Model Calibration in Network Models of HIV

Sally K. Slipher
Montana State University – Bozeman
Bozeman, Montana, USA
sallyslipher@montana.edu

Nicole Bohme Carnegie
Montana State University – Bozeman
Bozeman, Montana, USA
nicole.carnegie@montana.edu

ABSTRACT

Epidemiologic studies of infectious disease often involve capturing and modeling complex data of how people interact to contribute to transmission. Network models answer questions about the course of an epidemic within a population of interest, providing a realistic reproduction of the connections between contacts. The simulation of an epidemic over time provides estimates of infection, and by performing simulations with varying ranges of parameter values, parameters that contain uncertainty due to bias or lack of data can be calibrated. Latin hypercube sampling randomly draws parameter values in a way that distributes samples evenly over the space. Approximate Bayesian computation converges to subregions of the parameter space. We explore parameter calibration in network models using these two search strategies with a network model of HIV transmission.

KEYWORDS

HIV, agent-based models, epidemic network models, parameter search strategies

ACM Reference Format:
Sally K. Slipher and Nicole Bohme Carnegie. 2021. Model Calibration in Network Models of HIV. In *epiDAMIK 2021: 4th epiDAMIK ACM SIGKDD International Workshop on Epidemiology meets Data Mining and Knowledge Discovery*. ACM, New York, NY, USA, 5 pages. https://doi.org/xx.xxxx/xxxxxxxxx.xxxxxxx

1 INTRODUCTION

In this ongoing project, we investigate calibration of stochastic, agent-based network models of disease spread, specifically in the context of modeling the spread of HIV over sexual networks. The computational burden of this type of model restricts the feasibility of traditional methods of calibration using other deterministic or stochastic models. We compare a standard method to search a multi-dimensional parameter space, Latin hypercube sampling (LHS) [1, 2], with approximate Bayesian computation (ABC), a method that is commonly used to find a single set of calibrated parameter values [5, 10]. We will compare their performance calibrating a model of the HIV epidemic among men who have sex with men (MSM) in San Diego, California [8].

epiDAMIK 2021, Aug 15, 2021, Virtual

ACM ISBN 978-1-xxxx-XXXX-X…$15.00
https://doi.org/xx.xxxx/xxxxxxxxx.xxxxxxx

Network models of disease transmission simulate spread over a network of contacts that evolves over time. This framework has two advantages in modeling sexual transmission: contacts can persist across time points, and it is straightforward to incorporate highly heterogeneous risk behaviors, demographics, treatment use, etc. The drawback of allowing significant heterogeneity is the necessity of estimating inputs for a very large number of parameters.

Even when estimation of inputs is achievable, the possibility of bias in those estimates is substantial. Self-reported data are prone to bias, especially when dealing with stigmatized diseases and sensitive behaviors, as is the case with HIV. Moreover, there are typically strong selection effects, as we are modeling hard-to-reach populations and are unlikely to have representative samples from those populations [6, 7].

Model parameters are usually informed by experimental or observational data, analogous systems, statistical inference, expert opinion [11], biological constraints, variation in repeated measurements, or comparable data in scientific literature. When this information is lacking or not available, a range of reasonable values or a best guess can be chosen as initial inputs for a model [9].

Calibration of parameters is critical to obtain a model that better reflects reality and thus can produce reliable predictions and inferences about disease spread. This need is especially acute when unknown or uncertain parameters are influential to transmission and drive spread.

If input parameters control independent processes, they can be calibrated independently, and full exploration of the multi-dimensional parameter space is unnecessary. In network-based models, however, this is unlikely to be the case, and parameters have complex, nonlinear relationships with each other and model outputs. Hence, calibration must simultaneous for all parameters. This is a computationally intensive task, with the number of permutations to consider rapidly increasing for even a moderate number of parameters [1, 9].

Simulation-based calibration methods can be used to estimate likely values of outputs from complex, stochastic agent-based models. These methods use parameter-search strategies to find the joint distribution of model outputs and input parameters algorithmically, either by approximating the likelihood or by choosing combinations of parameters that produce desired network or epidemic metrics in simulated epidemics [3].

2 METHODS

We will apply two search algorithms to calibration of a large stochastic network-based model of HIV transmission among MSM in San Diego. In this section, we first present the two methods, Latin hypercube sampling and approximate Bayesian computation algorithms, then describe the parameterization of the epidemic model, and finally give the specifics of the simulation study.

2.1 Search Strategies

2.1.1 Latin Hypercube Sampling. One standard parameter-search strategy is Latin hypercube sampling (LHS). It provides a near-random sample of parameter values from a multidimensional distribution. This is accomplished by subdividing the parameter space into equally sized partitions and randomly sampling from each partition.

Parameters are not likely to be independent of each other, and their relationship with epidemic outcomes may be nonlinear. So, for multi-parameter searches, parameters must be sampled jointly. In addition, other search strategies may converge on a region of the parameter space that misses maxima in multimodal distributions. LHS handles both these issues by generating dependent samples in which each sample is random but does not occupy the same region of the parameter space, distributing samples more evenly over the space. This is an improvement over pure random sampling, which may produce samples that cluster in some areas of the parameter space (wasting computational resources), or highly dispersed samples in some regions (potentially missing features of the distribution) [1].

LHS is a stratified, Monte Carlo, sampling-based method. It proceeds by dividing the range of the n-dimensional parameter space into equally-probable intervals for each of n parameters being considered. The size of each interval is determined by the probability distribution chosen for each parameter space, and the use of uniform distributions results in intervals of equal size [9]. The number of intervals, i.e., the resolution of the grid of intervals, is determined by how many samples are desired. A parameter value is randomly drawn from each interval such that only one sample point occupies each row and column (for two-dimensional space), or each axis-aligned hyperplane (for an arbitrary number of dimensions). Samples are drawn sequentially, with new sample points taking into account the position of the previous sample points. The set of parameter values from a Latin hypercube sample is then used as input parameters and run in model simulation, and model outputs are obtained and assessed against target outputs.

2.1.2 Approximate Bayesian Computation. Approximate Bayesian computation (ABC) is another sampling-based method for calibrating parameters. It uses a rejection algorithm to sample parameter values from prior distributions of reasonable values. This approximates posterior distributions of the parameters by simulating from the model. Model outputs from the simulations under various combinations of parameters are compared with observed data, and the values of parameters that produce adequately-fit simulations are kept.

Like LHS, ABC methods can be used to estimate model parameters jointly to reveal dependencies between them. ABC assesses different models that are parameterized from empirical data as well as knowledge of how likely certain values of inputs might be [10]. Simulation results are compared to data through summary statistics in order to reduce the dimensionality of the complex results of the model, and only parameter values resulting in simulations within a given tolerance threshold are retained. ABC converges to subregions of the parameter space instead of randomly sampling the entire space [2], so it can be much more efficient in its search.

Parameter	Base Value	Parameter Space
Daily act rate	0.096(M), 0.136(C)	[0.5,1.5] multiplier
Initial prevalence	0.20	[0.10,0.30]
Mean testing interval (days)	198(B), 259(H), 234(O)	[0.5,1.5] multiplier
Days since last negative test	198(B), 259(H), 234(O)	[0.5,1.5] multiplier

Table 1: Parameters to estimate in the model calibration process. The initial model uses multipliers of 1 and an initial prevalence of 0.20. Where needed, rates are given by partnership type (main, casual) and by race (Black, Hispanic, Other).

This method is likelihood-free, so it can perform estimation when models are too complex to use standard likelihood techniques, as is the case here. This is achieved through approximating the likelihood functions by running simulations many times [9]. ABC methods not only give point estimates of parameter values but uncertainty measures, since the full posterior distribution is approximated. This uncertainty can then be propagated into the calibrated model's predictions [10].

ABC rejection sampling first samples plausible parameter values from prior distributions. A dataset is then simulated under the model using the sampled parameters as inputs. Target summary statistics that quantify defining aspects of the observed data are compared to summary statistics calculated from the simulated model. The parameter values are accepted if the summary measures match those of the data closely enough, within a distance smaller than a set threshold. This is repeated for many sets of parameters and their simulations. The accepted parameter values approximate the posterior distribution given observed data [4, 5].

2.2 Parameters

In the HIV model we are using for this study, there are over 150 input parameters. These parameters cover aspects of demographics (e.g., race, age), relationships (e.g., number of partnerships a person has, whether casual or long term, how often acts of sexual intercourse occur), and HIV transmission and progression (e.g., how often an individual is tested or treated for HIV, how long a person spends in different stages of the disease, viral load at each stage). Certain input parameters are well-supported by data, like age and race, or are well-characterized in the literature, like viral load. A number of parameters, however, have a potentially large impact on epidemic dynamics, but are difficult to pinpoint precisely.

For this investigation, we chose four parameters to calibrate: the initial prevalence of HIV in the population, the rate of how often sexual intercourse acts occurs, the average interval between HIV tests, and the average time since a person's last negative HIV test. The time since last negative test is uncertain but is not expected to strongly impact epidemic trajectory, and it is included as a check. The estimated value of this should not depend significantly on the values of other parameters. These chosen parameters are of interest since they contain bias and uncertainty, and most importantly, they are expected to greatly influence HIV prevalence over time (expect for time since last negative test).

The range of plausible values for each parameter of interest is given in Table 1. The simulated outputs from the epidemic model that will be used for evaluation are the final average prevalence of

HIV and the trend in prevalence, both calculated over the last 10 years of the simulation. We target 20% prevalence and a trend of 0 (stable prevalence) over this time period.

2.3 Simulation

Epidemic simulations for both LHS and ABC are run for 70 years in one-week time steps. The first 60 years are discarded as burn-in, and the average prevalence and slope of prevalence over the last 10 years are computed to evaluate model acceptability.

For the Latin hypercube sampling algorithm, input values for multipliers on act rate, testing interval, and last negative test were drawn sequentially from ranges of 0.5 to 1.5 and initial prevalence from a range of 0.10 to 0.30. A sample size of 1000 ($n = 1000$) was drawn for these four parameters ($k = 4$), so the joint parameter space is in a 4-dimensional space of real numbers ($\mathbb{R}^4$). Figure 1 shows a visualization of the 1000 samples of the three parameters that are expected to influence model outputs in the joint space. The sampled parameters are input into the simulation along with the other model parameters, and the average prevalence and trend are computed as above. Results within 1 percentage point from the target prevalence of 20% and within 0.00005 from the target trend of 0 are considered acceptable and the corresponding parameter combinations are retained.

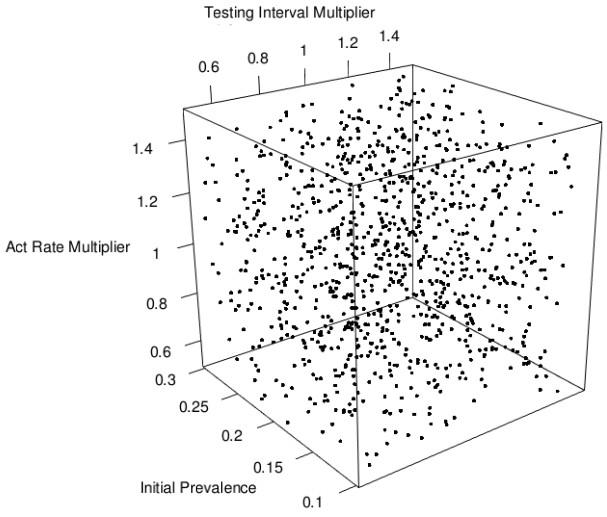

Figure 1: 1000 Latin hypercube samples with 3 parameters: act rate multiplier, testing interval multiplier, and initial prevalence. The parameter space for the multiplier on the last negative test interval is not shown.

For the alternative sampling method, approximate Bayesian computation, we first set marginal prior distributions for all four parameters of interest. We chose non-informative priors using uniform distributions from the minimum to maximum of the ranges of plausible values given in Table 1. The algorithm first draws from the marginal priors, those values are input to a simulation, and the average prevalence and trend are computed. If these are not within 10% of the targets of 0.20 and 0 (respectively), then the set of parameters is rejected. Parameters thereafter are sampled in a biased fashion from an area of the parameter space where simulations are frequently close to the targets. This process is repeated for narrower tolerances, drawing from the posterior distribution of the previous stage. The process of drawing parameters, simulating the epidemic, and rejecting simulations outside of the threshold is repeated until 100 accepted simulations are obtained, approximating the joint posterior distribution of model parameters. A relatively large threshold is used initially to find the target region, and the threshold becomes gradually reduced in the sequential stages of the procedure. This results in comparable thresholds to those used in LHS parameter selection.

3 RESULTS & DISCUSSION

3.1 LHS

The resulting accepted parameter combinations from the 1000 LHS samples are given in Table 2. These combinations produce a simulated epidemic over 70 years in which the last 10 years have an average prevalence of 19%-21%, and the slope of the prevalence is -0.0005 to 0.00005 % per year. These are relatively broad targets, since available data leaves much unknown about the true epidemic trajectory, though more specific or additional targets may be chosen to produce a smaller set of acceptable parameter values.

The act rate multiplier is consistently estimated to be near 0.70-0.74, below the base value multiplier of 1. The estimated initial prevalence is consistently high (26% to 30%) compared to the base value of 20%. One parameter combination shows the reverse pattern of these estimates, where act rate is high (0.78) and initial prevalence is low (20%), as seen in row 3 of Table 2. The estimated multiplier on testing interval does not show much of a pattern, with either a substantial increase (multiplier >1) or decrease (multiplier <1) from the base value to achieve acceptable epidemic trajectories. As expected, the estimated values for multiplier on last negative test interval range across almost the entire parameter space and show no clear patterns.

Figure 2 shows the accepted marginal values for each parameter relative to the prevalence and trend produced by all simulations. Accepted values are highlighted in blue. The point estimates for the multiplier on act rate are very consistent, implying that act rate is very important in simulating prevalence, and that it is a major driver of epidemic transmission. Initial prevalence is also relatively consistent, and its value drives the epidemic's trajectory, though to less of an extent than act rate. These two parameters show a dependent relationship, where combinations with higher act rate tend to be paired with lower initial prevalence (and the reverse). The other parameters, however, are spread relatively evenly across the space within the region of the target outcomes, indicating that their value is jointly-dependent on the others and not a major driver of epidemic trajectory.

The prevalence curves of the epidemics produced by the 1000 LHS samples are plotted in Figure 3. The simulated epidemic trajectories with parameter combinations that produce acceptable target outcomes are highlighted in green. It is of note that that there are clearly unacceptable parameter combinations that produce prevalence curves very different than what is observed in the data. The LHS sampling algorithm is illuminating in that, when parameters

Act Rate Multiplier	Initial Prevalence	Testing Interval Multiplier	Last Negative Test Interval Multiplier	Prevalence	Trend
0.721	0.296	0.790	0.697	0.202	0.00005
0.702	0.272	1.072	1.222	0.196	0.00003
0.784	0.200	0.580	1.252	0.198	0.00004
0.710	0.283	1.106	0.653	0.190	0.00005
0.733	0.276	0.800	0.594	0.199	0.00004
0.735	0.259	0.887	1.349	0.204	0.00004
0.700	0.299	1.243	1.073	0.193	0.00004
0.714	0.298	0.985	0.551	0.206	0.00004

Table 2: Accepted parameter combinations from LHS simulations that produce epidemic simulation outputs within 1 percentage point of prevalence target (20%) and 0.00005 of trend target (0).

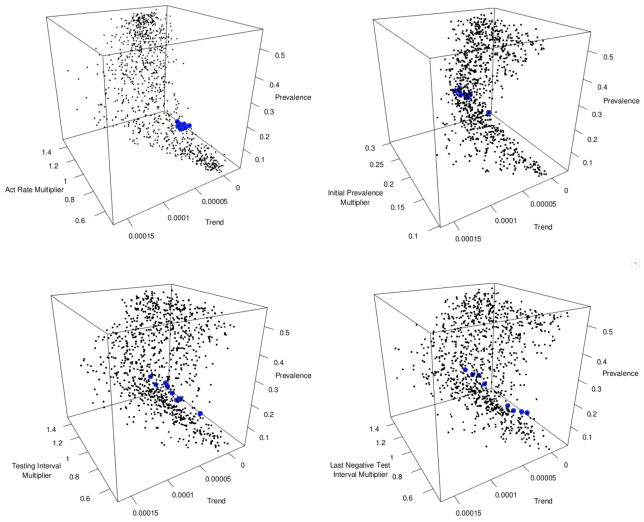

Figure 2: 3-D plots of simulation output values of prevalence and trend given LHS parameter combinations of act rate multiplier, initial prevalence, testing interval multiplier, and last negative test interval multiplier. Plots show marginal values of each parameter (y-axis) versus corresponding prevalence (z-axis) and trend (x-axis) outputs. Blue points are marginal parameter values that produced target model outputs within the specified thresholds.

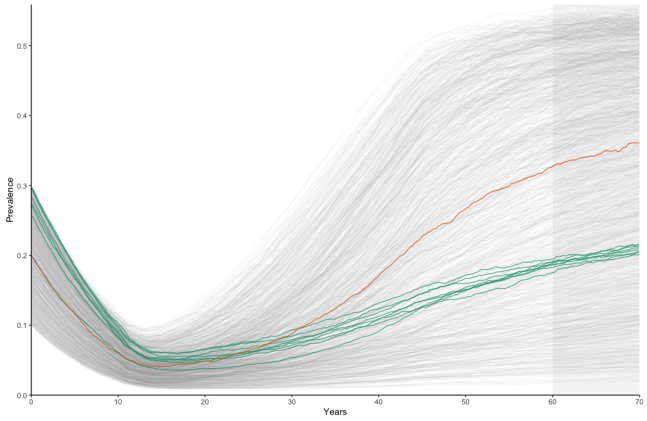

Figure 3: 1000 Simulated epidemics over 70 years using LHS samples of parameter values. Gray lines are simulated prevalence over time that did not meet target prevalence and trend thresholds. Green lines show accepted simulations. The red line is the initial simulation using base parameter values. The gray panel highlights the last 10 years, used for computing targets.

are unknown, there are multiple acceptable combinations of parameters that lead to model calibration. This is in contrast to one ideal combination (and its approximate distribution), chosen by approximate Bayesian computation.

3.2 ABC

Simulation using approximate Bayesian computation are in process. This algorithm produces posterior distributions of parameter values. Marginal distributions should center on values of the parameters that produce acceptable epidemic trajectories. Thus, different statistics (e.g., minimum, mean, median) of the posterior distributions may be used for model inputs. We will evaluate whether ABC misses global or local maxima that are captured by LHS.

4 CONCLUSION

In summary, epidemic network models are powerful, detailed models that capture dynamic and heterogeneous individual behaviors which contribute to disease spread. Since these models aim to reflect the reality of a specific disease's transmission, model parameters with uncertainty, variability, or bias need to be calibrated make meaningful inference from a model. Parameter estimation with Latin hypercube sampling is more dispersed and better covers the entire parameter space, while approximate Bayesian inference creates a focused distribution of values and is more computationally efficient.

ACKNOWLEDGMENTS

Research reported in this publication was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under award number R01AI147441. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

REFERENCES

[1] Collins, A. J., Seiler, M. J., Gangel, M., and Croll, M. Applying latin hypercube sampling to agent-based models: Understanding foreclosure contagion effects. *International Journal of Housing Markets and Analysis 6*, 4 (2013), 422–437.

[2] Donovan, D., Burrage, K., Burrage, P., McCourt, T., Thompson, B., and Yazici, E. Estimates of the coverage of parameter space by latin hypercube and orthogonal array-based sampling. *Applied Mathematical Modelling 57* (2018), 553–564.

[3] Hazelbag, C. M., Dushoff, J., Dominic, E. M., Mthombothi, Z. E., and Delva, W. Calibration of individual-based models to epidemiological data: A systematic review. *PLoS computational biology 16*, 5 (2020), e1007893.

[4] Heard, D. Statistical inference using agent-based models. https://www.samsi.info/wp-content/uploads/2016/03/Heard_may7_2014.pdf, 2014. Accessed: 2020-11-21.

[5] Jabot, F., Faure, T., Dumoulin, N., and Albert., C. *EasyABC: Efficient Approximate Bayesian Computation Sampling Schemes*, 2015.

[6] Keeling, M. J., and Eames, K. T. Networks and epidemic models. *Journal of the Royal Society Interface 2*, 4 (2005), 295–307.

[7] Lloyd, A. L., and Valeika, S. Network models in epidemiology: an overview. In *Complex population dynamics: nonlinear modeling in ecology, epidemiology and genetics*. World Scientific, 2007, pp. 189–214.

[8] Panichsillapakit, T., Smith, D. M., Wertheim, J. O., Richman, D. D., Little, S. J., and Mehta, S. R. Prevalence of transmitted hiv drug resistance among recently infected persons in san diego, california 1996-2013. *Journal of acquired immune deficiency syndromes (1999) 71*, 2 (2016), 228.

[9] Thiele, J. C., Kurth, W., and Grimm, V. Facilitating parameter estimation and sensitivity analysis of agent-based models: A cookbook using netlogo and r. *Journal of Artificial Societies and Social Simulation 17*, 3 (2014), 11.

[10] van der Vaart, E., Beaumont, M. A., Johnston, A. S., and Sibly, R. M. Calibration and evaluation of individual-based models using approximate bayesian computation. *Ecological Modelling 312* (2015), 182–190.

[11] Wu, J., Dhingra, R., Gambhir, M., and Remais, J. V. Sensitivity analysis of infectious disease models: methods, advances and their application. *Journal of The Royal Society Interface 10*, 86 (2013), 20121018.