci.* **180**, 29–48. (doi:10.1016/S0025-5564(02)00108-6)