

Quantifying Uncertainty in Epidemiological Models

Sumit K. Jha
Electrical Engineering and Computer
Science Department
University of Central Florida, Orlando,
Florida
Email: jha@eecs.ucf.edu

Arvind Ramanathan
Computer Science and Engineering
Division
Oak Ridge National Laboratory,
Oak Ridge, Tennessee 37830
Email: ramanathana@ornl.gov

ABSTRACT

Modern epidemiology has made use of a number of mathematical models, including ordinary differential equation (ODE) based models and agent based models (ABMs) to describe the dynamics of how a disease may spread within a population and enable the rational design of strategies for intervention that effectively contain the spread of the disease. Although such predictions are of fundamental importance in preventing the next global pandemic, there is a significant gap in trusting the outcomes/predictions solely based on such models. Hence, there is a need to develop approaches such that mathematical models can be calibrated against historical data. In addition, there is a need to develop rigorous uncertainty quantification approaches that can provide insights into when a model will fail and characterize the confidence in the (possibly multiple) model outcomes/predictions, when such retrospective analysis cannot be performed. In this paper, we outline an approach to develop uncertainty quantification approaches for epidemiological models using formal methods and model checking. By specifying the outcomes expected from a model in a suitable spatio-temporal logic, we use probabilistic model checking methods to quantify the probability with which the epidemiological model satisfies a given behavioral specification. We argue that statistical model checking methods can solve the uncertainty quantification problem for complex epidemiological models.

1. INTRODUCTION

Global pandemics threaten the very survival of our human race: the 1918 flu epidemic which resulted in nearly 50-100 million deaths [30], and the more recent SARS epidemics in 2003 [5] and the H1N1 flu epidemic in 2009 serve as vital reminders of how contagious diseases can alter the world demographic in a relatively short time. The spread of an epidemic amidst a large population can have disastrous effects on a country and the world as a whole; leading to widespread social, economic and political turmoil. Therefore, in an effort to bolster public health security, considerable efforts are underway to obtain accurate insights into the mechanisms that govern the dynamics of the spread of diseases.

Clinical research targeting infectious diseases remains at the fore-

front of developing novel approaches to thwart such global threats. With advances in whole genome sequencing, drug discovery techniques, data management practices, rapid diagnostics, we have been able to develop effective intervention systems through which rapid progress of infections can be successfully contained/controlled. However, with the world becoming increasingly connected and growing threats of biological/chemical warfare, there is an ever-increasing need to develop predictive models of how epidemics may spread and what strategies can be effective in combating such emerging pandemics.

Mathematical/computational modeling of infectious diseases is an integral part of modern epidemiology [16]. Mathematical models play a vital role in understanding scenarios of disease spread, as well as testing specific hypotheses of disease control based on clinical data. Integrated with rigorous statistical analyses tools, these models continue to provide relevant insights into patterns and dynamics of disease spread along with the efficacies of different intervention strategies. Although predictions from models are central to epidemiological studies, *there is a significant knowledge gap in trusting the (possibly multiple) outcomes/predictions of these models*. In particular, it has become particularly difficult to (a) identify, characterize and quantify the failure conditions for these models; (b) automatically evaluate the behavioral landscape of these models, based on the number of parameters involved; and, (c) characterize the confidence in model outcomes for a given scenario by using formal metrics of uncertainty, such as probability measures associated with different outcomes.

In this paper, we present an overview of approaches that are particularly well suited to develop a verification and validation framework for epidemiological models. In particular, we focus on developing robust epidemiological models that take into account the parameter space and behavioral landscape that allow one to quantify uncertainty in these models. Traditional approaches have relied on the use of historical data (from clinical studies and other sources) to validate epidemiological models and only more recently sensitivity analysis techniques have become popular [13]. However, retrospective data analysis for several epidemics may not be available (especially, in the context of emerging diseases such as H1N1 or H5N1). Hence, there is a need to develop integrated tools that allow for validation and verification of epidemiological models using qualitative expert insights (in the absence of such retrospective data) as well as the ability to automatically determine the uncertainty in the predictions made by these models. We first outline the mathematical approaches used in epidemiology and then present an overview of computational tools that can be used to quantify the uncertainty in these models, especially in the context where historical/retrospective data analysis cannot be performed.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

VVEPIDEMIOLOGY '12 Washington, DC USA

Copyright 2012 ACM X-XXXXXX-XX-X/XX/XX ...\$15.00.

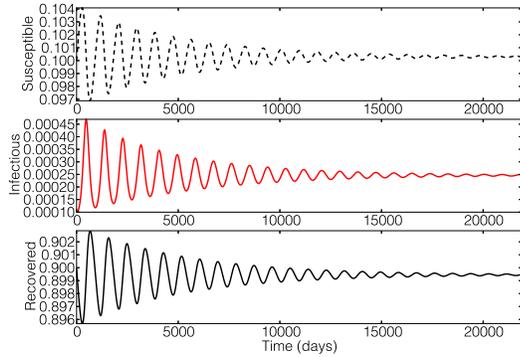


Figure 1: Numerical simulation results from a SIR model with demography. Note that even the simple model defined in Model 1 leads to an oscillatory behavior, with the disease cycling ~ 1200 days. The values of μ , β and γ were set to 3.91×10^{-5} , 1.4246 and 0.1428 respectively.

Ordinary Differential Equation (ODE) based models

In the simplest possible models, disease spread dynamics is described by segregating a population into different categories (or *compartments*) and studying the disease spread as a dynamical system of these interacting compartments. These compartments describe if the population was (a) previously unexposed to the pathogen, referred to as susceptible (S); (b) infected by the pathogen (I); (c) infected by the pathogen, but are not infectious yet, referred to as exposed (E) and (d) recovered from the infection (R). The dynamics of the interacting compartments was encoded using ODEs [25]. For example, for most infectious diseases such as flu and cholera, a commonly used model, referred to as the SIR model, can be used to describe the disease spread dynamics. The SIR model consists of an ODE system described as follows:

$$\frac{dS}{dt} = \mu - \beta SI - \mu S; \quad (1)$$

$$\frac{dI}{dt} = \beta SI - \gamma I - \mu I; \quad (2)$$

$$\frac{dR}{dt} = \gamma I - \mu R. \quad (3)$$

As shown above, the constant μ is referred to as the mortality rate, which represents the rate at which the demography is altered by a simple measure of birth/death in the population. The constant γ refers to the rate at which patients recover (it is the reciprocal of the mean time of infection, which is easily accessible from clinical data) and β refers to the rate of infection. Depending on β and γ , the sub-groups within a population can ‘exchange’, leading to the dynamics in disease spread and containment. An analytical solution to this simple system of equations is difficult to obtain; however, numerical simulations can indicate how the different sub-groups (S, I, R) would evolve with respect to time. As illustrated in figure 1, even this simple model possesses an ability to show oscillatory behavior, which is an important consequence of having introduced mortality rates into the population model.

The success of these deceptively simple dynamic models in interpreting infectious disease dynamics in the case of flu, rubella and chicken pox, led to the development of a variety of compartmental models including the SIR, SEIR, SIS and SII models [25]. These models have provided an effective computational infrastructure to describe several aspects of disease spread that are of interest to epi-

demologists and public health officials. These models can also predict when an epidemic will arise, how long the epidemic may last, how many secondary infections may occur and under what conditions the epidemic may cycle (recur). ODE models are also relatively flexible, in that they can incorporate a variety of scenarios, including host-heterogeneity (for e.g., different age-groups and demographics), multiple pathogens/host models (for e.g., vector-borne diseases, hosts with partial immunity, etc.), stochastic models of disease spread and spatial distribution of population.

ODE models of disease spread work under the assumption that populations are homogeneous and sufficiently large. While they are suitable to capture an average and collective behavior of disease spread, they cannot predict rare events or anomalies that may arise out of individual behavior/ activities. More generally, ODE based models are restricted in describing the nature of (a) heterogeneity in the population as well as geographic regions; (b) complex interactions between individuals; (c) activities of individual; (d) account for travel patterns of individuals, (e) the inherent randomness in the infection and recovery process, and (e) the evolution of pathogens, such as antibiotic resistance. These modeling limitations have allowed researchers to examine alternative strategies, primarily focused on agent based models (ABMs) to represent individual behaviors and interaction patterns.

These models have traditionally relied on the use of clinical data to interpret, predict and manage infectious disease spread dynamics. Although such data is difficult to obtain, many studies have shown that when such data is made available, one can indeed validate and use these models to predict the behavior of infectious diseases. It must however be noted that very few approaches in the literature have used formal methods to verify and validate ODE based disease spread models [7, 9, 19]. These approaches rely on extending current verification tools for biochemical networks to formally capture the behavioral properties of epidemiological models [7, 6, 9]. However, it must be noted that unlike a biochemical network, where compartments can be created out of different chemical species in the system, epidemiological models deal with more complex interactions and behaviors, which may not be sufficiently captured by such tools alone. Further, the nature of epidemiological studies requires the use of probabilistic correctness specifications with both spatial and temporal components. Traditional formal specification frameworks have mostly focussed on probabilistic temporal logics.

Agent-based Models (ABM) for epidemiology

ABMs are focused towards describing individual behavior and model the collective behavior of epidemic spread bottom-up, in contrast to ODE models that use a top-down approach. These models inherently capture interactions between a healthy and infected individuals, without taking into account a specific ODE like model [1]. ABMs have been adopted in order to simulate social and geospatial sciences to capture the individual interactions that tend to be dynamic and highly dependent on the spatial location of the individual. In particular, individual agents are micro-entities interacting and changing over time giving rise to a collective behavior, commonly referred to as a macro system. The individual agents change their behavior based in the environment and their response to input from other agents as well. The systems are inherently stochastic, especially since the interactions can be dynamic and are not controlled by any specific patterns input to the programs. ABMs are typically run several times in order to obtain a distribution of possible outcomes based on the initial conditions. Therefore ABMs represent a bottom-up approach to modeling systems, which at the lowest granularity of information are represented by the individual

agents.

ABMs in epidemiology have been popular, and have been widely used to study a variety of diseases including small-pox, west nile virus and influenza [3, 15, 10, 4]. In addition, there have also been efforts to validate ABMs against ODE-based models [2]. ABMs are typically attractive for their simpler representation of the system (as individual acting agents), and they can include data from a variety of sources, including national census records, clinical data and social networks [11]. Furthermore, they do not require any explicit empirical data and are more useful in the context when empirical data is unavailable (or even difficult to obtain).

However, it is also important to note that ABMs face a number of challenges, key of which is the fine granularity of data included in these models. Hence, the model complexity with information “overload” at the individual agent level can often obscure the outcomes of the model rather than keeping a simple outlook on the predictions that ABMs make. Often ABMs require multiple sources of data (i.e., individual national census reports, clinical inputs, possible activity records for individuals) that such detailed scenarios may thwart the very appeal of such models. A second challenge faced by ABMs is that with individual agents and interactions being the focus of these models, it is difficult to often estimate the robustness or uncertainty in these models. Apart from this, ABMs focus mostly on ‘worst-case’ or the ‘most-likely’ scenarios rather than a gamut of possible outcomes, which makes it difficult to estimate the state-space of such models. Given the intuitive appeal and the difficulty associated with understanding the complex landscape of ABMs, there is a necessity to develop approaches that can rigorously verify and validate these models.

2. MODEL CHECKING EPIDEMIOLOGICAL MODELS

While the development of complex differential equation and agent-based models has enjoyed substantial attention in the epidemiological community [16, 19, 3, 2, 1], the validation of such dynamical and stochastic models against expert insight and historical data has not been an active area of research. In this section, model checking is demonstrated as an effective algorithmic framework for two purposes: (i) it can be used to verify whether an epidemiological model satisfies the qualitative expert insight and numerical data that has been observed historically, and (ii) it can quantify the uncertainty in a prediction obtained from an epidemiological model that may include control measures designed to contain a disease.

2.1 Formal methods in epidemiology

Various models used to study the spread of diseases can be broadly classified into the following three classes of dynamical systems:

ODE-based dynamical systems. An ordinary differential equation model has an infinitely dense semantics. However, for epidemiological purposes, it is sufficient to consider a discretized trajectory of a dynamical system modeled as an ordinary differential equation. While an ODE model is indeed deterministic, many parameters in an epidemiological ODE model are only known up to their probability distribution. The shape and the hyper-parameters of the probability distribution can be obtained from any number of statistical and machine learning methods. Thus, an ODE epidemiological model $\mathcal{M}_{ode}(\vec{p})$ is parameterized by a vector of random variables $\vec{p} = (p_1, \dots, p_n)$, each of which is only known up to a distribution. For example, each parameter may be modeled as a normal distribution with a known mean and variance. An ODE epidemiological model with such probabilistically known parameters naturally forms a stochastic model that allows an epidemiologist to

explicitly capture the inherent uncertainty in our understanding of the spread of diseases.

Stochastic differential equations (SDE)-based dynamical systems. While ordinary differential equations with probabilistically distributed parameters permit the specification of the uncertainty in measuring such model parameters, they do not capture the fundamental time-varying uncertainty inherent in the natural process of infection and recovery. In other areas of systems biology, it has been shown that small random perturbations can substantially affect the behavior of a system. Thus, it is important to explore the impact of such physical “white-noise” noise on the impact of dynamical models in epidemiology. A convenient framework for specifying such models is the addition of a scaled Brownian motion to the standard SIR, SEIR, SIS and SII models.

ABMs. The finite-time behavior of an agent-based model can be described by the evolution of a parameterized discrete time Markov chain (DTMC). An agent-based model \mathcal{A} consists of a bounded number of m different types of agents. Thus, an agent A_i , ($1 \leq i \leq m$) may have at most n_i copies at any point during the finite simulation. Each agent A_i in an epidemiological model is a continuous chain Markov chain (CTMC) itself and evolves by jumping from one state to another with a probability that is a function of the current state and the transition been made. Thus, the agent-based model itself is a synchronous composition of the various agents, i.e.

$$\mathcal{M} = A_1^{n_1} \parallel A_2^{n_2} \parallel \dots \parallel A_m^{n_m}$$

Here, $A_i^{n_i}$ is a shorthand notation for the composition $A_i \parallel A_i \parallel \dots \parallel A_i$ (repeated n_i times). As the behavior of an agent may depend on one or more parameters, it is useful to explicitly represent such parameterization by representing the ABM as $\mathcal{M}_{abm}(\vec{p})$

Thus, the formal description of the spread of diseases naturally gives rise to the study of parameterized stochastic models, including parameterized ODEs, SDEs and ABMs.

2.2 Qualitative expert insight, quantitative historical data and probabilistic bounded spatio-temporal logic

While considerable effort has been directed at the development of epidemiological models, the specification of the correct behavioral outcomes from such models has not been a focus of active research. In this section, a formal definition of high-level behavioral specifications for epidemiological models is proposed. For the purposes of validating epidemiological models, the focus is on behavioral specifications whose truth value can be decided by observing only a finite prefix of model simulation. Such specifications are referred to as *adapted finitely monitorable* (AFM) [23] specifications.

For an epidemiological model \mathcal{M} , classical flavors of temporal logic [29, 17], such as Bounded Linear temporal Logic (BLTL) and Continuous Stochastic Logic (CSL), are not sufficient as space plays a crucial role in the evolution of the epidemiological models. Instead, we define a new logic - the Bounded Spatio-Temporal Logic (BSTL) - that includes the notion of both space and time.

The state variables \mathbf{V} in an epidemiological model form a finite discrete-valued set. A Boolean predicate over \mathbf{V} is a constraint and has form $x \leq v$ here $x \in \mathbf{V}$, and $v \in \mathbb{R}$. A spatial operator $\mathbf{P}(A, C)$ computes the number of entities of type A present in the compartment C ; for example, in an ABM model, it may compute the number of infected patients in a geographical location. Another spatial operator $\mathbf{N}(A, B, r)$ computes the number of entities of type A that lie within a radius r of one or more of the entities of type B . Here, we present the BSTL syntax using the following formal

grammar:

$$\phi ::= x \leq v \mid \mathbf{P}(A, C) \leq v \mid \mathbf{N}(A, B, r) \leq v \mid (\phi_1 \wedge \phi_2) \mid (\phi_1 \vee \phi_2) \mid \neg \phi_1 \mid (\phi_1 \mathbf{U}^t \phi_2),$$

where $x \in \mathbf{V}$, $v \in \mathbb{Q}$, C is a compartment, $r \in \mathbb{Q}$ is distance, t is time, A and B are types of agents. We can have shorthand temporal operators such as $\mathbf{G}^t \psi = \neg \mathbf{F}^t \neg \psi$, or $\mathbf{F}^t \psi = \mathbf{True} \mathbf{U}^t \psi$ in terms of the bounded until \mathbf{U}^t . The formula $\mathbf{G}^t \psi$ implies that ψ holds at all time for the next t time units into the future and the formula $\mathbf{F}^t \psi$ indicates that ψ holds sometime within t time units. We can also specify interval versions of the temporal operators. For example, the formula $\mathbf{F}^{[t_1, t_2]} \psi$ indicates that ψ holds sometime between t_1 and t_2 time units.

It can be shown that finite paths of the model \mathcal{M} with bounded duration are sufficient for checking BSTL formula on epidemiological models. If ϕ is a BSTL formula, $Pr_{\geq \rho}(\phi)$ denotes a Probabilistic Bounded Spatio-Temporal Logic (PBSTL) formula. The PBSTL specification is said to be true on a model if a randomly sampled execution of the model satisfies the PSTL specification ϕ with probability at least ρ .

EXAMPLE 1. A disease spreads to at least 1/10th of a population within 10000 time units if it reaches at least 1/100th of the population in 9000 time units. Formally,

$$\mathbf{F}^{9000}(I/N > 1/100) \implies \mathbf{F}^{10000}(I/N > 1/10)$$

Here, I is the number of infected individuals and N is the total number of individuals.

2.3 Probabilistic model checking for uncertainty quantification

The method for verifying epidemiological models proposed in this paper is a descendent of formal verification [8]. While there has been research into the model checking of quantitative models including ordinary differential equations [28] and Continuous Time Markov Chain models [18], these methods do not scale to the size of the models that must be analyzed for epidemiological research. There has been recent work on analyzing large and complex stochastic models using statistical hypothesis testing [27, 20]. In particular, model checking methods that can verify complex ABMs [21] have been developed.

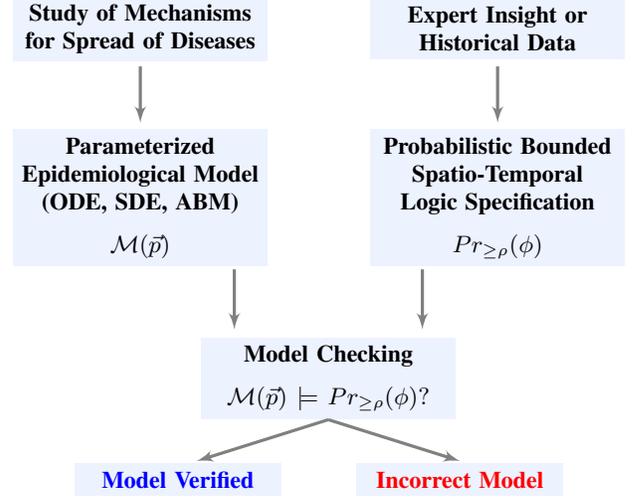
2.3.1 Statistical model checking

Statistical model checking [20] is an algorithmic framework that verifies a stochastic model against probabilistic behavioral specifications using a combination of independent and identically distributed sampling of executions from the model and sequential hypothesis testing. Epidemiological models, including parameterized ODEs, SDEs, and ABMs, can be simulated multiple number of times, and independent executions can be monitored algorithmically. Each execution is tested against the BSTL specification. If sufficient number of execution samples have been observed such that the probabilistic spatio-temporal logic specification can be either accepted or rejected statistically, the algorithm terminates. Otherwise, the algorithm will continue to explore additional behaviors of the epidemiological model.

2.3.2 Decision procedures based quantification of rare behaviors

Though the application of formal methods to other areas of Systems Biology has received considerable importance [12, 26], the focus has been on exploring the likely behaviors of stochastic models. For epidemiological models, the focus may often be on rare

but crucial behaviors that can have devastating socio-economic implications. Decision Procedures based on the Satisfiability Modulo Theory have been used previously for the formal verification of the correctness of the software and hardware models. Stochastic epidemiological models can also be analyzed using a combination of statistical model checking [22] and decision procedures [14]. The proposed method has become feasible due to the success of present research into nonlinear decision procedures [24].

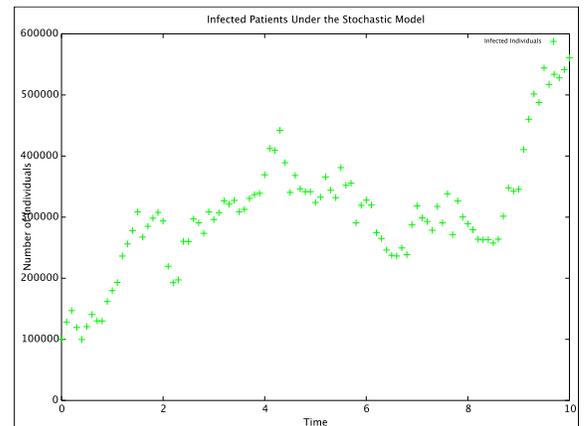


3. EXPERIMENTAL RESULTS

The following stochastic SIR model has been studied [25] in literature.

$$\begin{aligned} \frac{dX}{dt} &= \nu N - [\beta_{XY} N + f(X, Y)dW] - \mu X \\ \frac{dY}{dt} &= [\beta_{XY} N + f(X, Y)dW] - \gamma Y - \mu Y \\ \frac{dZ}{dt} &= \gamma Y - \mu Z. \end{aligned}$$

Initially, $X(0) = 10^6$, $Y(0) = 10^5$, and $Z(0) = 10^5$. Also, $N(t) = X(t) + Y(t) + Z(t)$, and dW is the standard Brownian Motion. Further, we assume that the transmission rate β_{XY} is 0.001, the mortality rate μ is 0.00005, and the infection rate ν is 0.00001, and the weighted influence of the random noise $f(X, Y)$ is set to be 10^5 .



From expert insight, we know that the spread of the disease in this area satisfies the following specifications:

1. An infection spreads to 5 times the initial infected population between 5 and 10 time units with probability 99%. Formally,

$$Pr_{\geq 0.99} (\mathbf{F}^{[5,10]} (Y > 5 Y(0)))$$

2. It is not the case that the infection spreads to more than 100 times the initial infected population within 5 and 10 time units with probability more than 1%. Formally,

$$Pr_{\leq 10^{-2}} (\mathbf{F}^{[5,10]} (Y > 100 Y(0)))$$

Using statistical model checking and numerical simulation of the stochastic model, it is observed that specification (1) is not true on the epidemiological model. The statistical model checking algorithm performed 182 numerical simulations of the model before arriving at this conclusion. Specification (2) is accepted by the model checking algorithm and shows that this model captures the scenario where the infection does not spread to more than 100 times the initial infected populations between 5 and 10 time units with probability more than 10^{-4} . The statistical model checking algorithm requires 113,984 independent and identically distributed observations to verify the second specification.

4. CONCLUSION AND FUTURE WORK

In this paper, we have presented a broad overview of mathematical models in epidemiology and discussed how we can quantify uncertainty within these models. We first highlighted the different types of models used by epidemiologists to study disease spread dynamics, including ODE based models and ABMs. ODE based models are relatively simple and are widely used to model disease spread within large populations. However, they tend to be restricted in terms of their ability to capture heterogeneity and the inherent randomness associated with disease spread. ABMs attempt to capture both the heterogeneity and randomness associated with these processes. However, ABMs suffer from the inherent complexity in the model as well as the amounts of data required to obtain a realistic picture of disease spread dynamics. Both models suffer from the lack of tools to quantify the uncertainty in their outcomes and predictions.

Given that both ODE and ABMs are going to play a greater role in predicting disease spread, designing intervention strategies, as well as shaping the public health policies in the future, uncertainty quantification is the required first step towards designing better models for epidemiology. We have illustrated how we can use a combination of statistical model checking and decision procedures-based verification methods to quantify the probabilistic outcomes from epidemiological models. While statistical model checking methods are particularly well-suited for complex spatio-temporal properties, decision procedures-based methods can uncover rare but socio-economically important consequence of epidemiological models.

In this paper, we did not present how epidemiological models can be verified against historical data. However, retrospective data analysis, in combination with the tools such as model checking for uncertainty quantification proposed here can be powerful tools to verify, validate and refine both ABMs and ODE based epidemiological models. We hope that such strategies will gain popularity within the epidemiological community in the near future.

Acknowledgment

The contents of this publication are the responsibility of the authors and

do not necessarily represent the official views of DTRA. ORNL is operated by UT-Battelle, LLC, for the U.S. Department of Energy under contract DE-AC05-00OR22725. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

5. REFERENCES

- [1] Amy H. Auchincloss and Ana V. Diez Roux. A new tool for epidemiology: The usefulness of dynamic-agent models in understanding place effects on health. *American Journal of Epidemiology*, 168(1):1–8, 2008.
- [2] R Bagni, R Berchi, and P Cariello. A comparison of simulation models applied to epidemics. *Journal of Artificial Societies and Social Simulation*, 5:33, 2002.
- [3] A Bell, A King, and K Pielak. Epidemiology of measles outbreak in british columbia- february 1997. *Canada Communicable Disease Report*, 23:49–51, 1997.
- [4] L Bian and D Liebner. Simulating spatially explicit networks for dispersion of infectious diseases. *GIS, Spatial Analysis and Modeling*, pages 245–264, 2005.
- [5] S. Cauchemez, P.-Y. Boëlle, C.A. Donnelly, N.M. Ferguson, G. Thomas, G.M. Leung, A.J. Hedley, R.M. Anderson, and A.-J. Valleron. Real-time estimates in early detection of sars. *Emerg Infect Dis*, 12(1):110–113, 2006.
- [6] F. Ciocchetta and J. Hillston. Bio-pepa: a framework for the modelling and analysis of biochemical networks. *Theoretical Computer Science*, 410((33-34)):3065–3084, 2009.
- [7] F. Ciocchetta and J. Hillston. Bio-pepa for epidemiological models. In *Proc. of PASM'09, ENTCS*, volume 261, pages 43–69, 2010.
- [8] Edmund M. Clarke, Orna Grumberg, and Doron A. Peled. *Model Checking*. The MIT Press, January 1999.
- [9] P. Drabik and G. Scatena. An application of model checking to epidemiology. In P. Milazzo and M. J. Perez Jimenez, editors, *Proceedings of Applications of Membrane computing, Concurrency and Agent-based modelling in Population biology (AMCA-POP)*, pages 90–97, 2010.
- [10] JB Dunham. An agent-based spatially explicit epidemiological model in mason. *Journal of Artificial Societies and Social Simulation*, 9, 2005.
- [11] Stephen Eubank, Hasan Guclu, V. S. Anil Kumar, Madhav V. Marathe, Aravind Srinivasan, Zoltan Toroczkai, and Nan Wang. Modelling disease outbreaks in realistic urban social networks. *Nature*, 429(6988):180–184, 05 2004.
- [12] François Fages. From syntax to semantics in systems biology towards automated reasoning tools. 3939:68–70, 2006.
- [13] Neil M. Ferguson, Matt J. Keeling, W. John Edmunds, Raymond Gani, Bryan T. Grenfell, Roy M. Anderson, and Steve Leach. Planning for smallpox outbreaks. *Nature*, 425(6959):681–685, 10 2003.
- [14] Arup K. Ghosh, Faraz Hussain, Sumit Kumar Jha, Christopher James Langmead, and Susmit Jha. Decision procedure based discovery of rare behaviors in stochastic differential equation models of biological systems. In *ICCABS*, pages 1–6, 2012.
- [15] TJ Gordon. A simple agent model of an epidemic. *Technological Forecasting and Social Change*, 70:397–417, 2003.
- [16] N.C. Grassly and C. Fraser. Mathematical models of infectious disease transmission. *Nat Rev Microbiol*, 6:477–487, 2008.
- [17] Hans Hansson and Bengt Jonsson. A logic for reasoning about time and reliability. *Formal Aspects of Computing*, 6:102–111, 1994.
- [18] J. Heath, M. Kwiatkowska, G. Norman, D. Parker, and O. Tymchyshyn. Probabilistic model checking of complex biological pathways. In C. Priami, editor, *Proc. Computational Methods in Systems Biology (CMSB '06)*, volume 4210 of *Lecture Notes in Bioinformatics*, pages 32–47. Springer Verlag, 2006.
- [19] S. Huang. Probabilistic model checking of disease spread and prevention. Master's thesis, Computer Science Department, University of Maryland, 2010.
- [20] Sumit Kumar Jha, Edmund M. Clarke, Christopher James Langmead,

- Axel Legay, André Platzer, and Paolo Zuliani. A bayesian approach to model checking biological systems. In *CMSB*, pages 218–234, 2009.
- [21] Sumit Kumar Jha, A. Donze, R. Khandpur, J. Dutta-Moscato, Q. Mi, Y. Vodovotz, G. Clermont, and Christopher James Langmead. Parameter estimation and synthesis for systems biology: New algorithms for nonlinear and stochastic models. *Journal of critical care*, April 2011.
- [22] Sumit Kumar Jha and Christopher James Langmead. Exploring behaviors of sde models of biological systems using change of measures. In *ICCABS*, pages 111–116, 2011.
- [23] Sumit Kumar Jha and Christopher James Langmead. Synthesis and infeasibility analysis for stochastic models of biochemical systems using statistical model checking and abstraction refinement. *Theor. Comput. Sci.*, 412:2162–2187, 2011.
- [24] Susmit Jha, Rhishikesh Limaye, and Sanjit Seshia. Beaver: Engineering an efficient smt solver for bit-vector arithmetic. In *Computer Aided Verification*, pages 668–674. 2009.
- [25] M.J. Keeling and P. Rohani. *Modeling Infectious Diseases in Humans and Animals*. Princeton University Press, 2008.
- [26] C.J. Langmead. Generalized Queries and Bayesian Statistical Model Checking in Dynamic Bayesian Networks: Application to Personalized Medicine. In *Proc. of the 8th International Conference on Computational Systems Bioinformatics (CSB)*, pages 201–212, 2009.
- [27] C.J. Langmead and S. K. Jha. Predicting protein folding kinetics via model checking. *Lecture Notes in Bioinformatics: The 7th Workshop on Algorithms in Bioinformatics (WABI)*, pages 252–264, 2007.
- [28] C. Piazza, M. Antoniotti, V. Mysore, A. Policriti, F. Winkler, and B. Mishra. Algorithmic Algebraic Model Checking I: Challenges from Systems Biology. *17th Internl Conf. on Comp. Aided Verification (CAV)*, pages 5–19, 2005.
- [29] Amir Pnueli. The temporal logic of programs. In *FOCS*, pages 46–57. IEEE, 1977.
- [30] J.K. Taubenberger and D.M. Morens. 1918 influenza: the mother of all pandemics. *Emerg Infect Dis*, 12(1):15–22, 2006.