

Chapter 8

Analyzing Global Epidemiology of Diseases Using Human-in-the-Loop Bio-Simulations

Dhananjai M. Rao, Alexander Chernyakhovsky and Victoria Rao

Abstract Humanity is facing an increasing number of highly virulent and communicable diseases such as influenza. Combating such global diseases requires in-depth knowledge of their epidemiology. The only practical method for discovering global epidemiological knowledge and identifying prophylactic strategies is simulation. However, several interrelated factors, including increasing model complexity, stochastic nature of diseases, and short analysis timeframes render exhaustive analysis an infeasible task. An effective approach to alleviate the aforementioned issues and enable efficient epidemiological analysis is to manually steer bio-simulations to scenarios of interest. Selective steering preserves causality, inter-dependencies, and stochastic characteristics in the model better than “seeding”, i.e., manually setting simulation state. Accordingly, we have developed a novel Eco-modeling and bio-simulation environment called SEARUMS. The bio-simulation infrastructure of SEARUMS permits a human-in-the-loop to steer the simulation to scenarios of interest so that epidemics can be effectively modeled and analyzed. This article discusses mathematical principles underlying SEARUMS along with its software architecture and design. In addition, the article also presents the bio-simulations and multi-faceted case studies conducted using SEARUMS to elucidate its ability to forecast timelines, epicenters, and socio-economic impacts of epidemics. Currently, the primary emphasis of SEARUMS is to ease global epidemiological analysis of avian influenza. However, the methodology is sufficiently generic and it can be adapted for other epidemiological analysis required to effectively combat various diseases.

D. M. Rao (✉)

CSE Department, Miami University, Oxford, OH 45056-1601, USA

e-mail: raodm@muohio.edu

A. Chernyakhovsky

Massachusetts Institute of Technology, Cambridge, MA 02139, USA

e-mail: achernya@mit.edu

V. Rao

Cybernetic Evolution, Mason, OH 45040, USA

8.1 Introduction

Communicable and vector-borne diseases, primarily induced by viral and bacterial infections, are the most common forms of ailments amongst humans and domesticated farm animals (Flint et al. 2004). Despite numerous advances in epidemiology and medicine many of these diseases, particularly zoonotic diseases, continue to defy human efforts to treat and control them (Flint et al. 2004). The source of such resistance to treatment and control is twofold. First, the incessant evolutionary processes such as genetic reassortment, recombination, and mutation induce changes to organisms, thereby constantly morphing their antigenic characteristics (Russell et al. 2008). Second, the vectors of the diseases, such as birds and animals, often disperse the pathogens over wide areas (Hagemeijer and Mundkur 2006; Normile 2006). The scenario is further convoluted due to complicated symbiotic processes between the vectors and other organisms including humans. Rapid changes in phylogeny and bio-diversity pose significant challenges in vaccine manufacturing (WHO 2007a). In addition, socio-political restrictions and technological limitations further magnify the challenges involved in controlling intercontinental spread of infectious diseases (WHO 2007a).

The aforementioned epidemiological, antigenic, and socio-political traits of communicable diseases are salient characteristics of several recent global epidemics, such as: severe acute respiratory syndrome (SARS) (Hufnagel et al. 2004), avian influenza (H5N1) outbreaks (WHO 2006), and swine flu (H1N1) outbreak (Garten et al. 2009; Smith et al. 2009). Such emergent viral epidemics cause significant morbidity and mortality, both in humans and livestock (CDC 2006; WHO 2005). For example, human influenza-A viruses causes more than 500,000 mortalities in humans annually (Russell et al. 2008). Its highly pathogenic avian homologue, namely H5N1, has caused economic devastation, particularly to poultry farming, in excess of 10 billion dollars (Kilpatrick et al. 2006).

The primary approach for combating emergent diseases, particularly viral diseases, is proactive, targeted antiviral prophylaxis (CDC 2006; WHO 2007b). Unfortunately, the constantly changing antigenic characteristics of viruses, particularly all influenza viruses, reduce efficacy of vaccinations (WHO 2006c). Moreover a myriad of technological issues pose serious hurdles to manufacturing and distribution of even small volumes of H5N1 vaccines (WHO 2006c). Furthermore, intercontinental migration of birds and international human travels further exacerbate effective containment and mitigation of epidemics.

8.1.1 Epidemiological Simulations: A Brief Overview

The technological, macro-economic, and socio-political issues surrounding emergent, viral and communicable diseases (discussed in Sect. 8.1) have led to a rapidly growing emphasis on analyzing the global epidemiology of diseases.

Epidemiology broadly covers the study of various facets of communicable diseases in a selected population, including: transmission characteristics of diseases, epidemic and endemic states, environmental factors impacting disease transmission, and impacts of containment and prophylactic interventions (Anderson and May 1992). Epidemiology is a multidisciplinary method involving biology, biostatistics, geographic information science, social science, and computer science (Anderson and May 1992; Daley and Gani 2001). It has steadily evolved from its inception in early nineteenth century and is now a mainstream methodology that forms the corner stone for public health measures and preventive medicine (Anderson and May 1992). Epidemiological analysis enables more accurate forecasting of epidemics and aids improving efficacy of geographically targeted antiviral prophylaxis (Epstein 2009; Ferguson et al. 2006; Halloran et al. 2008; Longini et al. 2005). Nevertheless, epidemiology is an active area of research as it continues to evolve in conjunction with breakthroughs in other disciplines, such as: microbiology, genetics, proteomics, and meta-genomics (Rao et al. 2007a).

A fundamental and widely used framework underlying epidemiological analysis is compartmentalized model of communicable diseases (Anderson and May 1992; Daley and Gani 2001). In a compartmentalized model, the sub-population that are in the same epidemic state are partitioned into non-intersecting sets called compartments (Anderson and May 1992; Daley and Gani 2001). The temporal progress of an epidemic is modeled by transitioning a suitable fraction of the population from one compartment to another (Anderson and May 1992; Daley and Gani 2001). The transition functions are typically modeled as ordinary differential equations (ODEs) with time as the independent variable. Constants in the ODEs are determined based on characteristics of the disease being analyzed. A more detailed description of the compartmentalized models is presented in Sect. 8.2.1.

Analysis of compartmentalized epidemiological models is performed by simultaneously solving the ODEs at different time steps to obtain characteristics of the epidemic being studied (Anderson and May 1992; Daley and Gani 2001). One of the most powerful and widely used approaches for such an analysis is computer-based simulation. Simulations have gained significant importance epidemiology because they are the only practical approach to analyze large and complex epidemiological models (Epstein 2009; Ferguson et al. 2006; Halloran et al. 2008; Longini et al. 2005; Rao and Chernyakhovsky 2008). Furthermore, the need to conduct a variety of multi-faceted analysis within short time frames necessitates the use of computer-based simulations. Moreover, simulation is a cost effective and non-destructive methodology from which results can be easily displayed in an intuitive form. More importantly, it enables effective explorations of policies and procedures associated with complicated control measures, such as: targeted layered containment [the main concept behind US government's containment strategies (Halloran et al. 2008)], quarantine, social distancing, school closing, and targeted antiviral prophylaxis (Epstein 2009; Ferguson et al. 2006; Longini et al. 2005; Rao et al. 2007a).

8.1.2 Motivation for Human-in-the-Loop Simulation

Currently, epidemiological models are used by many international agencies, including the World Health Organization (WHO) and the Centers for Disease Control (CDC), for large-scale, multi-faceted analysis required to propose and validate multi-national targeted layered containment policies as well as prophylactic measures (Epstein 2009; Ferguson et al. 2006; Longini et al. 2005; Rao et al. 2007a). Such epidemiological analysis is performed using detailed, stochastic models that involve complex, symbiotic interactions between the various entities involved in the simulation.

Typically, only a selected subset of scenarios is analyzed to evaluate effectiveness of a candidate set of containment strategies (Ferguson et al. 2006; Halloran et al. 2008; Longini et al. 2005). Narrowing the subset of scenarios to be analyzed is critical because hundreds of thousands of simulation runs are needed even to analyze a specific subset of scenarios. The numerous runs are necessary to provide sufficiently accurate and statistically significant results using realistic, stochastic models (Epstein 2009; Ferguson et al. 2006; Longini et al. 2005).

One of the important prerequisites of simulating selected scenarios is to have the model in a state that accurately reflects the conditions associated with the scenario. One strategy is to seed the simulation, i.e., manually initialize the state to appropriately model the scenario being analyzed. However, such an approach is typically tedious, cumbersome, and error prone. Moreover, in stochastic models, accurately capturing temporal, causal relationships (Lamport 1978) between entities in the model can prove to be a greater challenge. These issues often lead to conspicuous inefficiencies when multiple scenarios have to be analyzed. An alternative approach is to commence simulations in a given, verified initial state and then steer the simulations to desired scenarios of interest. Strategies for steering simulations can be broadly classified into three main categories, namely: automatic, semi-automatic, and manual steering. Note that this taxonomy is based on the degree of human involvement in the steering process.

The complex, stochastic nature of global epidemiological models necessitates the use of manual steering. Manual steering of a simulation requires a human-in-the-loop (HITL) to intermittently modify the state of the simulation; thereby altering its trajectory to the desired scenario. Such HITL epidemiological bio-simulations enable efficient generation of various scenarios to be analyzed. In addition, this approach preserves the causal inter-dependencies in a stochastic model. Moreover, simulating from a validated initial state minimizes modeling errors and reduces continued verification efforts. Last but not the least, HITL steering enables some degree of adaptation and modeling of unplanned or unanticipated scenarios that could occur in emergent epidemics. Consequently, HITL-simulations hold significant potential to play a vital role in epidemiology and in establishing pertinent public policies.

8.1.3 SEARUMS: A Human-in-the-Loop Epidemiological Simulator

Realizing the advantages of simulation-based epidemiological analysis requires the use of an effective software environment for Eco-modeling, HITL bio-simulation, and analysis (Rao et al. 2007a). In addition, the software environment must facilitate rapid simulation to minimize analysis time frames. We also envision the software must be portable and accessible to enable its widespread use (Rao et al. 2007a). Accordingly, we have endeavored to design and develop an Eco-modeling and bio-simulation software environment called SEARUMS.

Currently, SEARUMS is geared for analyzing the global epidemiology of avian influenza, with migrating waterfowl as the primary vectors for intercontinental spread of the disease. However, the methodology is sufficiently generic and can be easily adapted for other diseases. Accordingly, this chapter elucidates the modeling, HITL-simulation, and software design principles underlying SEARUMS so that the concepts can be readily adapted and applied to other diseases as well. Furthermore, SEARUMS is envisioned to serve as a global, multi-disciplinary environment that seamlessly integrates knowledge from various fields so that epidemiologists, economists, and disease control centers can collaboratively use it to combat global epidemics.

8.1.4 Section Organization and Audience

The principles underlying the Eco-modeling and HITL bio-simulation infrastructure of SEARUMS span a broad range of disciplines, including: mathematical modeling, statistical analysis, computer science, and software engineering. Furthermore, the multi-disciplinary nature of SEARUMS garners interest from a broad range of audience, including: epidemiologists, economists, software architects, and disease control centers. Consequently, the various sections in this chapter are organized to emphasize a specific aspect of SEARUMS. The objective of this organization is to enable readers from a particular discipline to focus on pertinent topics. Accordingly, common concepts that are pervasive across the various sections are discussed as the background information in Sect. 8.2. This section also reviews some of the pertinent related research investigations. Section 8.3 discusses the software architecture and design principles underlying SEARUMS. The procedure used for modeling, incorporation of real-world statistical data, and analysis using SEARUMS is presented in Sect. 8.4. Results from some of the analysis conducted using SEARUMS are summarized in Sect. 8.5. Section 8.6 concludes the chapter while emphasizing the utility and broader applicability of the HITL-simulations infrastructure of SEARUMS to other diseases.

8.2 Background and Related Research

Simulation-based analysis of global epidemiology of diseases is a multi-disciplinary task that involves ecology, mathematics, statistics, computer science, and software engineering (Anderson and May 1992; Epstein 2009; Ferguson et al. 2006; Halloran et al. 2008; Longini et al. 2005; Rao et al. 2009). Accordingly, this section covers relevant prerequisite information and terminology on the aforementioned topics along with brief surveys of closely related research investigations. Specifically, Sect. 8.2.1 covers the compartmental models used in epidemiology. Although the compartmental models can be applied to many diseases, in this article we focus on its use for analyzing global epidemiology of avian influenza. Consequently, the ecology of avian influenza is summarized in Sect. 8.2.2 to further motivate the Eco-modeling and bio-simulation environment constituting SEARUMS. Next Sect. 8.2.3 presents an overview of the Markov processes that are used to model the natural interactions occurring in the ecology. This section also illustrates the interactions between the compartmental models and Markov processes. Section 8.2.4 presents a brief survey of the concepts and software frameworks pertaining to HITL steering of simulators and software systems.

8.2.1 Compartmental Models in Epidemiology

The most widely used mathematical framework for epidemiological analysis are compartmentalized models (Anderson and May 1992; Daley and Gani 2001). In a compartmentalized model the population being analyzed is partitioned into a few non-intersecting subsets called compartments. Compartments are defined such that the sub-population within a compartment exhibits a vital disease characteristic (Anderson and May 1992; Daley and Gani 2001), such as:

- *Susceptible*: population lacks immunity
- *Exposed*: infected, but not yet infectious sub-population
- *Infected*: sub-population is actively spreading the disease, and
- *Recovered*: sub-population acquired immunity or died due to disease.

Additional compartments are used to appropriately model other epidemiological states of diseases depending on the analysis needs (Epstein 2009; Ferguson et al. 2006; Halloran et al. 2008; Longini et al. 2005; Rao et al. 2007a).

The classical susceptible-exposed-infected-removed (SEIR) compartmentalized mathematical model shown in Fig. 8.1 is used to model the epidemiology of various diseases (Anderson and May 1992). The characteristics of the SEIR model are represented using the following system of differential equations:

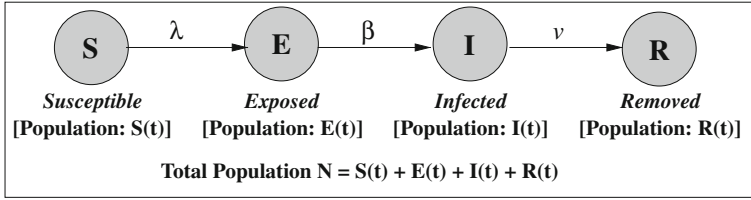


Fig. 8.1 A classical SEIR epidemiology model illustrating the four typical compartments used to describe the progression of a disease through a population (of N individuals)

$$\frac{dS}{dt} = \mu N - [\lambda + \mu]S(t)$$

$$\frac{dE}{dt} = \lambda S(t) - (\beta + \mu)E(t)$$

$$\frac{dI}{dt} = \beta E(t) - (\nu + \mu)I(t)$$

$$\frac{dR}{dt} = \nu I(t) - \mu R(t)$$

where, $S(t)$, $E(t)$, $I(t)$, and $R(t)$ represent the number of susceptible, exposed, infected, and removed hosts at any given instant of time t . The temporal progress of an epidemic is modeled by transitioning a fraction of the population from one compartment to another (Anderson and May 1992; Daley and Gani 2001). The transition functions are typically modeled as ODEs with time as the independent variable.

Constants in the ODEs are determined based on characteristics of the disease being analyzed. Specifically, the parameters μ , λ , β , and ν are:

1. μ : the per capita host birth/death rate,
2. λ : the force of infection
3. β : latency period
4. ν : per capita recovery rate

Typically, the ODEs involve stochastic components to account for uncertainties introduced by various environmental or external factors (Anderson and May 1992). An important and distinguishing property of the compartmentalized models and the transition functions is that the total population (N in Fig. 8.1) being modeled is held a constant (Anderson and May 1992; Daley and Gani 2001).

Adaptations of the aforementioned classical SEIR model have been widely used to model the epidemiology of influenza (Epstein 2009; Ferguson et al. 2006; Halloran et al. 2008; Longini et al. 2005; Rao et al. 2007a). Investigations conducted by Longini et al. (2005) and Ferguson et al. (2006) focus on analyzing pandemic mode of H5N1 in Thailand. In pandemic mode rapid and sustained human-to-human transmission is assumed. Since human-to-human transmission is assumed, these investigations use a highly detailed spatially explicit model based

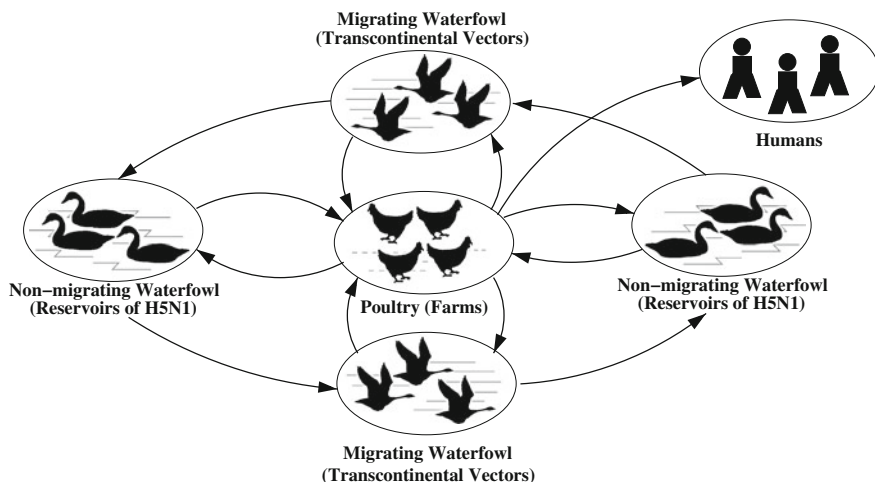


Fig. 8.2 The figure illustrates the salient pathways for transmission of H5N1 to various bird species and humans. The *arrows* in the diagram show the various directions of infection that occur in the ecology of avian influenza

on SEIR concepts. The modeling approach used by Halloran et al. (2008) is similar to those proposed by Ferguson et al. (2006) and Longini et al. (2005). Moreover, these three investigations are based on the premise that H5N1 has already mutated to a pandemic form and epidemics are being caused primarily due to human-to-human transmission.

In contrast, our investigations assume and reflect the current, real-world situation; i.e., H5N1 is yet to mutate into its pandemic state and human-to-human transmission is unsustainable. Furthermore, our research aims to utilize waterfowl migration data to forecast zoonotic epicenters and timelines of potential epidemics. In addition, we also emphasize the role of poultry (extensible to swine) as intermediate hosts. These aspects notably distinguish our efforts from the aforementioned recent related investigations.

8.2.2 Avian Influenza

Avian influenza is a viral disease caused by H5N1, a highly virulent strain of the influenza-A virus, that has the potential to cause a global pandemic (CDC 2006; WHO 2006c). The ecological interactions contributing to the transcontinental spread of the disease is illustrated in Fig. 8.2. As shown in the figure, infected migrating waterfowl, in which the virus is endemic, are the primary vectors for causing intercontinental spread of the disease (Normile 2006).

The virus rapidly spreads from waterfowl to poultry and humans through contaminated water, feed, feces, and surfaces. Once infected, the disease has a devastating impact on poultry farms causing 100% mortality within 48 h (WHO 2007b). In humans, the virus causes disease with a high mortality rate of nearly 60% (CDC 2006; Normile 2006). Furthermore, it is known to induce primary viral pneumonia in the host (Normile 2006). Researchers believe that avian influenza has significant potential to become one of the deadliest pandemics in human history (CDC 2006; WHO 2006c). This inference has been drawn based on statistics from recent epizootic outbreaks and the highly pathogenic characteristics of H5N1. Moreover, manufacturing and distribution of vaccinations is facing multi-faceted challenges (WHO 2007a). The aforementioned issues make it imperative the epidemiology of avian influenza is thoroughly analyzed in order to empower various national and international organizations with the knowledge to strategically combat the disease (Anderson and May 1992; Epstein 2009; Ferguson et al. 2006; Longini et al. 2005; GLiPHA 2007; Rao et al. 2009).

8.2.3 Markov Processes

The compartmental epidemiological models focus purely on the temporal progression of the disease in a given entity, may it be waterfowl, human, or poultry. However, the compartmental SEIR models (see Sect. 8.2.1) do not embody the complete ecology, such as seasonal migration, occurring in nature. Consequently, such ecological processes need to be suitably modeled in conjunction with the epidemiological process. Accordingly, we envision using the concept of a Markov processes to model the overall ecological processes. Furthermore, our Eco-modeling approach suitably incorporates the SEIR models to provide a complete, holistic representation of the real-world ecology.

A Markov process is a mathematical formalism used to describe changes occurring to the state of a stochastic system in discrete time steps (Solow and Smith 2006; Winston 1994). A Markov process consists of a number of states (or values) through which the system may transition at any given time. Mathematically, a Markov process is defined as a sequence of time dependent random variables X_0, X_1, X_2, \dots , where X_t is a random variable that describes the state of the process at discrete time t . The initial or starting state of the system is typically represented by X_0 . Transitions from one state to another are governed by the following three laws: (1) a Markov process may be in only one given state at any instant of time; (2) transition from one state to another occurs instantaneously in discrete time steps; and (3) the next state to which the process transitions is purely determined by the current state of the system and not its past. In other words, the past, present, and future states of a Markov process are independent of each other (Winston 1994).

In our approach, the SEIR operations are repeatedly preformed in an appropriate state in the Markov process. It must be noted that the SEIR models involving

ODEs are inherently based on the notion of continuous simulation-time (Anderson and May 1992). However, the Markov processes operate in discrete simulation-time steps (Winston 1994). Consequently, our Markov process-based approach approximates the SEIR model. However, the discrete simulation-time steps are chosen to be sufficiently small to provide adequate accuracy without degrading performance (Rao et al. 2009).

Another important aspect of our Eco-modeling methodology is the aggregate representation of sub-populations. At a global scale, the varying populations of entities are typically modeled at a coarser granularity or aggregate entities (Booth 1997; GROMS 2006; Law et al. 2005; Rao et al. 2009). For example, certain species of waterfowl that live and migrate as a large flock in nature are typically modeled as a single entity. A large collection of birds, such as a poultry farms are represented as a single entity. In an analogous manner, humans living in geographic proximity to each other are modeled as a single entity.

The motivation for such a coarser grained, aggregate representation is two fold (Booth 1997; GROMS 2006; Law et al. 2005; Rao et al. 2009). First, such an approach is necessary to reduce the size and complexity of the model to more tractable scales. Aggregated models continue to provide the necessary fidelity as long as the aggregation is performed at a sufficiently fine granularity (Law et al. 2005). Second, various real-world statistical data on population dynamics, seasonal migratory behaviors, and disease progresses are determined through random sampling of a large subset of the population. Such data is more meaningful and relatively straightforward to apply at an aggregate level.

In our software system the three primary aggregate entities, namely: (1) flock of waterfowl; (2) poultry; and (3) group of humans have been modeled using the aforementioned Eco-modeling approach. The Markov processes have been suitably implemented using the modeling framework provided by SEARUMS. The Markov process for the waterfowl entity is shown in Fig. 8.3. Readers are referred to the literature for more details on the Markov processes for various entities and the associated mathematical equations (Rao et al. 2009).

8.2.4 Human-in-the-Loop Simulation and Steering

In conjunction with advancements in microprocessor technologies, computer-based simulations have become an important, indispensable, and multi-disciplinary methodology for study and analysis of complex systems (Railsback et al. 2006). Today, simulations are widely used in fundamental sciences, applied sciences, engineering, medicine, economics, and in the military (Tobias and Hofmann 2004). During its inception, simulations were run purely as an offline or batch processing task without requiring any interaction with a human operator. However, the growing demand and diversification in needs in various disciplines have led to the development of simulations which involve interactions with a human operator (Rao et al. 2007b). Such simulations, which permit or require a human to interact,

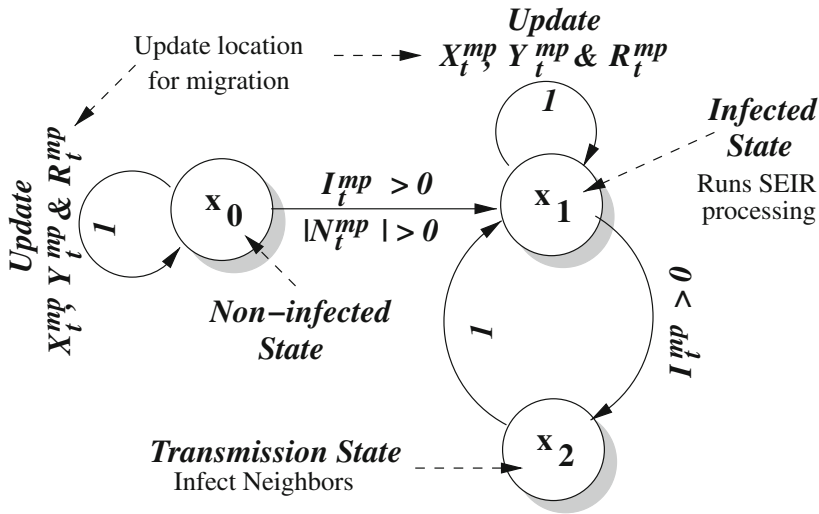


Fig. 8.3 Overview of Markov process for a waterfowl flock illustrating the states through which the process transitions to model the behavioral life cycle of waterfowl. The state of the process S_t^{mp} at discrete time t is represented by the four-tuple $S_t^{mp} = \langle X_t^{mp}, Y_t^{mp}, R_t^{mp}, I_t^{mp} \rangle$, where X_t^{mp} is the longitude, Y_t^{mp} is the latitude, R_t^{mp} is the instantaneous radius of the flock, I_t^{mp} is the current measure of infection in the flock. The set N_t^{mp} represents the neighborhood at time t . Readers are referred to the literature (Rao et al. 2009) for details on the mathematical notations, spherical geometric equations, and state transitions

are classified as HITL-simulations. The process of impacting the state of the simulation to divert the simulation to another state is called simulation steering. Simulation steering is similar to program steering, which is a more generic concept (Gu et al. 1994).

HITL-simulations can be further classified into constructive HITL-simulations and virtual simulations. In a virtual simulation, simulation-time progressed as the same rate as the real, wall-clock time (Rao et al. 2007b). Furthermore, interaction between the human and the simulation occurs in a synchronous manner. Virtual simulations are primarily used for training, gaming, and education (Rao et al. 2007b). On the other hand, in a constructive simulation, simulation-time progresses at a different rate than real time. Human interactions with a constructive simulation are typically intermittent or on a “as needed” basis. The interactions are triggered based on an user gesture or when predefined events occur in the simulation.

Constructive simulations are used for simulating large and complex systems such as the ecological and epidemiological models. Several such general purpose simulators have been described in the literature (Gilbert and Banks 2002; Railsback et al. 2006; WHO 2005). Some of the commonly used simulators include NetLogo, SWARM, SWARM-Java, Repast, and MASON. Railsback et al. (2006) highly recommended NetLogo for its ease-of-use. However, it uses a custom language for modeling and its source code is proprietary. On the other hand, the latter four platforms uses traditional programming languages and source

codes are freely available (Railsback et al. 2006). These four platforms essentially provide a core framework for model development and a collection of library modules. The library modules are built using the core framework and can be readily reused for modeling. Moreover, the simulators provide some support for HITL interactions.

The aforementioned simulators are mostly general purpose simulators and are not specifically designed for epidemiology. However, our objective is to minimize learning curves for both developers and users, maximize portability, include intuitive interfaces for modeling, and seamlessly incorporate epidemiological analysis tools. Furthermore, the simulators had some disadvantages (Rao et al. 2009). Consequently, we endeavored to develop SEARUMS, a custom Eco-modeling and HITL bio-simulation environment.

8.3 SEARUMS

SEARUMS is an Eco-modeling and HITL bio-simulation environment. Currently, it is optimized to enable study and analysis of global epidemiology of avian influenza. However, the design of its modeling and simulation framework is sufficiently generic. Therefore, it can be adapted for epidemiological analysis of other diseases. SEARUMS has been developed in Java by utilizing many of the language's object oriented programming features (Bloch 2001). SEARUMS is designed to be a user friendly, integrated, graphical modeling, simulation, visualization, and analysis environment for conducting epidemiological analysis. These design goals have been achieved by composing the system using a collection of interdependent but loosely coupled modules as shown in Fig. 8.4.

Each module shown in Fig. 8.4 has a well-defined functionality that can be accessed and utilized via a set of application program interface (API) method calls. APIs of the modules are Java interface classes that are implemented by each module. Interactions between modules are achieved through interface classes to ensure loose coupling. This approach permits seamless "plug and play" of modules and the environment is composed by loading suitable modules dynamically on-demand via Java's reflection API (Bloch 2001). Such an implementation approach has been adopted to ease customization and extension of SEARUMS without requiring changes to its design or impacting existing modules.

The modules constituting SEARUMS are broadly classified as core modules and graphical user interface (GUI) modules. The core modules of SEARUMS are the agent repository, agent customizer, persistence module, HITL steering module, simulation module, and logging module. These modules provide the core M&S functionality of SEARUMS. The GUI facilitates interactions with the core modules via convenient and intuitive user interfaces. The GUI modules can be further categorized into the editor subsystem, the simulation controller, and the visualization and analysis subsystem. The GUI presented by these modules is shown in Fig. 8.5. SEARUMS uses the model-view-controller pattern to couple the core

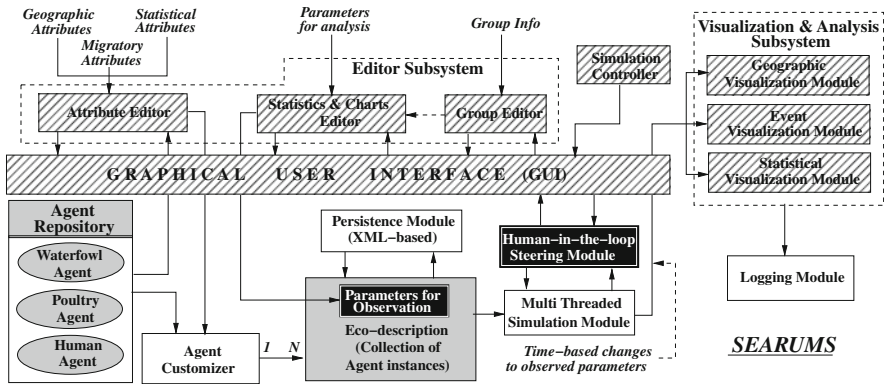


Fig. 8.4 Architectural overview of SEARUMS. GUI modules are highlighted with a striped background. SEARUMS can be downloaded from <http://www.searums.org>

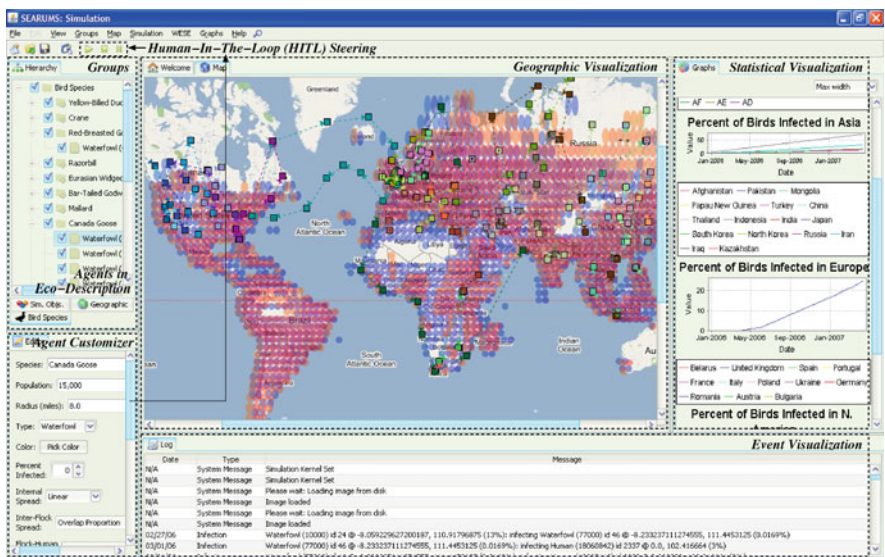


Fig. 8.5 The figure presents a screenshot of SEARUMS illustrating the graphical layout as seen by a user. The various modules constituting SEARUMS have been marked using black dashed lines. Purple circles indicate human groups and orange circles indicate poultry flocks. Variation in colors arises due to overlap of human and poultry flocks. Colored squares and corresponding colored dashed lines illustrate migration paths of waterfowl flocks

modules, the GUI models, and the Eco-description. The design permits the GUI modules to be easily replaced with a minimal command-line text interface for running SEARUMS in offline batch mode. The batch mode is useful for

performing repeated runs or analyzing different scenarios on computational clusters.

The modules and subsystems constituting SEARUMS cooperatively operate on a shared, in-memory representation of the model called the Eco-description. The Eco-description is a centralized data structure that includes all the information necessary for modeling, simulation, and analysis. It is composed using a collection of Java classes and that provide efficient access to data and information required by the various modules. The primary information encapsulated by the Eco-description parameters for analysis relates to the smart agents (Hare and Deadman 2004) that constitute the model. As shown in Fig. 8.4, the agents are organized into an AGENT REPOSITORY to facilitate instantiation and use via Java reflection API.

Currently, SEARUMS includes the following three smart agents: WATER-FOWL AGENT that represents a migrating waterfowl flock, POULTRY AGENT that models behavior of poultry flocks, and HUMAN AGENT that models humans. Each agent has its own behavior that reflects the characteristics of its real-world counterpart. The behaviors are customized to represent specific instances of an agent by specifying suitable values for the exposed attributes via the ATTRIBUTE EDITOR GUI module. The attributes of an agent include:

1. *Geographic attributes* that indicate the location (latitude and longitude) and logical association with countries and continents. In addition, each agent has a circle or influence that circumscribes its neighborhood.
2. *Migratory attributes* are specified only for agents whose location changes over the lifetime of the simulation. The migratory attributes are described as a sequence of migration points. Each migration point has geographical and chronological (arrival and departure dates) attributes associated with it. In SEARUMS, only one complete migration cycle needs to be specified. The software automatically reuses the information to simulate annual migratory cycles.
3. *Statistical attributes* for agent instances include their initial population, density and distribution, initial infection percentage, infection spread parameters, incubation periods, mortality rates, and population re-growth parameters.

The agents implement the conceptual, mathematical model of the system developed using Markov processes as described in Sect. 8.2.3. They are added to a model via suitable toolbar buttons or menu options provided by SEARUMS. Agent instances are created with default attributes from the agent repository by the AGENT CUSTOMIZER module using Java's reflection API. Once instantiated, the attributes for agents can be modified via the ATTRIBUTE EDITOR module. The agents are implemented as a family of Java classes by extending a common base class called AGENT. The AGENT class provides methods for interacting with the simulation kernel, inspecting the neighborhood, scheduling events, and interfacing with the GUI modules.

The agents in a model are logically organized into hierarchical sets called groups. SEARUMS permits multiple top-level groups with an arbitrary number of hierarchies, with one or more sub-groups at each hierarchical level. An agent can be a member of multiple groups. The groups serve several different purposes in

SEARUMS. A group can be used as a parameter for statistical analysis and for plotting charts. For example, a group called US can be created with 50 different sub-groups, one for each state, encompassing various agents. The main US group can be selected for plotting charts and SEARUMS automatically collates and plots data for each state. Note that even though graph plotting is restricted to one hierarchical level, statistics for plotting are collated in a recursive, depth-rst manner and includes data from all agents in underlying hierarchies. A modeler can use a combination of groups to perform multi-faceted analysis at different scales. In addition, groups can be included or excluded from simulations for analyzing different scenarios. The GUI modules utilize groups to provide control on visibility of agents to manage details displayed on the screen. The group editor module provides the user interface for managing group entries and hierarchies.

Once all the agent instances and groups have been established in a model, the parameters for observation are added to the Eco-description. These parameters are selected by the user via the Statistics and Charts editor from a list of options. The list includes the attributes of the agents and the groups in the Eco-description. Each parameter is configured to be sampled hourly, daily, or weekly in terms of simulation-time. Moreover, each parameter can be subjected to statistical operations, such as sum, mean, and median. SEARUMS can dynamically (i.e., during simulation) plot and save a variety of charts including: line graphs and pie charts. Multiple charts can be simultaneously used for analyzing a variety of data.

All of the aforementioned information is stored as an integral part of the Eco-description. The Eco-description can be saved for future reuse via the PERSISTENCE MODULE. The Eco-description is unmarshalled into a XML document that is compliant with a predefined XML schema. Serializing to a XML document provides a few advantages. First, it enables simple scripts to be developed that can modify specific values and perform multiple simulation runs in batch mode. Second, XML documents can be readily version controlled and archived using commonly available revision control systems such as CVS and subversion. Third, it eases documentation, validation, sharing, and reuse of valuable domain-specific statistical data collated by different researchers from diverse sources. Lastly, it is used to create checkpoints that reflect different scenarios. Such features play an important role in facilitating large-scale, collaborative epidemiological studies.

The SIMULATION MODULE performs the task of conducting a discrete event simulation (DES) using the Eco-description. This module utilizes a multi-threaded DES kernel that manages and schedules the discrete events generated by the agents. Multi-threading enables the DES kernel to exploit the compute power of multiprocessor or multi-core machines thereby reducing the wall-clock time for simulation. The number of threads spawned by the DES kernel is configurable. Each thread processes concurrent events (events with the same timestamp) in parallel without violating the causal constraints between events.

Table 8.1 The different agent instances used to develop the Eco-description used for case studies

Description of agent type	No. of instances	Total population	No. of countries
Bar-tailed godwit	4	40,000	18
Canada goose	16	231,700	5
Common crane	9	22,500	21
Eurasian widgeon ^a	3	1,296,000	17
Great knot	3	231,000	8
Mallard ^a	1	5,000	1
Razorbill	1	148,000	4
Red-breasted goose ^a	1	44,000	4
Red-crowned crane	1	15,000	4
Siberian crane	3	30,000	12
Yellow-billed duck	2	20,000	8
Total waterfowl flocks	44	4,371,000	40
Total poultry flocks	1,315	18,136,146,826	All
Total human groups	1,314	6,646,739,849	All
Total	2,673	24,787,572,675	All

Note that each agent is used to represent a group

^a High risk waterfowl species (Hagemeijer and Mundkur 2006). The total of 44 waterfowl flocks with different migratory pathways were used. The total population column shows the sum of the populations of all agent instances in each category

8.3.1 Human-in-the-Loop Steering Module

In our current research, we have drawn inferences and conclusions from a variety of case studies conducted using a calibrated Eco-description. First an initial Eco-description was developed via SEARUMS using the modeling methodology described in Sect. 8.3. Table 8.1 lists the waterfowl species, including high risk species (Hagemeijer and Mundkur 2006), used to develop the Eco-description. The migratory flyways of the waterfowl and their population have been collated from data published by various organizations (CDC 2006; WHO 2006c; GROMS 2006; GLiPHA 2007, Hagemeijer and Mundkur 2006). For modeling and simulation purposes the dates for migration were approximated to the middle of the months reported in the statistics. Due to the significant variation in migration patterns the approximated migration dates are expected to have deviations of ± 2 weeks which are accounted for through stochastic changes in migration dates each time a simulation is performed.

8.4 Methods

In our current research, we have drawn inferences and conclusions from a variety of case studies conducted using a calibrated Eco-description. First an initial Eco-description was developed via SEARUMS using the modeling methodology

described in Sect. 8.3. Table 8.1 list the waterfowl species, including high risk species (Hagemeijer and Mundkur 2006), used to develop the Eco-description. The migratory flyways of the waterfowl and their population have been collated from data published by various organizations (CDC 2006; GLiPHA 2007; GROMS 2006; Hagemeijer and Mundkur 2006; USCB 2006; WHO 2006c). For modeling and simulation purposes the dates for migration were approximated to the middle of the months reported in the statistics. Due to the significant variation in migration patterns the approximated migration dates are expected to have deviations of ± 2 weeks which are accounted for through stochastic changes in migration dates each time a simulation is performed. The radius of the waterfowl agents were computed using the population, density, and dispersion data obtained from published statistics (GROMS 2006; Hagemeijer and Mundkur 2006).

The dispersion of poultry population in different continents has been approximated to circular regions with even density (GLiPHA 2007; Law et al. 2005; Booth 1997). Such a modeling approach is commonly used in spatially explicit ecological models (Booth 1997; Hare and Deadman 2004; Law et al. 2005; Winston 1994; Wolfram MathWorld 2006). Global poultry and human population density data have been collated from statistics published by national organizations and government databases (GLiPHA 2007; SEDAC 2007; USCB 2006). As shown in Table 8.1, our model includes the complete human population (approximately 6.646 billion) humans represented by 1,314 agents. On an average, each human agent models 5.058 million humans living in a contiguous circular region. However, the precise population represented by an agent varies depending on the density of the region it models. Agents modeling dense metropolitan areas have higher human populations while agents modeling rural areas of the world have lower population. In contrast, the radius of all the human agents in the model is equal. The radius was computed using the grid size of gridded human population data from SEDAC (2007).

A similar strategy has also been applied to distribute the 18.136 billion poultry birds to 1,315 poultry agents as shown in Table 8.1. All the waterfowl agents have equal radius as determined from the grid size of the gridded poultry data obtained from GLiPHA (2007). However, the poultry population represented by each agent varies depending on the world region being modeled by the agent. Note that our Eco-description includes only a selected subset of the waterfowl as complete migration data is unavailable. However, to the best of our knowledge, it is the most comprehensive model of its kind reported to date. Furthermore, it can be readily extended to include additional waterfowl entities from other parts of the world.

8.4.1 Human-in-the-Loop Steering and Calibration

The first step in our study was to calibrate the model to ensure that it provides a sufficiently accurate representation of real-world epidemiology. The calibration was performed in conjunction with verification and validation of the model.

Table 8.2 Comparison of chronology of significant Real-world outbreaks as reported by WHO (2006) against the simulated outbreaks

Incident	Real-world date	Simulated date	Error (days)
Outbreak in Indonesia	23-Jan-2006	01-Jan-2006	22
Infection in Iraq/Iran	01-Mar-2006	25-Mar-2006	24
Infection in China	27-Apr-2006	02-Apr-2006	-25
Infection in Egypt	11-Oct-2006	14-Sep-2006	-27

The data was recorded after the model was calibrated

We verified the accuracy and fidelity of the aforementioned Eco-description by performing extensive HITL-simulations with initial source of infection set to outbreak in Indonesia, a notable epicenter of H5N1 epidemics (WHO 2006a). The Eco-description was interactively calibrated by suitably tuning the following attributes: start date for simulation, initial infection percentage, intra-flock disease spread rate, and inter-flock transmission mechanism. Note that we calibrated only the attributes that were indirectly derived from published statistics. We established validity of the Eco-description and SEARUMS by confirming that the timing and chronology of several outbreaks observed in the simulations correlate with the significant real-world incidents reported (WHO 2006a) as shown in Table 8.2.

It must be noted that the HITL-simulations steering played an important role in calibrating the model, particularly given its size and complexity. Using the validated and calibrated Eco-description we performed several case studies to analyze the spread and pandemic threat posed by avian influenza to US. Some of these case studies are discussed in Sect. 8.5.

8.5 Experiments

The calibrated Eco-description (see Sect. 8.4 for details) has been used to conduct several experiments to analyze the potential impact of avian influenza to poultry farming in the US. The study was conducted using a number of bio-simulations with three different experimental groups of migrating waterfowl. The three flocks were chosen based on their close proximity to known primary sites of disease outbreaks (Rao et al. 2009). The initial infection in each experimental group was varied for analysis.

Figure 8.6 illustrates one of the trans-Atlantic transmission pathways to the continental US. We observed that the spread was determined by migratory pathways and timelines of different species of waterfowl rather than initial infection percentages (Rao et al. 2009). One of the interesting observations is that our experiments correctly predicted an outbreak in the UK (WHO 2006c). The graph in Fig. 8.7 presents the impact of avian influenza outbreaks on poultry population in the continental US. Decrease in poultry population corresponds to H5N1 induced death and culling of birds to control the disease. Increase in poultry

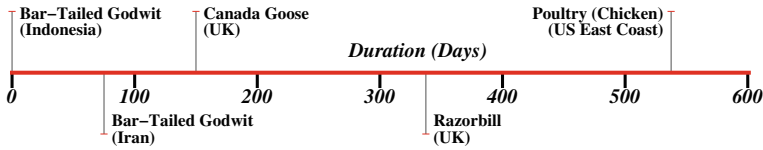
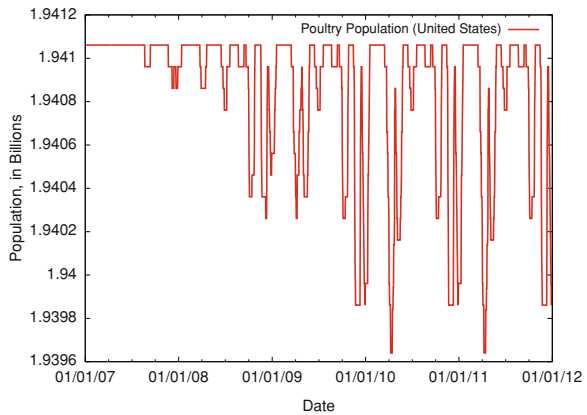


Fig. 8.6 Timelines for H5N1 infection spread from Indonesia (a known epicenter of H5N1 outbreaks) to US. The time chart also illustrates the epicenters of the intermediate infections

Fig. 8.7 Fluctuation in poultry population due to H5N1 outbreak which is controlled by mass culling of infected poultry



population reflects regeneration of poultry flocks after an outbreak. As illustrated by the graph, infections in poultry also follow a cyclic pattern that correlate with annual migration of waterfowl. The mortality figures can be translated to corresponding dollar figures for financial analysis. Additional case studies and experiments conducted using SEARUMS are discussed in the literature (Rao et al. 2009).

8.6 Conclusion

This article motivated the need for HITL-simulation steering (see Sect. 8.1.2) to effectively analyze the global epidemiology of emergent diseases. As discussed in Sect. 8.1, epidemiology is a vital discipline with far reaching impacts on human health and economics. Specifically, the article discussed the conceptual mathematical models (refer to Sects. 8.2.1 and 8.2.3) and their implementation in a custom Eco-modeling and HITL bio-simulation environment called SEARUMS. The article also discussed the software architecture of SEARUMS and its design in Sect. 8.3.

The procedure involved in utilizing SEARUMS for Eco-modeling and model calibration using HITL-simulation steering was discussed in Sect. 8.4. HITL-simulation steering played a vital role in enabling calibration of the model. The

HITL infrastructure was developed because conventional approaches for suitably seeding the model were unsuccessful despite hundreds of attempts and simulation runs. Furthermore, the ability to create checkpoints using the HITL infrastructure enabled rapid analysis of various scenarios.

The Eco-modeling and HITL-simulation infrastructure of SEARUMS has been used to model and analyze the global epidemiology of avian influenza. It must be noted that, even though the current emphasis of SEARUMS is on avian influenza, the underlying conceptual models and the software infrastructure are generic. Consequently, the concepts and software tools can be adapted for epidemiological analysis.

Some of the experiments conducted using SEARUMS were presented in Sect. 8.5 to illustrate the multi-disciplinary applicability of SEARUMS. Researchers, epidemiologists, and ornithologists can utilize HITL-simulations for rapid “what-if” types of analysis to study impacts of other factors influencing epidemics. It can be used to analyze other scenarios such as those simulated by Los Alamos National Laboratory (2006). SEARUMS and our Eco-description provide an excellent foundation for further enhancements. Note that use of SEARUMS does not require any special computing infrastructure or programming knowledge. Consequently, experts from multiple domains can collaboratively use SEARUMS to perform various types of analysis on a global scale, assess threats, and measure effectiveness of countermeasures. Our methodology and HITL-simulation environment will enable mankind to strategically invest precious time and resources to combat avian influenza, minimize its impacts on human life and global economy thereby averting a pandemic.

References

- Anderson RM, May RM (1992) *Infectious diseases of humans: dynamics and control*. Oxford University, Oxford University Press
- Bloch J (2001) *Effective Java programming language guide*, 1st edn. Pearson, Prentice Hall
- Booth G (1997) Gecko: a continuous 2d world for ecological modeling. *Artif Life* 3(3):147–163
- CDC (2006) Centers for disease control & prevention: avian Influenza: current situation. <http://www.cdc.gov/flu/avian/outbreaks/current.htm>
- Daley DJ, Gani J (2001) *Epidemic modelling: an introduction*. Cambridge University Press, Oxford
- Epstein JM (2009) Modelling to contain pandemics. *Nature* 460:687–687. doi:10.1038/nature460687a
- Ferguson NM, Cummings DAT, Fraserl C, Cajka JC, Cooley PC, Burke DS (2006) Strategies for mitigating an influenza pandemic. *Nature* 442:448–452. doi:10.1038/nature04795
- Flint SJ, Enquist LW, Racaniello VR, Skalka AM (2004) *Principles of virology*, 2nd edn. American Society for Microbiology (ASM), Washington, DC
- Garten RJ, Davis CT, Russell CA, Shu B, Lindstrom S, Balish A, Sessions WM, Xu X, Skepner E, Deyde V, Okomo-Adhiambo M, Gubareva L, Barnes J, Smith CB, Emery SL, Hillman MJ, Rivaller P, Smagala J, de Graaf M, Burke DF, Fouchier RAM, Pappas C, Alpuche-Aranda CM, Lopez-Gatell H, Olivera H, Lopez I, Myers CA, Faix D, Blair PJ, Yu C, Keene KM, Dotson JP, David, Boxrud D, Sambol AR, Abid SH, St George K, Bannerman T, Moore AL, Stringer DJ, Blevins P, Demmler-Harrison GJ, Ginsberg M, Kriner P, Waterman S, Smole S, Guevara HF,

- Belongia EA, Clark PA, Beatrice ST, Donis R, Katz J, Finelli L, Bridges CB, Shaw M, Jernigan DB, Uyeki TM, Smith DJ, Klimov AI, Cox NJ (2009) Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. *Science* 325(5937):197–201. doi:10.1126/science.1176225, <http://www.sciencemag.org/cgi/reprint/325/5937/197.pdf>
- Gilbert N, Bankes S (2002) Platforms and methods for agent-based modeling. In: Proceedings of the National Academy of Sciences of the USA, vol 99, pp 7197–7198. doi:10.1073/pnas.072079499
- GLiPHA (2007) Global livestock production and health atlas (GLiPHA): animal production and health division of food and agriculture organization of the UN. <http://www.fao.org/ag/aga/glipha/index.jsp>
- GROMS (2006) Global register of migratory species: summarising knowledge about migratory species for conservation. <http://www.groms.de/>
- Gu W, Vetter J, Schwan K (1994) An annotated bibliography of interactive program steering. *ACm SIGPLAN notices* 29(9):140–148. doi:10.1145/185009.185038
- Hagemeijer W, Mundkur T (2006) Migratory flyways in Europe, Africa, and Asia and the spread of HPAI H5N1. In: International scientific conference on avian influenza and wild birds, FAO and OIE, Rome, Italy. <http://www.fao.org/ag/againfo/subjects/en/health/diseases-cards/conference/documents/hagemeijer-mundkur.pdf>
- Halloran ME, Ferguson NM, Eubank S, Ira M Longini J, Cummings DAT, Lewis B, Xu S, Fraser C, Vullikanti A, Germann TC, Wagener D, Beckman R, Kadau K, Barrett C, Macken CA, Burke DS, Cooley P (2008) Modeling targeted layered containment of an influenza pandemic in the US. *Proc Natl Acad Sci USA* 105(12):4639–4644. doi:10.1073/pnas.0706849105, <http://www.pnas.org/cgi/doi/10.1073/pnas.0706849105>
- Hare M, Deadman PJ (2004) Further towards a taxonomy of agent-based simulation models in environmental management. *Mathematics and computers in simulation* 64(1):25–40. doi:10.1016/S0378-4754(03)00118-6
- Hufnagel L, Brockmann D, Geisel T (2004) Forecast and control of epidemics in a globalized world. *Proc Natl Acad Sci USA* 101(42):15124–15129. doi:10.1073/pnas.0308344101, <http://www.pnas.org/content/101/42/15124.abstract>
- Kilpatrick AM, Chmura AA, Gibbons DW, Fleischer RC, Marra PP, Daszak P (2006) Predicting the global spread of H5N1 avian influenza. *Proc Natl Acad Sci USA* 103(51):19368–19373. doi:10.1073/pnas.0609227103
- Lampert L (1978) Time, clocks, and the ordering of events in a distributed system. *Commun ACM* 21(7):558–565
- LANL (2006) Los Alamos National Laboratory: avian flu modeled on supercomputer, explores vaccine and isolation options for thwarting a pandemic. http://www.lanl.gov/news/index.php?fuseaction=home.story&story_id=8171
- Law R, Dieckmann U, Metz JA (2005) The geometry of ecological interactions: simplifying spatial complexity. Cambridge University Press, New York
- Longini IM, Nizam A, Xu S, Ungchusak K, Hanshaoworakul W, Cummunings DAT, Halloran ME (2005) Containing pandemic influenza at the source. *Science* 309(5737):1083–1087. doi:10.1126/science.1115717
- Normile D (2006) Avian influenza: evidence points to migratory birds in H5N1 spread. *Science* 311(5765):1225. doi:10.1126/science.311.5765.1225
- Railsback SF, Lytinen SL, Jackson SK (2006) Agent-based simulation platforms: review and development recommendations. *Simulation* 82(9):609–623. doi:10.1177/0037549706073695
- Rao DM, Chernyakhovsky A (2008) Parallel simulation of the global epidemiology of avian influenza. In: Proceedings of the 2008 winter simulation conference, pp 1583–1591
- Rao DM, Chernyakhovsky A, Rao V (2007a) SEARUMS: studying epidemiology of avian influenza rapidly using simulation. In: Proceedings of ICMHA'07. Berkeley, pp 667–673
- Rao DM, Martin Stieger DDH Jr, Johnson CB, Kidambi P, Narayanan S (2007b) Design and implementation of virtual and constructive simulations using open EAAGLES. Linus Publications, New York

- Rao DM, Chernyakhovsky A, Rao V (2009) Modeling and analysis of global epidemiology of avian influenza. *Environ Model Softw* 24(1):124–134. doi:[10.1016/j.envsoft.2008.06.011](https://doi.org/10.1016/j.envsoft.2008.06.011)
- Russell CA, Jones TC, Barr IG, Cox NJ, Garten RJ, Gregory V, Gust ID, Hampson AW, Hay AJ, Hurt AC, de Jong JC, Kelso A, Klimov AI, Kageyama T, Komadina N, Lapedes AS, Lin YP, Mosterin A, Obuchi M, Odagiri T, Osterhaus ADME, Rimmelzwaan GF, Shaw MW, Skepner E, Stohr K, Tashiro M, Fouchier RAM, Smith DJ (2008) The global circulation of seasonal influenza A (H3N2) viruses. *Science* 320(5874):340–346. doi:[10.1126/science.1154137](https://doi.org/10.1126/science.1154137)
- SEDAC (2007) SocioEconomic Data and Applications Center (SEDAC): gridded population of the world. <http://sedac.ciesin.columbia.edu/>
- Smith GJD, Vijaykrishna D, Bahl J, Lycett SJ, Worobey M, Pybus OG, MaSK Cheung CL, Raghwani J, Bhatt S, Peiris JSM, Guan Y, Rambaut A (2009) Origins and evolutionary genomics of the 2009 swine-origin h1n1 influenza a epidemic. *Nature* 459:1122–1125. doi:[10.1038/nature08182](https://doi.org/10.1038/nature08182)
- Solow AR, Smith WK (2006) Using markov chain successional models backwards. *J Appl Ecol* 43(1):185–188. doi:[10.1111/j.1365-2664.2005.01127.x](https://doi.org/10.1111/j.1365-2664.2005.01127.x)
- Tobias R, Hofmann C (2004) Evaluation of free java-libraries for social-scientific agent based simulation. *J Artif Soc Soc Simul(JASS)* 7(1). <http://jass.soc.surrey.ac.uk/7/1/6.html>
- USCB (2006) U.S. Census Bureau: 100 largest counties. <http://www.census.gov/popest/counties/CO-EST2006-07.html>
- WHO (2005) World Health Organization: ten things you need to know about pandemic influenza. <http://www.who.int/csr/disease/influenza/pandemic10things/>
- WHO (2006a) World Health Organization: avian influenza: updates. http://www.who.int/csr/disease/avian_influenza/updates/en/
- WHO (2006b) World Health Organization: H5N1 avian influenza: timeline. http://www.who.int/entity/csr/disease/avian_influenza/timeline.pdf
- WHO (2006c) World Health Organization (WHO): avian influenza. http://www.who.int/csr/disease/avian_influenza/en/
- WHO (2007a) A description of the process of seasonal and H5N1 influenza vaccine virus selection and development. http://www.who.int/entity/csr/disease/avian_influenza/Fluvac-cvirusselection.pdf
- WHO (2007b) World Health Organization: avian influenza: situation updates. <http://www.who.int/entity/csr/disease/influenza/H5N1-9reduit.pdf>
- Winston WL (1994) *Operations research (applications and algorithms)*, 3rd edn. Duxbury Press, Belmont
- Wolfram MathWorld (2006) Mercator projection. <http://mathworld.wolfram.com/MercatorProjection.html>