Environmental Modelling & Software 24 (2009) 124–134

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Modeling and analysis of global epidemiology of avian influenza

Dhananjai M. Rao [a,*], Alexander Chernyakhovsky [b], Victoria Rao [c]

[a] CSA Department, Miami University, 205 Benton Hall, Oxford, OH 45056, USA
[b] Mason High School, Mason, OH 45040, USA
[c] Cybernetic Evolution Inc., Mason, OH 45040, USA

ARTICLE INFO

Article history:
Received 18 September 2007
Received in revised form 14 June 2008
Accepted 15 June 2008
Available online 9 August 2008

Keywords:
Influenza
H5N1
Spatially explicit model
Agent-based model
Discrete event simulation

ABSTRACT

The World Health Organization has activated a global preparedness plan to improve response to avian influenza outbreaks, control outbreaks, and avoid an H5N1 pandemic. The effectiveness of the plan will greatly benefit from identification of epicenters and temporal analysis of outbreaks. Accordingly, we have developed a simulation-based methodology to analyze the spread of H5N1 using stochastic interactions between waterfowl, poultry, and humans. We have incorporated our methodology into a user friendly, extensible software environment called SEARUMS. SEARUMS is an acronym for Studying the Epidemiology of Avian Influenza Rapidly Using Modeling and Simulation. It enables rapid scenario analysis to identify epicenters and timelines of H5N1 outbreaks using existing statistical data. The case studies conducted using SEARUMS have yielded results that coincide with several past outbreaks and provide non-intuitive inferences about global spread of H5N1. This article presents the methodology used for modeling the global epidemiology of avian influenza and discusses its impacts on human and poultry morbidity and mortality. The results obtained from the various case studies and scenario analyses conducted using SEARUMS along with verification experiments are also discussed. The experiments illustrate that SEARUMS has considerable potential to empower researchers, national organizations, and medical response teams with timely knowledge to combat the disease, mitigate its adverse effects, and avert a pandemic.

1. Introduction and motivation

Avian influenza commonly refers to the disease caused by H5N1, a highly virulent strain of the influenza-A virus (CDC, 2006; Normile, 2006d; United States Department of Agriculture, 2007; WHO, 2006c, 2005). It is known that the virus has become endemic to waterfowl in certain areas and it readily transmits within the Anatidae family, primarily through contaminated feed and feces (Chen et al., 2005; CDC, 2006; Normile, 2006c; Liu et al., 2005). The virus has a devastating impact on poultry causing 100% mortality within 48 h of infection (WHO, 2007). Moreover, the pathogen also spreads to humans through direct contact with infected poultry and contaminated surfaces (WHO, 2007). Human-to-human transmission of the disease has been noted but it has been rare and unsustained (CDC, 2006; Normile, 2006b,c; WHO, 2007). However, the number of human fatalities has almost doubled in 2006 with 80 human deaths versus 42 in 2005 (Brahmbhatt, 2006). In spring 2006, it has been established that the disease is spread to other parts of the world by infected migrating waterfowl (Chen et al., 2005; Liu et al., 2005; Normile, 2006a). The aforementioned factors indicate that all prerequisites to start an H5N1 pandemic have been met with the exception of sustained human-to-human transmission (WHO, 2007; Normile, 2006b; Zambon, 2007).

In the past, pandemics have taken the world largely by surprise (WHO, 2007; Zambon, 2007). However, in the case of avian influenza the world has been given a clear warning, precious time, and a unique opportunity to defend itself. Accordingly, in 2004, the World Health Organization (WHO) activated a pandemic preparedness plan, alerted its network laboratories, and placed response teams on standby (WHO, 2007; Normile, 2006b). WHO's plan includes the following three objectives: (i) avert a pandemic; (ii) control the outbreaks in humans; and (iii) conduct the research to monitor the situation and improve preparedness, including development of vaccines (WHO, 2007).

Vaccines are the primary mechanism for the prevention of influenza (CDC, 2006; WHO, 2007). Unfortunately, a myriad of technological and socio-political issues have rendered manufacturing and distribution of H5N1 vaccine a significant challenge (CDC, 2006; WHO, 2007). Some of these issues are: (i) rapid mutations in the virus make it practically infeasible to manufacture large volumes of vaccine; (ii) reverse genetic techniques that are employed for vaccine manufacturing give raise to serious bio-safety concerns in addition to significant economic, political,

* Corresponding author. Tel.: +1 513 529 0335.
E-mail address: raodm@muohio.edu (D.M. Rao).

1364-8152/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.envsoft.2008.06.011

and intellectual property issues; (iii) current vaccine manufacturing facilities do not have the capacity to meet the needs of even small seasonal epidemics let alone a pandemic; and (iv) targeted distribution of limited quantities of the vaccine is a huge challenge in itself (WHO, 2007; Normile, 2006b).

The current level of preparedness and aforementioned issues require international planning and strategic deployment of suitable countermeasures. The effectiveness of the current strategies, including WHO's preparedness plan, will greatly benefit from prediction and forecasting of epicenters and global timelines of H5N1 outbreaks. In addition, detailed scenario analysis is necessary to plan suitable prophylaxis strategies to contain outbreaks. Currently, there are no methodologies to analyze the intercontinental spread of H5N1 via migrating waterfowl and its associated impacts on poultry and humans.

In an endeavor to empower researchers, governments, and health organizations with the knowledge required to avert a pandemic, we have developed a groundbreaking methodology based on computer simulations to study epidemiology of avian influenza. It allows prediction of chronology of outbreaks, identification of high risk epicenters of the disease, and analysis of economic impacts. Our simulations capture the complex, symbiotic interactions between migrating waterfowl, poultry, and humans. The conceptual model for simulation is based on Markov processes. Parameters for the model are determined using real-world statistical data on:

- Waterfowl migration published by Global Register Of Migratory Species (GROMS) (GROMS, 2006).
- Waterfowl species that are at higher risk to carry virus (Brahmbhatt, 2006),
- Global poultry population and distribution published by the Food and Agriculture Organization (FAO) of the United Nations (GLiPHA, 2007).
- Global human population data published by the SocioEconomic Data and Applications Center (SEDAC) (SEDAC, 2007).
- Information on major metropolitan areas in United States obtained from the Census Bureau (USCB, 2006).

We have incorporated our methodology into a graphical, portable, extensible, scalable, and intuitive simulation environment called SEARUMS (Rao et al., 2007). SEARUMS is an acronym for Studying the Epidemiology of Avian Influenza Rapidly Using Modeling and Simulation. SEARUMS is a multi-disciplinary modeling, simulation, and analysis environment that seamlessly integrates knowledge from various fields so that epidemiologists, economists, and disease control centers can collaboratively use it and combat the disease.

The remainder of this article presents our methodology, simulation environment, and the case studies conducted as a part of our research. Section 2 compares and contrasts our research to some of the closely related investigations reported in the literature. This section also motivates the design rationale for developing SEARUMS. Section 3 presents the conceptual model based on Markov processes used to characterize the epidemiology of avian influenza. This section is followed by Section 4 that presents a brief overview of SEARUMS, our modeling, analysis, and simulation environment. The discussion of SEARUMS is brief because the focus of this article is on the conceptual model and case studies conducted using SEARUMS. However, a detailed description of the architecture of SEARUMS along with a discussion on its design and implementation issues are available in the literature (Rao et al., 2007). Our case studies along with some significant findings are discussed in Section 5. This section also presents the experiments conducted to verify and validate the model used for conducting the case studies. Section 6 concludes this article by summarizing our findings and provides pointers to future research.

2. Related research

The conceptual, declarative model of avian influenza has been developed using a combination of Markov processes and spherical geometry. The Markov processes describe the epidemiological states of the three salient entities, namely: waterfowl, poultry, and humans. Spherical geometry, based on great-circle distances, has been used to model and trigger interactions between the entities. The Markov processes reflect the classical SIR (Susceptible–Infected–Removed) mathematical model used to model epidemiology of various diseases (Anderson and May, 1992). Specifically, the states of each Markov process reflect the susceptible, infected, and removed states in the classical SIR model shown in Fig. 1. The characteristics of the SIR model is represented using the following system of differential equations:

$$\frac{dX}{dt} = \mu N - [\lambda(t) + \mu]X(t) \tag{1}$$

$$\frac{dY}{dt} = \lambda X - (\nu + \mu)Y(t) \tag{2}$$

$$\frac{dZ}{dt} = \nu Y - \mu Z(t) \tag{3}$$

where, $X(t)$, $Y(t)$, and $Z(t)$ represent the number of susceptible, infected, and removed hosts at any given instant of time t. The parameters μ, λ, and ν are the per capita host birth/death rate, the force of infection, and per capita recovery rate, respectively. Adaptations of the aforementioned classical SIR model have been used to model the epidemiology of influenza (Longini et al., 2005; Ferguson et al., 2006). Section 2.1 presents an overview of the closely related investigations, in addition to comparing and contrasting them to our approach.

In conjunction with the aforementioned temporal state transitions, we have also introduced state transitions in the Markov processes to model spatial interactions between entities. Spatial interactions between a pair of entities arise when the areas occupied by the entities overlap. Since space is an explicit axis in the model, our mathematical model falls under the category of spatially explicit models (Law et al., 2005; Railsback et al., 2006; Rao and Wilsey, 2005). Accordingly, the modeling and simulation infrastructure of SEARUMS leverages individual, Agent-based Spatially Explicit (ASE) modeling methodology and Discrete Event Simulation (DES) technique to implement our conceptual model. Impetus for such a Modeling and Simulation (M&S) approach, involving ASE models and DES, stems from its advantages (Law et al., 2005; Railsback et al., 2006; Rao and Wilsey, 2005) and popularity (Law et al., 2005; Railsback et al., 2006). A few ecological M&S

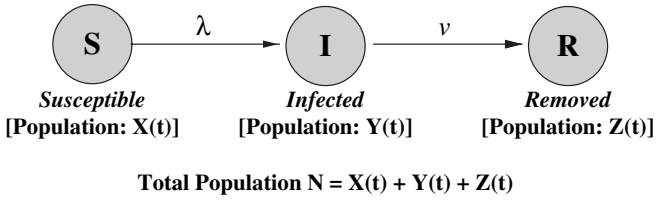

Fig. 1. Overview of the classical SIR (Susceptible–Infected–Removed) mathematical model used in epidemiological analysis.

environments and frameworks that utilize such an approach have been reported. Some of the related M&S approaches and software, that are similar in philosophy to SEARUMS, are briefly described in Section 2.2. Furthermore, Section 2.2, also compares and contrasts existing M&S with SEARUMS to elucidate some of the factors motivating the design and implementation of SEARUMS.

2.1. Related epidemiological investigations

Avian influenza, caused by a highly pathogenic virus called H5N1, continues to pose a serious public health threat (CDC, 2006; WHO, 2007). Currently, epizootic outbreaks, particularly in poultry farms, are not expected to diminish (CDC, 2006). Moreover, many researchers believe the virus has the potential to mutate into a pandemic form and readily transmit between humans in a sustained manner (Longini et al., 2005; Ferguson et al., 2006). Consequently, research into epidemiology of avian influenza has gained significant momentum in the recent past.

Prophylactic analysis of the pandemic mode of the H5N1 virus is an active area of research and investigation. Longini et al. (2005) and Ferguson et al. (2006) analyze various potentially feasible intervention strategies (Longini et al., 2005; Ferguson et al., 2006) using Thailand as an example. Recently, Halloran et al. (2008) analyzed and reported the use of several Targeted Layered Containment (TLC) methods using detailed, spatially explicit SIR models of metropolitan areas similar to Chicago (Halloran et al., 2008). These investigations analyze the use of prophylaxis, quarantine, isolation, and social distancing to control influenza epidemics. In these studies, the stochastic nature of the system is analyzed by simulating a large number of model realizations involving various parameters.

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 on SIR concepts. In their models, each individual human is modeled using an agent. The agents are created using detailed census data, demographic information, and social network data from the Thai census bureau. The human agents are heuristically seeded to account for household size, generational age structure within household, school sizes, workplace data, and travel data. The infection parameters for the virus such as latent period, infectiousness over time, and transmission parameters have been estimated from data available for prior influenza outbreaks. 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. However, such a scenario continues to remain only a possibility at the time of this writing.

Unlike the investigations reported by Longini et al. (2005), Ferguson et al. (2006), and Halloran et al. (2008) our research assumes and reflects the current and more realistic situation, i.e., H5N1 is yet to mutate into its pandemic state and human-to-human transmission is unsustained. Consequently, we do not model human interactions in great detail. Instead, we emphasize seasonal migration patterns of waterfowl that are the primary vectors for intercontinental spread of the disease. Furthermore, our research also accounts for epizootic outbreaks in poultry farms and subsequent sporadic transmission to humans. The aforementioned aspects notably distinguish our efforts from those reported by Longini et al. (2005), Ferguson et al. (2006), and Halloran et al. (2008). However, similar to these three investigations, we also use the classical SIR mathematical models along with agent-based, spatially explicit models.

Upadhyaya et al. (2008) present a statistical transmission of avian influenza based on investigations of a week-long outbreak in India. Their non-linear model includes parameters on poultry transportation distance, lifetime of the virus, immunity to the disease, and a threshold value. Their simulation-based empirical evaluations indicate that their model yields satisfactory results for a given set of parameter values. Iwami et al. (2007) present a SIR-based mathematical model for epidemic and pandemic mode of the virus along with proofs to support inferences drawn from their model. In contrast to these two investigations, our agent-based, spatially-explicit SIR model uses Markov processes to model various epidemiological scenarios. Moreover, in our model, waterfowl migration is considered as the dominant factor causing spread of the virus while the investigations by Upadhyaya et al. (2008) and Iwami et al. (2007) focus on poultry trade.

2.2. Related software environments

Several different approaches can be used to develop a software implementation of a conceptual spatially explicit model. We surveyed several implementation alternatives prior to developing our custom modeling, simulation, visualization, and analysis environment called SEARUMS. Some of the related modeling & simulation approaches and software, that were considered in our survey are briefly described in this section. Furthermore, the section compares and contrasts the surveyed environments with SEARUMS to elucidate the important factors motivating the design and implementation of SEARUMS.

Gilbert and Bankes (2002), Railsback et al. (2006), and Tobias and Hofmann (2004) have reviewed five popular, general purpose software platforms for scientific, agent-based modeling and simulation (Railsback et al., 2006). The common platforms considered 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. Therefore, it poses interoperability issues. On the other hand, the latter four platforms use 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.

SWARM, written in Objective-C, is reported to be one of the most mature environments (Gilbert and Bankes, 2002; Railsback et al., 2006). However, Objective-C is not widely used is not native to all platforms, and lacks developer tools (Railsback et al., 2006). SWARM-Java provides a Java interface to SWARM's Objective-C libraries to improve portability and interoperability. Nevertheless, Railsback et al report that Java SWARM does not combine the advantages of the two languages well. Furthermore, SWARM was reported to have poor performance for larger models (Railsback et al., 2006). Hence SWARM and SWARM-Java were not viable candidates for developing SEARUMS, that we envision to be fast, portable, and interoperable. On the other hand, Repast and MASON are pure Java-based environments with slightly different design objectives. The focus of Repast is primarily on the domain of social science and includes tools specific to that domain (Gilbert and Bankes, 2002; Tobias and Hofmann, 2004; Railsback et al., 2006). Although highly rated, Repast is tightly coupled and dependent on the Eclipse framework. This makes Repast large with voluminous Application Program Interface (API) making it a complicated system with daunting learning curves to effectively extend and implement all the envisioned features of SEARUMS.

MASON, on the other hand, has been designed as a smaller and faster alternative to Repast with focus on computationally demanding models (Luke et al., 2005). Although MASON is the fastest simulator amongst these software environments, it is the least mature (Railsback et al., 2006). Furthermore, MASON does not provide a native threading capability like Repast. Instead, the responsibility to achieve multi-threading is left to the modeler (Luke et al., 2005). In this scheme concurrent events (discrete events with the same timestamp) for various agents would not be automatically scheduled to execute in parallel without significant kludges. Lack of fine grained, event-level parallelism prevents effective use of inherent parallelism in the model to accelerate simulations using multiple threads. Consequently, the design of SEARUMS was steered away from MASON.

The recently standardized, High Level Architecture (HLA) was also explored as a candidate to aid design and implementation of SEARUMS. HLA is a general purpose Application Program Interface (API) for developing cross-platform, multi-language, distributed simulations. The standard is mandated and widely used by the Department of Defense (DoD). Unfortunately, it requires a Runtime Infrastructure (RTI) software and a federate library that is no longer freely available from the Department of Modeling and Simulation (DMSO). Commercial RTI implementations can be purchased but they can be cost prohibitive and therefore this approach was not adopted. Rao and Wilsey (2005) present a web-based, parallel environment called WESE to analyze the epidemiology of Lyme disease (Rao and Wilsey, 2005). WESE has been developed in C++ and provides a framework for developing agents (Rao and Wilsey, 2005). It also supports multi-scale, parallel simulations. However, it is designed for batch simulations and does not provide an effective mechanism to interface with a GUI.

The aforementioned drawbacks of various software systems motivated us to custom develop SEARUMS. The objective was to minimize learning curves for both developers and users, maximize portability, include intuitive interfaces for modeling, and seamlessly incorporate analysis tools. These objectives have been realized through the design and implementation of SEARUMS as discussed in the next section.

3. Conceptual, mathematical model

Global epidemiological analysis of avian influenza is still nascent and several aspects of the disease have not yet been formalized. Only a few highly specific prescriptive models of the global epidemiology of the disease have been reported thus far. However, individual life cycle and behavior of the three primary biological entities are relatively well understood. The three biological entities being waterfowl flocks, poultry flocks, and human groups. Note that each entity is not an individual bird or human but a flock of birds or a group of humans with common characteristics. For example, a specific species of waterfowl that live and migrate as a large flock in real life are considered as a single entity. Poultry entities represent a large collection of birds such as a poultry farm. In an analogous manner, humans living in geographic proximity to each other are considered as a single entity. Such a representation is necessary to reduce the size and complexity to more tractable scales.

In the conceptual model, the temporal characteristics of individual entities has been modeled using discrete-time Markov process (Solow and Smith, 2006; Winston, 1994). Life cycle events of the entities are modeled via probabilistic state changes occurring within each Markov process. The spatial interactions between the entities are modeled using principles of spherical geometry. Accordingly, the Earth's surface has been modeled as a sphere. Such an approach has been employed to reduce the mathematical and computational complexity of ecological models (Law et al., 2005; Winston, 1994; Booth, 1997; Hare and Deadman, 2004; Wolfram MathWorld, 2006). Spatial interactions essentially alter the probabilities of state transitions impacting the life cycle activities. The spatial interactions and many temporal state transitions are common characteristics of all three Markov processes. Consequently, these common model characteristics and associated formalisms are initially presented in Section 3.1. The different Markov processes for waterfowl, poultry, and humans are discussed in Sections 3.2–3.4, respectively.

3.1. Common characteristics of various Markov processes

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 $X_0, X_1, X_2, \ldots$, 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: (i) a Markov process may be in only one given state at any instant of time; (ii) transition from one state to another occurs instantaneously in discrete-time steps; and (iii) 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). Formally, this is defined using the following conditional probability relationship:

$$\Pr(X_{n+1} = x | X_t = x_t, \ldots, X_1 = x_1, X_0 = x_0) = \Pr(X_{t+1} = x | X_t = x_t) \quad (4)$$

In our research, the SIR (Susceptible–Infected–Removed) life cycles of the three main entities namely: waterfowl ($\mathrm{MP_{wf}}$), poultry ($\mathrm{MP_{po}}$), and humans ($\mathrm{MP_{hu}}$) have been modeled as three different Markov processes represented by the set $\mathrm{MP} = \{\mathrm{MP_{wf}}, \mathrm{MP_{po}}, \mathrm{MP_{hu}}\}$. The Ecosystem of interest M is modeled as a set of interacting instances of these three types of Markov processes represented as $\mathbb{M} = \{\mathrm{mp}_1, \mathrm{mp}_2, \ldots, \mathrm{mp}_n\}(\mathbb{M} \neq \{\phi\}$ and $\forall \mathrm{mp} \in \mathbb{M}, \{\mathrm{mp}\} \subset \mathrm{MP})$. The temporal state of each Markov process is represented by the 5-tuple, $\forall \mathrm{mp} \in \mathbb{M}$, $S_t^{\mathrm{mp}} = < X_t^{\mathrm{mp}}, Y_t^{\mathrm{mp}}, R_t^{\mathrm{mp}}, I_t^{\mathrm{mp}}, P_t^{\mathrm{mp}} >$. The variables $X_t^{\mathrm{mp}}, Y_t^{\mathrm{mp}}$, and R_t^{mp} collectively model the geographic and migratory characteristics. The state variable I_t^{mp} represents a measure of H5N1 infection in entity $(\mathrm{mp} \in \mathbb{M})$ at time t. Recollect that entities represent a large collection of individuals. For example, a waterfowl entity represents a large flock of birds. The actual population of individuals represented by an entity is represented using the time varying state variable P_t^{mp}. Variations with time permit modeling birth and attrition of birds due to disease and other natural causes as per the SIR model discussed in Section 2.

The X_t^{mp} and Y_t^{mp} state variables represent the instantaneous geographic attributes of entity mp at time t. These values correspond to the longitude and latitude of an entity. The state variable R_t^{mp} represents the radius of the approximated circular surface area occupied by the given entity. Such approximations are widely used to reduce the mathematical and computational complexity of ecological models (Law et al., 2005; Winston, 1994; Booth, 1997; Hare and Deadman, 2004; Wolfram MathWorld, 2006). Migratory behavior of a process $\mathrm{mp}(\mathrm{mp} \in \mathbb{M})$ is reflected by periodic, time-dependent changes to its state (S_t^{mp}) variables X_t^{mp} and Y_t^{mp}. The values for X_t^{mp} and Y_t^{mp} are selected from a predefined migration path. The migration path is represented by a digraph $G(V_{\mathrm{mt}}, E)$ consisting of a set of migration-time dependent vertexes V_{mt}

connected by directed edges E. Information regarding location (latitude and longitude) of vertexes, migration-times (mt), and direction of edges required to develop the digraph are obtained by piecewise linearization (Law et al., 2005; Rao, 2003) of the migratory flyways and timelines published by various organizations (GROMS, 2006; Hagemeijer and Mundkur, 2006).

State transitions also occur due to transmission of infection via interactions between processes. Interactions occur when the time-dependent neighborhood of a process $N_t^{\mathrm{mp}}(\forall \mathrm{mp} \in \mathbb{M})$ is a non-empty set as per the following equation:

$$\begin{aligned} \forall \mathrm{mp} \in \mathbb{M}, \quad N_t^{\mathrm{mp}} &= \{\mathrm{mp}_a, \mathrm{mp}_b, \ldots, \mathrm{mp}_z\}\left(|N_t^{\mathrm{mp}}| \geq 0\right) \\ &\quad \text{and} \quad \forall \mathrm{nmp} \in N_t^{\mathrm{mp}} \quad \mathrm{nmp} \in \mathbb{M} \\ &\quad - \{\mathrm{mp}\}\left(N_t^{\mathrm{mp}} \subset \mathbb{M}\right) \end{aligned} \tag{5}$$

$$2\tan^{-1}\left(\frac{\sqrt{a}}{\sqrt{1-a}}\right)R \leq \left[R_t^{\mathrm{nmp}} + R_t^{\mathrm{mp}}\right] \tag{6}$$

where,

$$\begin{aligned} a &= \sin^2\left(\tfrac{\Delta\mathrm{lat}}{2}\right) + b\sin^2\left(\tfrac{\Delta\mathrm{lon}}{2}\right) \\ b &= \cos\left(Y_t^{\mathrm{mp}}\right) \quad \cos\left(Y_t^{\mathrm{nmp}}\right) \\ \Delta\mathrm{lat} &= Y_t^{\mathrm{mp}} - Y_t^{\mathrm{nmp}} \\ \Delta\mathrm{lon} &= X_t^{\mathrm{mp}} - X_t^{\mathrm{nmp}} \end{aligned}$$

R=Radius of earth (3958.75587 miles).

Eq. (6) defines the instantaneous neighborhood N_t^{mp} of a process $\mathrm{mp} \in \mathbb{M}$ at time t to be a subset of processes such that their areas overlap. The distances between entities are the great-circle distance computed using the great-circle distance formula. Using the time varying neighborhood function, transmission of infection is modeled as a change in the measure of infection of the interacting entities as per the following equation:

$$\begin{aligned} &\forall \mathrm{nmp} \in N_t^{\mathrm{mp}}, \\ &I_t^{\mathrm{nmp}} = \begin{cases} I_t^{\mathrm{nmp}} & \text{if } I_t^{\mathrm{mp}} = 0 \\ I_f(\mathrm{mp}, \mathrm{nmp}) & \text{otherwise} \end{cases} \end{aligned} \tag{7}$$

Eq. (7) states that if a given entity is not infected, i.e., $I_t(\mathrm{MP}_i) = 0$, then it does not spread infection to its neighborhood. However, if the entity is infected, then it spreads infection to its neighbors based on an abstract function $I_f(\mathrm{mp}, \mathrm{nmp})$. In our current implementation, I_f is defined as the maximum percentage of overlap between mp and nmp occurring in consecutive set of interactions. In the case of poultry processes, the infection percentage is also used to cull poultry after 48 h of infection (WHO, 2007). The specific characteristics and state transitions for each type of Markov process are presented in the following subsections.

3.2. Markov process for waterfowl flock

The Markov process representing the SIR life cycle and migration of a generic waterfowl flock ($\mathrm{MP}_{\mathrm{wf}}$) is shown in Fig. 2. As shown in the figure, the Markov process consists of three states, namely: x_0, x_1, and x_2. The significance of each state and the probabilistic transitions between them are as follows:

- *State* x_0 (*Susceptible*): This state is the initial state in the life cycle of a waterfowl. In this state, the waterfowl is not infected with the avian influenza virus. In this state, the process is set to transition back to state x_0 each time the waterfowl changes its coordinates $(X_t^{\mathrm{mp}}, Y_t^{\mathrm{mp}})$ at discrete-time t to reflect its migration. However, if the bird comes in contact with an infected flock and gets infected $(I_t^{\mathrm{mp}} > 0)$ then it transitions to the infected state x_1.
- *State* x_1 (*Infected*): This state indicates the state in which the waterfowl is infected and the infection is spread within the flock. Note that once the flock is in this state it never transitions back to the non-infected (x_0) state as the virus becomes endemic. Accordingly, there is no removed state in this Markov process as the disease does not cause mortality in waterfowl. Similar to the earlier state, transitions occur to reflect migration of the waterfowl by changing its coordinates $(X_t^{\mathrm{mp}}, Y_t^{\mathrm{mp}})$. Moreover, when the flock comes in contact with other flocks $(|N_t^{\mathrm{mp}}| > 0)$ it transitions to state x_2 to reflect inter-flock infection transmissions.
- *State* x_2 (*Transmitting*): A flock of waterfowl remains in this state only for a brief period of time to represent inter-flock infection transmission. In this state, it spreads the infection to other entities in its neighborhood as per Eq. (7) may they be waterfowl, poultry, or humans. Recollect that the neighborhood of the flock is determined by Eq. (6). Once the infection transmission phase has been completed, the flock immediately transitions back to x_1 state (see Fig. 2).

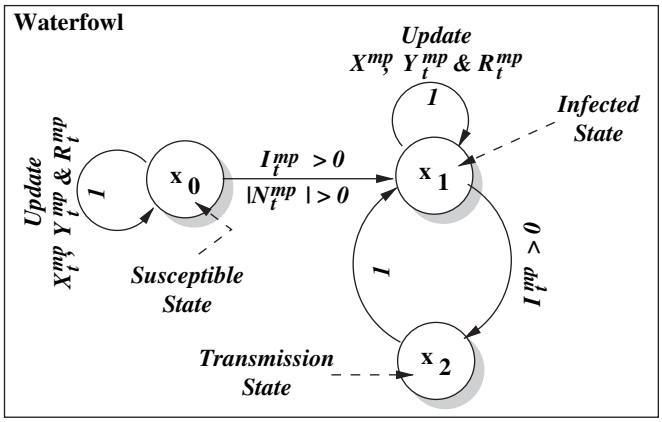

Fig. 2. Overview of Markov process for a Waterfowl Flock ($\forall \mathrm{mp} \in \mathbb{M}$ and $\{\mathrm{mp}\} \subset \{\mathrm{MP}_{\mathrm{wf}}\}$) illustrating the states through which the process transitions to model the behavioral life cycle of waterfowl. The state of the process (S_t^{mp}) at discrete-time t is represented by the 4-tuple $S_t^{\mathrm{mp}} = < X_t^{\mathrm{mp}}, Y_t^{\mathrm{mp}}, R_t^{\mathrm{mp}}, I_t^{\mathrm{mp}} >$, 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.

3.3. Markov process for poultry

The Markov process representing the salient SIR aspects of a poultry flock's ($\mathrm{MP}_{\mathrm{po}}$) life cycle is illustrated in Fig. 3. As illustrated by the figure, the process is similar to the waterfowl process and it transitions through three states. The poultry process begins in the initial susceptible state State x_0. In this state, the poultry is not infected and there is no attrition in population. However, the population may periodically increases to model regeneration of poultry after culling or death that occurs in the infected state x_1. Regeneration occurs only if the poultry population is below a given threshold value. As shown in Fig. 3, on being infected $(I_t^{\mathrm{mp}} > 0)$ due to other infected entities in its neighborhood N_t^{mp} the processes transitions to the infected state x_1. A poultry flock typically remains only for a short period of time in this state until the infected birds die or they are intentionally culled to control spread of the disease. Once all the infected birds have died (i.e., $I_t^{\mathrm{mp}} = 0$), the process transitions back to the non-infected x_0 state (see Fig. 3). However, if the neighborhood of the flock as defined by Eq. (6) is non-zero, then

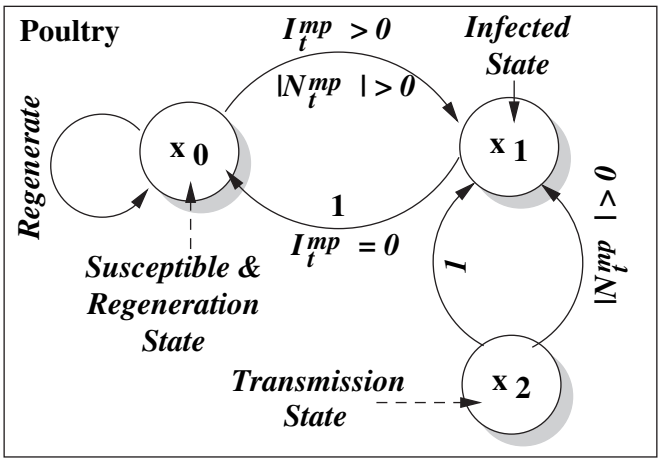

Fig. 3. Overview of Markov process for Poultry flocks ($\forall mp \in \mathbb{M}$ and $\{mp\} \subset \{MP_{po}\}$) illustrating the states through which the processes transitions to model its SIR behavioral life cycles. The state of the process (S_t^{mp}) at discrete-time t is represented by the 4-tuple $S_t^{mp} = < X_t^{mp}, Y_t^{mp}, R_t^{mp}, I_t^{mp} >$, where X_t^{mp} is the longitude, Y_t^{mp} is the latitude, R_t^{mp} is the radius. I_t^{mp} is the current measure of infection in the entity. The set N_t^{mp} represents the neighborhood at time t.

it temporarily transitions through state x_2 to model the inter-flock disease transmissions. Once the infection transmission phase has been completed, the flock immediately transitions back to x_1 state.

3.4. Markov process for humans

Fig. 4 illustrates the Markov process representing humans (MP_{hu}). State x_0 is the initial state in which the human Markov process (MP_{hu}) commences. In this state the humans represented by the process are assumed to be disease free but susceptible to infection. However, the process transitions to state x_1, when an infection is transmitted via interactions with poultry or waterfowl in the neighborhood. State x_1 is used to represent the healing phase after infection. This state also includes possible human fatalities that may occur. The third state x_2 is used to model epidemic state of the infection where the H5N1 readily transmits from humans to

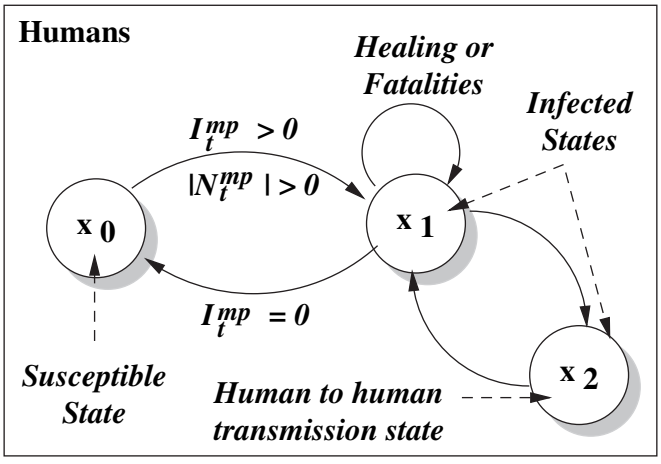

Fig. 4. Overview of Markov process for Human groups ($\forall mp \in \mathbb{M}$ and $\{mp\} \subset \{MP_{hu}\}$) illustrating the states through which the processes transitions to model its SIR behavioral life cycles. The state of the process (S_t^{mp}) at discrete-time t is represented by the 4-tuple $S_t^{mp} = < X_t^{mp}, Y_t^{mp}, R_t^{mp}, I_t^{mp} >$, where X_t^{mp} is the longitude, Y_t^{mp} is the latitude, R_t^{mp} is the radius. I_t^{mp} is the current measure of infection in the entity. The set N_t^{mp} represents the neighborhood at time t.

humans. When the processes periodically transitions to state x_2 where the infection increases. Once the infection in state x_1, the human process drops back to zero (i.e., $I_t^{hu} = 0$), the process transitions back to the initial x_0 state. Currently, our cases studies do not utilize state x_2 where sustained human-to-human transmissions occur in order to accurately model current real-world epidemics (Normile, 2006c).

4. SEARUMS: the modeling, simulation, and analysis environment

SEARUMS is an acronym for Studying Epidemiology of Avian influenza Rapidly Using Modeling and Simulation. It is a user friendly, integrated, graphical modeling, simulation, and visualization environment for study and analysis of epidemiology of avian influenza and the impacts of the disease. SEARUMS has been developed from the ground-up, in Java, due to the reasons discussed in Section 2. SEARUMS provides an intuitive research flow, from initial setup to data visualization, via user friendly graphical interface. SEARUMS essentially implements our descriptive, conceptual model presented in Section 3. Each Markov process in the mathematical model is implemented as a smart agent using the Application Program Interface (API) provided by SEARUMS.

The overall ecological model is represented using a collection of interacting smart agents. The agent-based, spatially explicit model embodies the epidemiology of H5N1 virus (CDC, 2006; WHO, 2005). SEARUMS uses a multi-threaded discrete event simulation (Rao, 2003; Rao and Wilsey, 2005) engine to efficiently implement the mathematical model and effectively capture the stochastic, symbiotic interactions between the individuals in the Ecosystem, namely: migrating waterfowl, poultry, and humans (CDC, 2006; WHO, 2005). The modeling methodology and its implementation in SEARUMS enables efficient use of real-world statistical data of H5N1 outbreaks (WHO, 2006b) along with waterfowl migration flyways (GROMS, 2006) to predict intercontinental transmission pathways, timelines, epicenters of outbreaks, and economic impacts of avian influenza.

Since the focus of this article is on the mathematical model and experimental analysis using SEARUMS, only a brief description of SEARUMS, pertinent to the scope of this article is presented in this section. However, a detailed description of the design and implementation of the software is available in the literature (Rao et al., 2007). Readers are referred to the literature and the SEARUMS website (SEARUMS, 2008) for details on its design, implementation, installation, and use. An overview of the procedure for research and analysis using SEARUMS is illustrated in Fig. 5. The research flow with SEARUMS is broadly classified into the following three phases: Phase 1: Development of Eco-description using graphical interface, Phase 2: Simulation and Data collection phase and Phase 3: Data Visualization and Analysis phase that partially overlaps with Phase 2.

In Phase 1, the Eco-description for simulation-based study and analysis is developed. The Eco-description consists of a collection of smart *agent* (Hare and Deadman, 2004) instances that model the behavior of entities involved in the epidemiology of avian influenza (CDC, 2006; WHO, 2005). The agents are organized into an Agent Repository for rapid access, extensibility, and reuse. Currently, SEARUMS includes the following three smart agents: Waterfowl 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 different exposed attributes via the Attribute Editor module. The screenshot of SEARUMS shown in Fig. 6 illustrates the Graphical User Interface (GUI) presented by the Attribute Editor. As

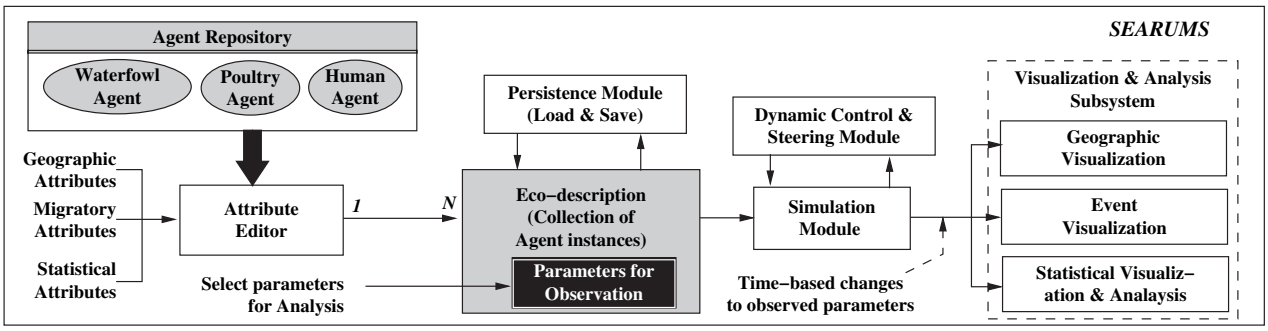

Fig. 5. The figure presents an architectural overview of SEARUMS. It illustrates the research flow and interactions between the modules constituting SEARUMS. A freely available version of SEARUMS for use and validation by the scientific community is also included in the supporting materials. Refer to SEARUMS website for latest software updates (SEARUMS, 2008).

indicated in Fig. 5, the attributes of an agent include geographic attributes, migratory attributes, and statistical attributes.

The geographic attributes indicate the location (latitude and longitude) and radius of each entity. Entities are displayed as circles on the graphical view as shown in Fig. 6. In addition, entities are logically placed into hierarchical, intersecting sets called groups. Groups are used to categorize and classify entities to facilitate aggregate analysis. For example, poultry in a country can be placed in a single group to analyze impacts at a country level. Similarly, a variety of hierarchy of groups can be used to collate and analyze statistical data. Groups are created and managed via the Group Editor module shown in Fig. 6.

The migratory attributes are specified only for agents whose location changes over the lifetime of the simulation. In our studies, migratory attributes have been specified only for waterfowl instances. However, the feature can be applied to human and poultry instances as well to reflect travel and trade. 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. We have obtained migration attributes from the Global Register of Migratory Species (GROMS) (GROMS, 2006).

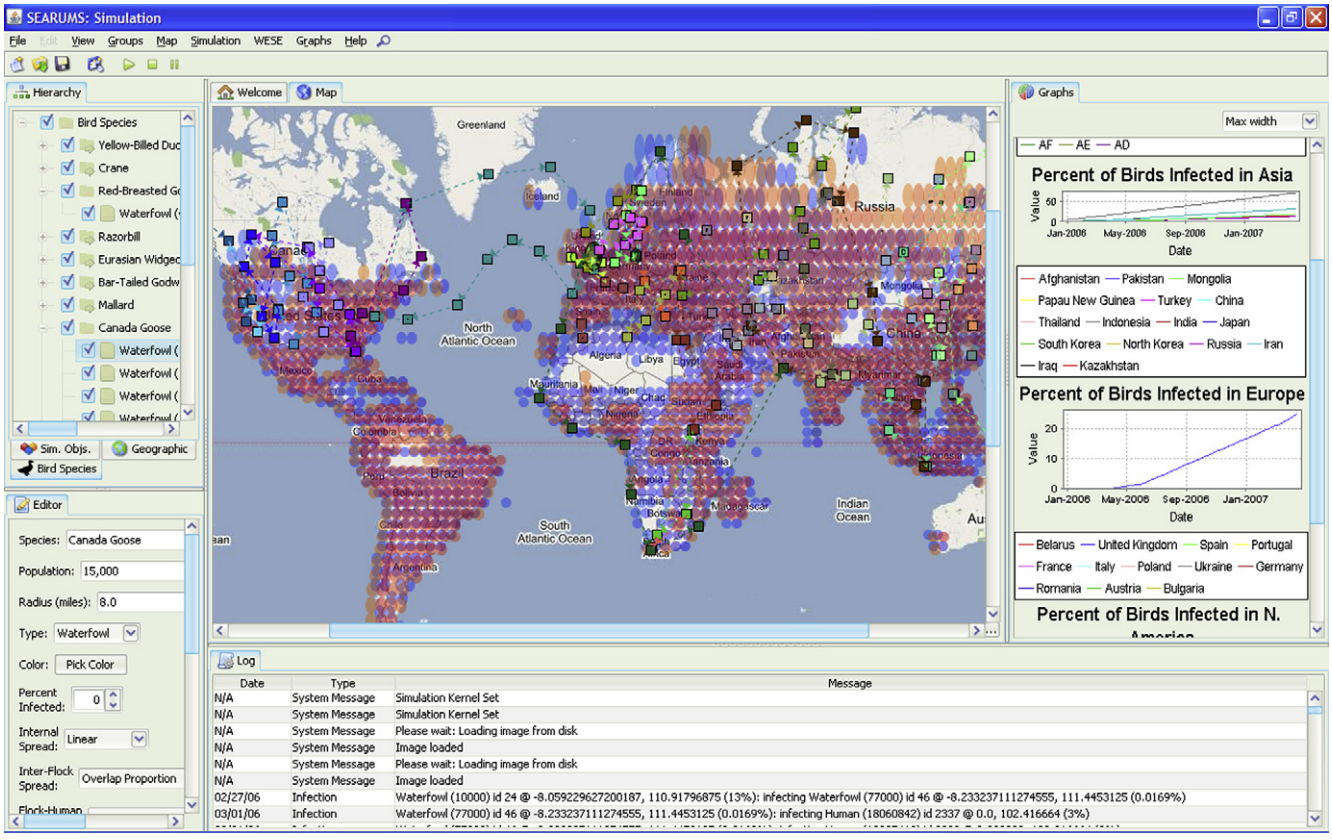

Fig. 6. The figure presents a screenshot of SEARUMS illustrating the graphical layout as seen by a user. The various modules constituting SEARUMS have been marked using black dashed lines. Purple circles indicate human groups and orange circles indicate poultry flocks. Variation in colors arises due to overlap of human and poultry flocks. Colored squares and corresponding colored dashed lines illustrate migration paths of waterfowl flocks.

The statistical attributes for agent instances include their initial population, density and distribution, initial infection percentage, infection spread parameters, incubation periods, mortality rates, and population regrowth parameters. The initial values are set with respect to the logical time when the simulation commences. In our case studies, we have either directly obtained or derived these attributes from data published by various national organizations such as: the World Heath Organization (WHO) (WHO, 2006b,c, 2007), Centers for Disease Control and Prevention (CDC), The U.S. Census Bureau (USCB, 2006), FAO's Global Livestock Production and Health Atlas (GLiPHA) database (GLiPHA, 2007), and GROMS (GROMS, 2006).

Fig. 6 presents a geographic view of various agents constituting a given model. Once all the instances have been configured, parameters for observation are added to the Eco-description. These parameters are selected by the user from a list of options offered by SEARUMS. The parameters can at individual entity level or at a group level. Each parameter is configured to be sampled hourly, daily, or weekly in terms of simulation time. In addition, each parameter can be subjected to statistical operations and they can be plotted using a variety of charts provided by SEARUMS. 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 the software equivalent of the individual-based (Law et al., 2005; Rao and Wilsey, 2005; Booth, 1997) spatially explicit (Law et al., 2005; Rao and Wilsey, 2005; Booth, 1997) mathematical model described in Section 3.

In Phase 2, the Eco-description developed in the first phase is loaded and simulated by SEARUMS' Simulation Module. The Simulation Module performs the task of triggering and coordinating the time-based interactions between the various agent instances constituting the Eco-description. The Simulation Module uses discrete event simulation methodology for achieving its functionality (Rao, 2003; Rao and Wilsey, 2005). In this approach, each agent's time-based behavior and causal interactions with other agents are triggered using simulation-time stamped events (Rao, 2003; Rao and Wilsey, 2005). The events from various processes are causally coordinated (Lamport, 1978) using a global simulation time that monotonically increases (Lamport, 1978) in discrete steps as the simulation progresses. The Dynamic Control & Steering module enables a user to dynamically (i.e., during simulation) perform various tasks such as: start, stop, pause the simulation, trigger new infections that were not previously configured in the Eco-description, change selected entity attributes, and save snapshots of the simulation. As shown in Fig. 5, the Simulation Module provides time-based changes in attribute values to the Visualization & Analysis subsystem of SEARUMS.

In Phase 3, the Visualization & Analysis subsystem is used to observe, analyze, and infer results from the simulation. This subsystem consists of three modules, namely: the Statistical Analysis & Visualization module, the Geographic Visualization module, and Event Visualization module. The Statical Analysis & Visualization module performs the task of plotting graphs and charts. Recollect that the parameters plotted by this module are preselected in Phase 1 and included in the Eco-description. Fig. 6 illustrates the GUI presented by this module. The Geographic Visualization module presents a continuously updated cartographic view of the agent instances constituting the Eco-description. The Event Visualization module provides a chronological log of events and interactions occurring between the agent instances. The models in the Visualization & Analysis subsystem dynamically interact with the Simulation Module to collate and visualize data. Consequently, this part of Phase 3 overlaps with Phase 2. However, upon completion of simulation, these modules operate independently and can be used to further analyze the data for drawing inferences and conclusions.

5. Experiments

In our current research, we have drawn inferences and conclusions from a variety of case studies conducted using a single Eco-description. The Eco-description has been developed via SEARUMS using the modeling methodology described in Section 4. Table 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; USCB, 2006; 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 is 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 GROMS (2006) and 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 1, our model includes the complete human population (approximately 6.646 billion) humans represented by 1314 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 are equal. The radius was computed using the grid size of gridded human population data from SEDAC (2007). The human agents also include the 26 major metropolitan areas of the United States, which has been the focus of one of our case studies discussed further below.

A similar strategy has also been applied to distribute the 18.136 billion poultry birds to 1315 poultry agents as shown in Table 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.

We verified the accuracy and fidelity of the aforementioned Eco-description by performing extensive simulations with initial source of infection set to outbreak in Indonesia, a notable epicenter of H5N1 epidemics (WHO, 2006a). The Eco-description was 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 following significant real-world incidents reported by WHO (2006a):

Table 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	2,31,700	5
Common crane	9	2,25,000	21
*Eurasian widgeon	3	1,296,000	17
Great knot	3	231,000	8
*Mallard	1	5000	1
Razorbill	1	1,48,000	4
*Red-breasted goose	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	1315	18,136,146,826	All
Total human groups	1314	6,646,739,849	All
Total	2673	24,787,572,675	All

Note that each agent is used to represent a group. Entries prefixed with *indicate 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.

(i) Initial infection source corresponding to outbreak in Indonesia reported on 6 December 2005 (WHO update 45 (WHO, 2006a)).
(ii) Avian influenza situation in Iraq reported on 1 March 2006 (update 4).
(iii) Avian influenza situation in Indonesia reported on 8 May 2006 (update 11).
(iv) Avian influenza situation in China reported on 16 June 2006 (update 12).

Using the validated and calibrated Eco-description we performed three extrapolative case studies to analyze the spread and pandemic threat posed by avian influenza to United States. Each of the three case studies is discussed below.

5.1. Case study 1: factors influencing spread of H5N1 to the United States

The first part of our investigations aimed to answer the important question as to if-and-when avian influenza will be transmitted to the United States through migrating waterfowl. In order to conduct this case study, we analyzed the impact of varying the percentage of infection in waterfowl flocks within close proximity to known primary sites of disease outbreaks in south-east Asia. The three experimental groups of waterfowl that were varied in this study are:

- Experimental Group 1 (EG1): Common Crane (*Grus grus*) near Hong Kong.
- Experimental Group 2 (EG2): Red-Crowned Crane (*Grus japonensis*) in South Korea.
- Experimental Group 3 (EG3): Bar-Tailed Godwit (*Limosa lapponica*) in Indonesia.

Three independent sets of experiments were conducted using these experimental groups. In each experimental group, the initial infection percentage was varied from 5% to 100% in increments of 5%. None of the other waterfowl flocks carried any initial infections in these experiments. In all these experiments, the simulations were run for a period of 5 years. The infection transmission chains and timelines to spread the disease to the United States for the three experimental groups are shown in Fig. 7. The data shown for each experimental group is average from 30 different stochastic runs of the model. The 95% confidence interval for the averages is shown as gray bars in Fig. 7.

Contrary to our hypothesis, we noted that the arrival dates of the infection in United States were independent of the percentage of birds infected initially in each experimental group. Instead, the spread was determined by the migratory pathways and timelines of the different species of waterfowl. Moreover, the experimental groups EG1 and EG2 share a common trans-Pacific migration sub-path. Conversely, experimental group EG3 took a trans-Atlantic path and arrived in June, 2007. One of the notable observations is that experimental group EG3 correctly predicted an outbreak in the United Kingdom.

5.2. Case study 2: assessing a pandemic threat

The objective of this case study was to assess a pandemic threat with focus on outbreaks in human population in 26 major metropolitan areas of United States (USCB, 2006). The source of human infections were infected waterfowl and poultry population in a given area. The percentage of humans infected and human fatalities were computed based on statistics reported by WHO for prior outbreaks in other parts of the world (WHO, 2006a). The study used the current epidemiological nature of H5N1 that sustained human-to-human transmission does not occur. Since humans are not vectors of H5N1 it was unnecessary to incorporate mobility of humans in this case study. The initial infection percentage was set to 5% for all experimental groups of waterfowl (EG1, EG2, and EG3) introduced in case study 1. The simulations were performed for a period of 5 years. The list of infections and human fatalities observed in the simulations for various metropolitan areas are shown in Table 2.

As illustrated by the data in Table 2, this scenario analysis indicates that a pandemic is unlikely given the assumption that sustained human-to-human transmission of H5N1 does not occur. It was observed that the entire continental United States will be affected with coastal areas being the entry points for the infection. This case study also highlights potential epicenters for the disease. Moreover, we observed cyclical patterns of human infections correlating with the annual migration patterns of waterfowl.

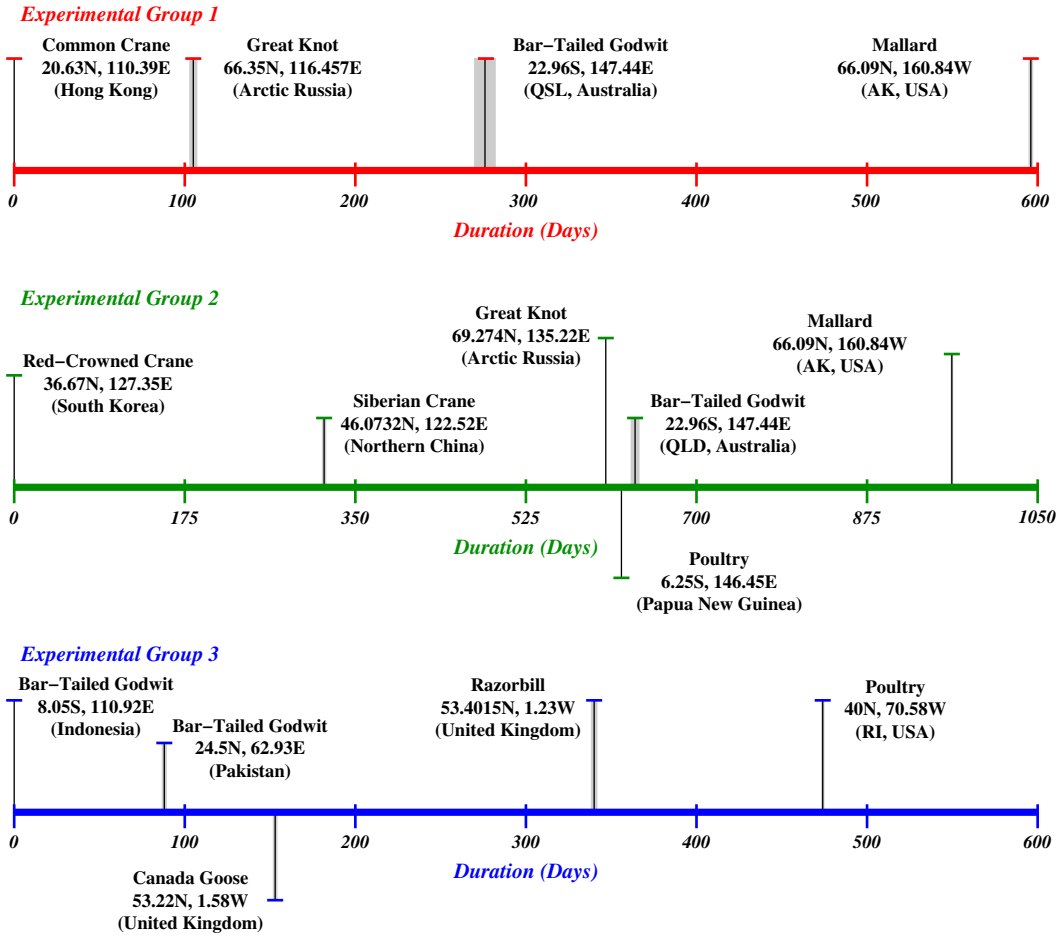

Fig. 7. Timelines for H5N1 infection spread. All of the experiments were logically set to start on day zero of simulation. The data shown for each experimental group is the average value computed from 30 simulation runs. The gray bars show the 95% confidence interval for each data point.

Table 2
Simulated human morbidity and mortality rates in U.S. Metropolitan areas over a period of 4 years

City	Year #2		Year #3		Year #4		Year #5	
	Morbidity	Mortality	Morbidity	Mortality	Morbidity	Mortality	Morbidity	Mortality
Atlanta (Georgia)	0 (±0)	0 (±0)	1 (±0.08)	0.4 (±0.17)	1.8 (±0.14)	0.67 (±0.22)	1.85 (±0.04)	0.83 (±0.24)
Baltimore (Maryland)	69.1 (±5)	30 (±2.54)	73.4 (±5.4)	35.25 (±3)	77.9 (±0.7)	36.1 (±1.1)	78.1 (±0.58)	35.5 (±1.3)
Boston (Massachusetts)	10.25 (±0.75)	5.4 (±0.47)	11.4 (±0.8)	7.2 (±0.65)	11.9 (±0.14)	7.45 (±0.4)	11.8 (±0.02)	7.25 (±0.23)
Chicago (Illinois)	0 (±0)	0 (±0)	0.9 (±0.1)	0 (±0)	1.7 (±0.34)	0.25 (±0.23)	1.8 (±0.3)	1.2 (±0.3)
Cincinnati (Ohio)	0 (±0)	0 (±0)	1.0 (±0.08)	0.4 (±0.2)	1.8 (±0.14)	0.66 (±0.22)	1.85 (±0.03)	0.8 (±0.2)
Cleveland (Ohio)	1.6 (±0.1)	0.4 (±0.17)	6.8 (±0.5)	3.3 (±0.46)	9.2 (±0.5)	4.75 (±0.5)	9.5 (±0.14)	4 (±0.4)
Dallas (Texas)	0 (±0)	0 (±0)	0.17 (±0.06)	0 (±0)	0.7 (±0.03)	0 (±0)	0.7 (±0.03)	0.2 (±0.14)
Houston (Texas)	0 (±0)	0 (±0)	0 (±0)	0 (±0)	0 (±0)	0 (±0)	0 (±0)	0 (±0)
Denver (Colorado)	0 (±0)	0 (±0)	0 (±0)	0 (±0)	1.9 (±0.1)	0.625 (±0.17)	2.1 (±0.08)	0.875 (±0.25)
Detroit (Michigan)	1.5 (±0.1)	0.42 (±0.17)	7 (±0.5)	3.3 (±0.45)	9.5 (±0.6)	4.75 (±0.5)	9.8 (±0.14)	4.0 (±0.44)
Minneapolis (Michigan)	0 (±0)	0 (±0)	2.3 (±0.18)	0.41 (±0.17)	3.5 (±0.2)	1.5 (±0.3)	3.6 (±0.18)	1.95 (±0.37)
Los Angeles (California)	0.1 (±0.01)	0 (±0)	1.5 (±0.042)	0.58 (±0.17)	2.9 (±0.05)	1.25 (±0.2)	2.9 (±0.04)	1.75 (±0.27)
Miami (Florida)	0 (±0)	0 (±0)	0.36 (±0.11)	0.083 (±0.1)	0.28 (±0.02)	0 (±0)	0.3 (±0.02)	0 (±0)
New York (New York)	35.3 (±2.6)	15.125 (±1.25)	35.6 (±2.6)	17.5 (±1.5)	37.14 (±0.3)	16.5 (±0.7)	37.16 (±0.3)	16.8 (±0.8)
Phoenix (Arizona)	0 (±0)	0 (±0)	0.1 (±0.03)	0 (±0)	2.2 (±0.06)	0.9 (±0.25)	2.2 (±0.05)	1 (±0.2)
Pittsburgh (Pennsylvania)	1.5 (±0.1)	0.4 (±0.17)	5.75 (±0.42)	2.875 (±0.4)	7.5 (±0.42)	4.0 (±0.44)	7.6 (±0.13)	3.2 (±0.4)
Portland (Oregon)	0.23 (±0)	0 (±0)	1.3 (±0.06)	0.6 (±0.17)	1.7 (±0.055)	0.9 (±0.2)	1.7 (±0.05)	0.7 (±0.2)
Riverside (California)	0 (±0)	0 (±0)	0 (±0)	0 (±0)	0.2 (±0)	0 (±0)	0.2 (±0)	0 (±0)
Sacramento (California)	0.12 (±0)	0 (±0)	1 (±0.03)	0 (±0)	1.6 (±0.03)	0.4 (±0.17)	1.6 (±0.02)	0.7 (±0.2)
San Diego (California)	0 (±0)	0 (±0)	0 (±0)	0 (±0)	0.2 (±0)	0 (±0)	0.2 (±0)	0 (±0)
San Fransisco (California)	0 (±0)	0 (±0)	0.17 (±0)	0 (±0)	0.3 (±0)	0 (±0)	0.3 (±0)	0 (±0)
Seattle (Washington)	0.22 (±0.0)	0 (±0)	1.3 (±0.06)	0.58 (±0.17)	1.7 (±0.05)	0.9 (±0.2)	1.7 (±0.05)	0.7 (±0.24)
St.Louis (Missouri)	0 (±0)	0 (±0)	0.2 (±0.04)	0 (±0)	0.3 (±0.09)	0.04 (±0.07)	0.31 (±0.09)	0.041 (±0.07)
Tampa (Florida)	0 (±0)	0 (±0)	0.36 (±0.12)	0.08 (±0.1)	0.3 (±0.02)	0 (±0)	0.3 (±0.02)	0 (±0)
Washington DC	33.85 (±2.5)	14.95 (±1.5)	37.7 (±2.8)	17.6 (±1.8)	40.8 (±0.52)	19.54 (±0.85)	40.95 (±0.3)	18.6 (±0.95)

The values shown are averages obtained from 30 independent simulation runs. The value in parentheses indicates the 95% confidence interval for each average value. Note that during year #1 there were no infections in the United States in this case study.

5.3. Case study 3: analyzing impacts to US economy

The negative socio-economic impacts of avian influenza include (i) large scale culling of birds to control spread of disease; (ii) market loss due to culling and import/export restrictions; (iii) drop in consumption of poultry; (iv) human infections and deaths through contact with infected birds; (v) decline in tourism to affected areas; (vi) and endangered biodiversity (Brahmbhatt, 2006). In this case study we analyze the impacts to US poultry industry due to large scale poultry culling that is currently required to contain H5N1 infections (WHO, 2007; Brahmbhatt, 2006). In these experiments, the source of poultry infection was infected waterfowl (WHO, 2007; Chen et al., 2005). The percentage of poultry infection and culling was computed using published statistics (WHO, 2007; Brahmbhatt, 2006) in conjunction with the duration and degree of contact between poultry and infected waterfowl. Culled flocks were set to regenerate at a fixed rate. Similar to case study 2, the initial infection percentage was set to 5% for all experimental groups (EG1, EG2, and EG3). The simulations were performed for a period of 5 years.

The graph in Fig. 8 presents the relative fluctuation of poultry population in continental United States. The blank line is the average value obtained from 30 independent runs of the model. The gray area illustrates the corresponding 95% confidence interval for the average values plotted in Fig. 8. Decrease in poultry population corresponds to culling of birds while increase in population corresponds to regeneration of poultry flocks. 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. Note that in this study, we have only accounted for direct infections from waterfowl to poultry. However, in the US, poultry litter and byproducts are used in animal feed that may cause secondary infections and other economic losses.

6. Conclusion

This article described a novel methodology involving computer-based simulations to analyze the epidemiology of avian influenza and its impacts. An overview of SEARUMS, the modeling and simulation environment, developed as a part of this research was presented. SEARUMS has been used to develop an Eco-description that was employed to conduct several case studies. The results from three case studies were discussed. Based on the results obtained from the case studies the following conclusions and inferences have been drawn. First and foremost, the spread of avian influenza to USA is inevitable. Moreover, the infection will recur year after year based on the cyclical migration patterns of the infected waterfowl. From the first case study, we inferred that controlling the population of infected waterfowl will not slow down inter-continental spread of avian influenza.

On a positive note, our studies indicate that given the current form with unsustained human-to-human transmissions, an H5N1 pandemic in humans is unlikely. However, human infections and mortality will occur. Our methodology provides an effective mechanism to predict the timelines and epicenters of the infections. Forecasting epicenters of infections will enable national and international organizations like the Centers for Disease Control (CDC) and the WHO to strategically deploy limited quantities of vaccines and countermeasures in a timely manner to contain outbreaks and save human lives. The information can also be utilized by poultry farmers and government agencies such as Food And

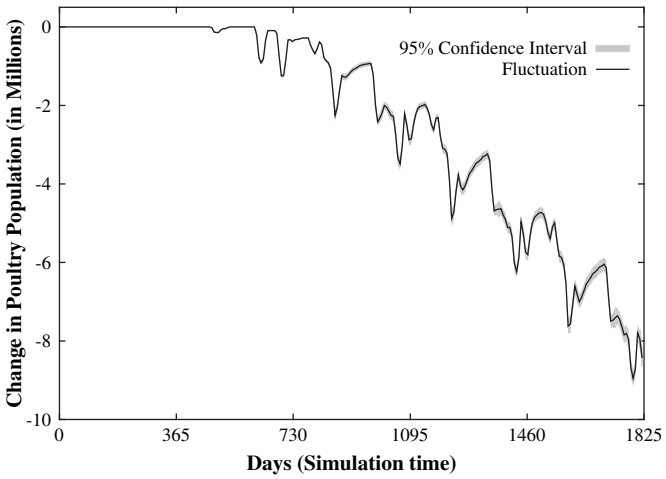

Fig. 8. Fluctuation in U.S. poultry population. The data (blank line) is average values (along with 95% confidence interval plotted in gray) obtained from 30 simulation runs. Note that the fluctuations commence only after H5N1 has been transmitted to the United States. Until such time change in population due to H5N1 epidemics is zero.

Drug Administration (FDA) to enforce suitable preventive measures to combat H5N1 outbreaks in poultry.

The multi-disciplinary nature of our methodology enables analyzing the socio-economic impacts of avian influenza. Researchers, epidemiologists, and ornithologists can utilize the simulations for rapid "what-if" types of analysis to study impacts of other factors influencing H5N1 outbreaks. Our methodology can be readily extended to include additional aspects of avian influenza as-and-when further information about the disease is discovered. It can be used to analyze other scenarios such as those simulated by Los Alamos National Laboratory (LANL, 2006). SEARUMS and our Eco-description provides an excellent foundation for further enhancements. We kindly implore the scientific community to contribute their enhancements and models for further research via SEARUMS website (SEARUMS, 2008). 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. The proposed methodology and software 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, R.M., May, R.M., 1992. Infectious Diseases of Humans: Dynamics and Control. Oxford University Press, Walton Street. Oxford OX 2 6DP.

Booth, G., 1997. Gecko: a continuous 2d world for ecological modeling. Artificial Life 3 (3), 147–163.

Brahmbhatt, M., June 2006. Economic impacts of avian influenza propagation.

CDC, August 2006. Centers for Disease Control & Prevention: Avian Influenza: Current Situation Available from: http://www.cdc.gov/flu/avian/outbreaks/current.htm.

Chen, H., Smith, G., Zhang, S.Y., Qin, K., Wang, J., Li, K.S., July 2005. Avian flu: H5n1 virus outbreak in migratory waterfowl. Nature 436, 191–192.

Ferguson, N.M., Cummings, D.A.T., Fraser1, C., Cajka, J.C., Cooley, P.C., Burke, D.S., 2006. Strategies for mitigating an influenza pandemic. Nature 442, 448–452.

Gilbert, N., Bankes, S., 2002. Platforms and methods for agent-based modeling. In: Proceedings of the National Academy of Sciences of the United States of America. 99. pp. 7197–7198.

GLiPHA, 2007. Global Livestock Production and Health Atlas (GLiPHA): Animal Production and Health Division of Food and Agriculture Organization of the United Nations Available from: http://www.fao.org/ag/aga/glipha/index.jsp.

GROMS, April 2006. Global Register of Migratory Species: Summarising Knowledge about Migratory Species for Conservation Available from: http://www.groms.de/.

Hagemeijer, W., Mundkur, T., May 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 Available from: http://www.fao.org/ag/againfo/subjects/en/health/diseases-cards/conference/documents/hagemeijer-mundkur.pdf.

Halloran, M.E., Ferguson, N.M., Eubank, S., Ira, M., Longini, J., Cummings, D.A.T., Lewis, B., Xu, S., Fraser, C., Vullikanti, A., Germann, T.C., Wagener, D., Beckman, R., Kadau, K., Barrett, C., Macken, C.A., Burke, D.S., Cooley, P., March 2008. Modeling targeted layered containment of an influenza pandemic in the united states. Proceedings of the National Academy of Sciences of the United States of America 105 (12), 4639–4644. Available from: www.pnas.org/cgi/doi/10.1073/pnas.0706849105.

Hare, M., Deadman, P.J., 2004. Further towards a taxonomy of agent-based simulation models in environmental management. Mathematics and Computers in Simulation 64 (1), 25–40.

Iwami, S., Takeuchi, Y., Liu, X., 2007. Avianâ-human influenza epidemic model. Mathematical Biosciences 207, 1–25.

Lamport, L., July 1978. Time, clocks, and the ordering of events in a distributed system. Communications of ACM 21 (7), 558–565.

LANL, April 2006. Los alamos National Laborartory: Avian Flu Modeled on Supercomputer, Explores Vaccine and Isolation Options for Thwarting a Pandemic Available from: http://www.lanl.gov/news/index.php?fuseaction=home.story&story_id=8171.

Law, R., Dieckmann, U., Metz, J.A., 2005. The Geometry of Ecological Interactions: Simplifying Spatial Complexity. Cambridge University Press, 32 Avenue of the Americas, New York, NY. 10013-2473, USA.

Liu, J., Xiao, H., Lei, F., Zhu, Q., Qin, K., w. Zhang, X., Zhang, X.-l., Zhao, D., Wang, G., Feng, Y., Ma, J., Liu, W., Wang, J., Gao, G.F., August 2005. Highly pathogenic H5N1 influenza virus infection in migratory birds. Science 309 (5738), 1206.

Longini, I.M., Nizam, A., Xu, S., Ungchusak, K., Hanshaoworakul, W., Cummunings, D. A.T., Halloran, M.E., 2005. Containing pandemic influenza at the source. Science 309 (5737), 1083–1087.

Luke, S., Cioffi-Revilla, C., Panait, L., Sullivan, K., Balan, G., July 2005. Mason: a multiagent simulation environment. Simulation 81 (7), 517–527.

Normile, D., March 2006a. Avian influenza: evidence points to migratory birds in H5N1 spread. Science 311 (5765), 1225.

Normile, D., June 2006b. Avian influenza: human transmission but no pandemic in Indonesia. Science 312 (5782), 1855.

Normile, D., March 2006c. Avian influenza: studies suggest why few humans catch the H5N1 virus. Science 311 (5768), 1692.

Normile, D., January 2006d. Avian influenza: who proposes plan to stop pandemic in its tracks. Science 331 (5759), 315–316.

Railsback, S.F., Lytinen, S.L., Jackson, S.K., September. 2006. Agent-based simulation platforms: review and development recommendations. Simulation 82 (9), 609–623.

Rao, D.M., 2003. Study of dynamic component substitution. Ph.D. thesis, University of Cincinnati.

Rao, D.M., Chernyakhovsky, A., Rao, V., October 2007. SEARUMS: studying epidemiology of avian influenza rapidly using simulation. In: Proceedings of ICMHA'07. Berkeley, CA, USA, pp. 667–673.

Rao, D.M., Wilsey, P.A., April 2005. Accelerating spatially explicit simulations of spread of lyme disease. In: Proceedings of the 38th Annual Simulation Symposium, San Diego, California, USA, pp. 251–258.

SEARUMS, April 2008. SEARUMS Website. Available from: http://www.searums.org/.

SEDAC, December 2007. SocioEconomic Data and Applications Center (SEDAC): Gridded Population of the World Available from: http://sedac.ciesin.columbia.edu/.

Solow, A.R., Smith, W.K., 2006. Using Markov Chain successional models backwards. Journal of Applied Ecology 43 (1), 185–188.

Tobias, R., Hofmann, C., January 2004. Evaluation of free java-libraries for social-scientific agent based simulation. Journal of Artificial Societies and Social Simulation (JASS) 7 (1) Available from: http://jasss.soc.surrey.ac.uk/7/1/6.html.

United States Department of Agriculture, March 2007. Avian influenza (bird flu). Available from: http://www.usda.gov/wps/portal/usdahome%3F%20navid=AVIAN_INFLUENZA.

Upadhyaya, R.K., Kumaria, N., Rao, V.S.H., 2008. Modeling the spread of bird flu and predicting outbreak diversity. Nonlinear Analysis: Real World Applications 9 (4), 1638–1648.

USCB, July 2006. U.S. Census Bureau: 100 Largest Counties Available from: http://www.census.gov/popest/counties/CO-EST2006-07.html.

WHO, October 2005. World Health Organization: Ten Things You Need to Know About Pandemic Influenza Available from: http://www.who.int/csr/disease/influenza/pandemic10things/.

WHO, July 2006a. World Health Organization: Avian Influenza: Updates Available from: http://www.who.int/csr/disease/avian_influenza/updates/en/.

WHO, May 2006b. World Health Organization: H5N1 Avian Influenza: Timeline Available from: http://www.who.int/entity/csr/disease/avian_influenza/timeline.pdf.

WHO, July 2006c. World Heath Organization (WHO): Avian Influenza Available from: http://www.who.int/csr/disease/avian_influenza/en/.

WHO, April 2007. World Health Organization: Avian Influenza: Situation Updates Available from: http://www.who.int/entity/csr/disease/influenza/H5N1-9reduit.pdf.

Winston, W.L., 1994. Operations Research (Applications and Algorithms), third ed. Duxbury Press, Belmont, California.

Wolfram MathWorld, December 2006. Mercator Projection Available from: http://mathworld.wolfram.com/MercatorProjection.html.

Zambon, M., April 2007. Lessons from the 1918 influenza. Nature Biotechnology 25, 433–434.