Bifurcation analysis of individual-based models in population dynamics

Ivo Siekmann[a,b,*]

[a] *National ICT Australia, Victorian Research Laboratory, Melbourne, Australia, Tel.: (+61) 3 903 58819*
[b] *Department of Electrical and Electronic Engineering, The University of Melbourne, Melbourne, Australia, Tel.: (+61) 3 903 58819*

Abstract

Individual-based models (IBM) enable us to investigate the effects of inter-individual differences within a population much more easily than traditional modelling approaches based on differential equations. However, the greater flexibility of IBMs makes it difficult to systematically analyse the parameter dependency of the model behaviour so that an IBM may be hard to interpret. In this article, bifurcation analysis techniques for investigating models based on ordinary differential equations (ODE) are transferred to IBMs. For this purpose, we infer stationary solutions of the IBM from the asymptotic dynamics. The stability of these stationary solutions can then be studied depending on model parameters. As shown previously for ODE models (Siekmann et al., 2010; Siekmann, 2013), stationary solutions S_i^S can be used as bifurcation parameters which allows us to predict survival or extinction of populations by simple algebraic relationships. This is demonstrated with the example of a simple two-strain infection IBM. Moreover, analysing model behaviour based on stationary solutions provides a unified representation of different models that allows us to rigorously compare IBMs with other modelling frameworks like, for example, ODE models. A comparison of the IBM to a population-based ODE model of a two-strain infection leads to similar predictions although both models were built with very different modelling approaches.

Keywords: Individual-based models (IBM), Multi-strain infections, bifurcation analysis, consumer-resource interactions, ecological motifs

*Corresponding author

1. Introduction

The classical framework for models in population dynamics goes back to Lotka (1925) and Volterra (1926). Neglecting inter-individual differences, they regarded the members of a population as a collection of "particles" with identical properties. This allowed them (and others who use the same framework until today) to build elegant mathematical models where all relevant processes in ecological and epidemiological population dynamics—growth, death, infection transmission, predation—are translated to contact rates, as in models of chemical reactions. Thus, the basic unit of models of this type is the population—in the following we will refer to these models as population-based.

In contrast, the basic unit of individual-based models (IBM) in population dynamics are the states of individuals that form a population. States may be properties like age, physiological parameters such as size, an individual's spatial location or simple attributes such as "being infected". Thus, IBMs have great potential for investigating the influence of inter-individual differences on the dynamics of the total population—they do not rely on the strong assumption that population dynamics of heterogeneous populations will be accurately captured by populations consisting of "representative individuals" with identical properties. In order to make this more explicit, let us consider a few representative examples of IBMs from epidemiology. Eubank et al. (2004) simulate a population and their movement patterns based on empirical data. The resulting time series of moving individuals is transformed to a contact graph on which infections can spread. The authors investigate general graph-theoretic properties of the contact network and develop a model of smallpox transmission. The model is then used for investigating the effect of different vaccination strategies on smallpox spread.

Ferguson et al. (2005) use similar data as Eubank et al. (2004) as the basis of their IBM influenza model but without making the intermediate step of creating a contact network. Instead, the focus is on the geographical spread of influenza in Thailand. Again, the effectiveness of different interventions (vaccination, quarantine zones, closing down of schools and workplaces) is assessed by simulation studies. See Smieszek et al. (2011) for a more recent study, similar in spirit, where an influenza model is validated with empirical data of an influenza epidemic in Switzerland.

Rolls et al. (2012) consider the spread of hepatitis C infections in a social rather than a transport network. Because hepatitis C is highly prevalent in injecting drug users and is often transmitted via needle sharing the authors use an interaction network of injecting drug users that was derived from empirical data in a previous study.

In this article a simple two-strain infection model will be studied. An example for a very detailed multi-strain model of influenza is Roche et al. (2011). The authors validate the population dynamics of their very detailed IBM by comparison with traditional population-based multi-strain models based on differential equations, see the relevant chapters in Brauer et al. (2008) or Keeling and Rohani (2008). Cisternas et al. (2004) present a less detailed multi-strain model for influenza. This study is interesting because the authors use "equation-free" modelling, an approach that can be used for computational analysis of IBMs. I will review "equation-free" modelling and "coarse" bifurcation analysis—a numerical framework developed by Kevrekidis and co-workers in the last 15 years—in more detail in the discussion. Common to all IBMs presented here is that they allow us to investigate how dynamics observed at the population level arises from individual interactions. However, the detailed representation of the individual level makes it difficult to analyse IBMs systematically.

In contrast to IBMs, the analysis of population-based models is a routine exercise, thanks to the well-developed methods of nonlinear dynamics such as linear stability analysis and bifurcation theory, see, for example Wiggins (2003); Kuznetsov (1995). Population-based models are usually formulated by a system of ordinary differential equations (or, analogously, difference equations) of the form

$$\frac{d\mathbf{P}(t)}{dt} = f(\mathbf{P}(t), \lambda), \quad \mathbf{P}(0) = \mathbf{P}^0. \tag{1}$$

Here, $\mathbf{P}$ is the vector $(P_1, \ldots, P_n)$ of n interacting populations and λ is a vector of parameters. The question under which conditions a population P_i survives or goes extinct can be answered by analysing the stationary solutions $\mathbf{P}^s$ of (1) which are obtained from the system of nonlinear equations

$$f(\mathbf{P}^S, \lambda) = 0. \tag{2}$$

Choosing a suitable bifurcation parameter is an important modelling decision. Usually, a specific parameter (such as a predator mortality or the transmission rate of an infection) is selected. This type of bifurcation analysis reveals the influence of a particular process on the system behaviour. In contrast, Siekmann et al. (2010) and Siekmann (2013) have shown that choosing the stationary solutions S_i^S of a resource S as bifurcation parameters gives insight into the effect of simultaneously varying all parameters, or, in other words, the response of the system under the influence of all relevant processes. For example, if two consumers compete indirectly via depleting a shared resource S the consumer that reduces the resource to the lower

steady state level drives its competitor to extinction (exploitative competition). Thus, consumer 1 survives if $S_1^S < S_2^S$ holds and consumer 2 displaces consumer 1 if this inequality is reversed. Similar survival conditions were found for the most common example systems in ecology and epidemiology including food chains, competition models and intraguild predation. A great advantage of choosing stationary solutions S_i^S as aggregated bifurcation parameters is that unlike in a classical bifurcation study we do not have to restrict our attention to one process related to a specific model parameter. Moreover, stability conditions depending on stationary solutions are not specific for ecological or epidemiological systems so that Siekmann et al. (2010) and Siekmann (2013) could use this approach for comparing the complete parameter space of seemingly different systems from ecology, epidemiology and eco-epidemiology. It was shown that specific stability conditions are associated with different types of interactions like food chains, various forms of competition and intraguild predation. This implies that for two different system that are characterised by the same type of interaction the parameter spaces are structured in a similar way. Therefore qualitatively similar dynamics is expected when two predators compete for the same prey or when two disease strains compete for the same host because both systems are examples for exploitative competition. Since the qualitative dynamics observed thus to a large extent depends on the type of interaction rather than on the underlying system I proposed to refer to these patterns as "motifs" (Siekmann, 2013). Whereas in previous publications all models were based on differential equations, I will show in this article that the approach can also be applied for comparing models constructed with different modelling frameworks, namely stochastic IBMs and deterministic differential equations.

Analysing system behaviour by choosing stationary solutions as bifurcation parameters allows us to carry out a bifurcation analysis of an IBM. Because IBMs are not implemented as a system of equations, stationary solutions cannot be calculated as in (2). Nevertheless, *stable* stationary solutions—in the following denoted steady states—can be obtained by simulating an IBM for a sufficient number of iterations. In this article we demonstrate how this approach can be implemented in practice by investigating a simple example of an individual-based two-strain infection model, see Figure 1 for a schematic representation of the model structure.

Our model, which is an extension of the `Virus` model (Wilensky, 1998) provided with `NetLogo` (Wilensky, 1999), is an example for an exploitative competition system because both strains compete for a population of susceptibles without interacting directly with each other. I deliberately chose this very simple model in order to make the workflow as clear as possible. Moreover, I have made available the code for download from

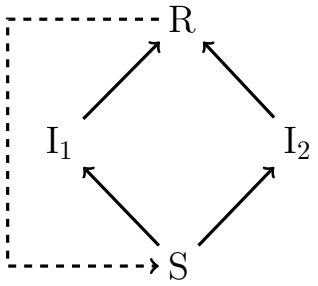


Figure 1: Model structure of the individual-based two-strain infection model and the ODE model (3)-(5) analysed in this article. Solid arrows stand for state transitions from susceptibles S to one of the strains I_1 or I_2, and from infected to resistant R. The dashed arrow accounts for susceptible offspring produced by the resistant population.

`https://github.com/merlinthemagician/2strainIBM` so that readers can follow the steps of the analysis as described in this article and reproduce the results.

After a short description of the two-strain model in Section 2, the one-strain version of the model is thoroughly analysed by determining the (stochastic) steady state S^* depending on the parameters `infectiousness`, `duration` and `chance-recover`. Representing these results as a bifurcation study with bifurcation parameter S^* enables us to compare the IBM with a population-based two-strain infection model based on ordinary differential equations (ODE), see (3)-(5).

In Section 4 we show that the two-strain IBM behaves according to the exploitative competition motif so that survival and extinction only depend on the simple condition $S_1^S \lessgtr S_2^S$ where S_i^S, $i = 1, 2$ are the stationary solutions for the susceptible population if only strain 1 or 2 are present, respectively. Based on this condition we can also infer for a given parameter set how much the competition of the two strains will be influenced by stochastic events. We discuss our results in Section 5 and give an outlook to future work.

2. A simple individual-based infection model

The two-strain infection model investigated in this article (see Figure 1 for a schematic representation) extends the example `Virus` (Wilensky, 1998) provided with the `NetLogo` software (Wilensky, 1999). `Virus` is a stochastic spatio-temporal model of how an infectious disease spreads in a population.

The population consists of susceptible, infected and resistant individuals that move randomly on a spatial grid.

2.1. Growth

Only susceptible and resistant members of the population can reproduce. At each time step, there is a 4% chance of producing offspring, in this case a new-born susceptible is placed next to its "mother". All members have a limited `lifespan`—in the absence of infection they die after exactly 100 time units. In order to avoid unlimited growth, no more offspring are produced if the total population has reached a `carrying-capacity` of 750.

2.2. Infection transmission

It is assumed that infection is transmitted by contacts between susceptibles and infected. When a susceptible is closer than a certain critical distance to an infected individual (in `NetLogo` terms this means that the susceptible has entered the "patch" of the infected) the infection is transmitted with probability `infectiousness`. As mentioned above, infected lose the capability of producing offspring.

2.3. Immunity or dying from infection?

Infection continues for at least `duration` time steps. When the infection ends depends on the number of time steps that an infected individual has been sick (`sick-count`). By sampling at each time step

$$x \sim U(0, \texttt{sick} - \texttt{count} - 1)$$

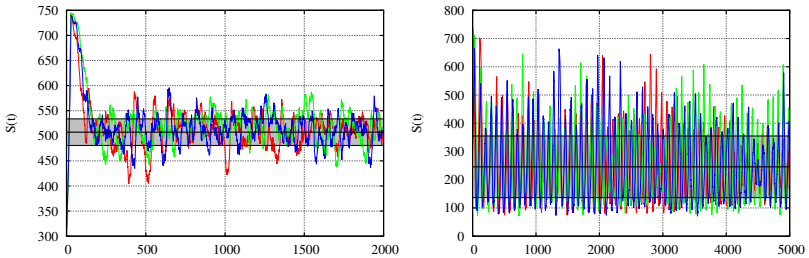
and requiring

$$x > \texttt{duration}$$

for the infection to end, the probability that the infection ends increases with `sick-count`. Infections can end with immunity or death; which of the two occurs depends on the probability `chance-recover`.

2.4. Two-strain model

For the purpose of this article this model is extended by a second strain of the disease. The two strains A and B each have parameters `infectiousness_a`, `duration_a`, `chance-recover_a` and `infectiousness_b`, `duration_b`, `chance-recover_b`, respectively. I further assume *cross-immunity*, i.e. that carriers of strain A cannot be infected by strain B and vice versa. For simplicity, it is assumed that infected of any strain that gain immunity are protected against the other strain as well. The current version of the `NetLogo` source code is provided under the link `https://github.com/merlinthemagician/2strainIBM`.



(a) stochastic steady state (b) complex fluctuations

Figure 2: Simulations of the same parameter set but for different seeds of the random number generator lead to qualitatively similar asymptotic dynamics. (a) shows an example where the solutions tend towards a stochastic steady state. The trajectories are already very close in the transient phase for $t < 200$ and seem to approach the same stochastic steady state. The mean for the last 1,000 time points of the red population is marked by the bold black line, the grey area indicates deviations of one standard deviation. Parameters: $\beta = 23\%$, $\tau = 9$, $\rho = 0\%$. (b) For different parameters ($\beta = 50\%$, $\tau = 10$, $\rho = 10\%$), the solutions exhibit complex dynamics similar to chaotic oscillations. Nevertheless, all trajectories seem to stay in the same range. Here, mean and standard deviation were calculated from the last 4,000 time points of the red population.

3. Asymptotic dynamics of the one-strain model

Our analysis of the IBM is based on asymptotic dynamics, i.e. for behaviour that is observed after running the IBM sufficiently long to reach a stochastic steady state. It is striking that although individual trajectories for the susceptibles population are clearly stochastic, the time course of different realisations (different initial conditions, different seeds for the random number generator) for the same parameter set are nevertheless very similar. In particular, all trajectories approach the same stochastic steady state (Figure 2). One remark that has to be made here is that our model has absorbing states like extinction of the infection or even the total population. For $k \to \infty$ the system must tend towards one of the absorbing states. However, the expected number of iterations it takes for reaching an absorbing state may be very large. Thus, technically, whenever I refer to "steady states" what, in fact, is meant is a quasi-steady state. Having said this, absorbing states are always available to the system and become relevant if fluctuations drive populations close to extinction.

The asymptotic dynamics was characterised by computing the arithmetic mean of populations S_k for an appropriate number of subsequent iterations k; standard deviations were calculated in order to obtain an estimate of the magnitude of fluctuations. For some parameter regions the populations do not tend towards a stochastic steady state but show regular or even chaotic oscillations. However, especially for small amplitude fluctuations it is hard to distinguish between deterministic and stochastic sources of variability. For this reason, I decided not to treat oscillatory behaviour and steady states differently—the standard deviation accounts for the magnitude of stochastic noise as well as deterministic fluctuations. By comparing several realisations for each parameter set, multi-stability could be excluded.

The one-strain version of the model was systematically simulated for 10 different levels of `infectiousness`—$\beta = 10\% - 90\%$, 7 levels of `duration`—$\tau = 10 - 70$ and 5 levels of `chance-recover`—$\rho = 10\%, 30\%, 50\%, 70\%, 90\%$. Each simulation was repeated with three different seed values for the `NetLogo` random number generator. Because we are interested in dynamics where the infection is present, the population was initialised at its carrying capacity $P = 750$; this usually ensures that the infection persists. All simulations were run for at least 10,000 iterations.

3.1. Steady state S^* unaffected by `chance-recover` ρ

A first observation from our parameter study is that `chance-recover` ρ hardly influences the mean value of the susceptibles S. Nevertheless ρ plays an important role in determining if susceptibles tend towards a steady state

or oscillate. Increasing ρ decreases fluctuations, so higher levels of ρ stabilise the system (Figure 3).

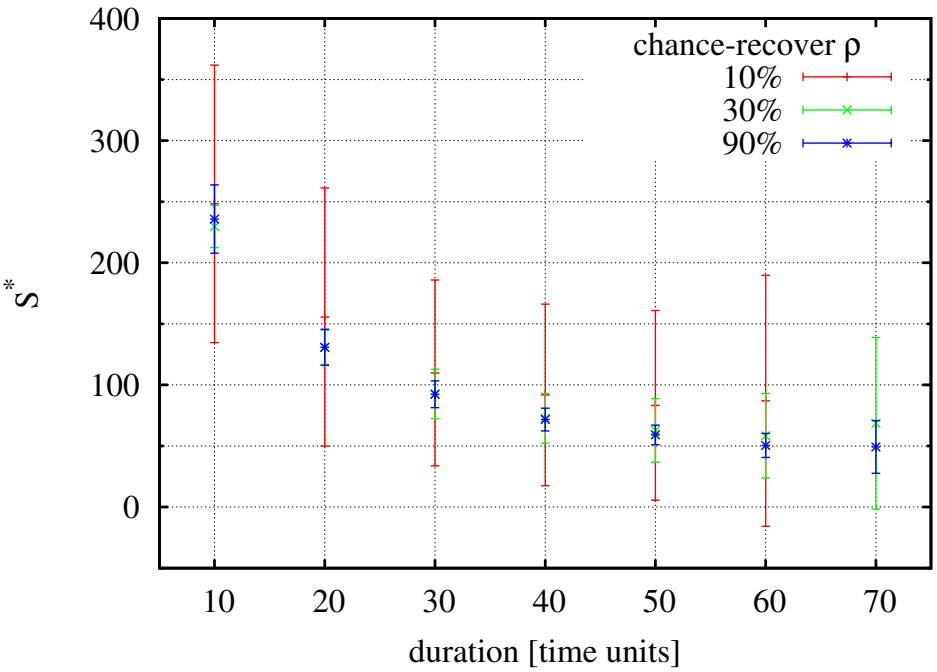


Figure 3: Dependency of the average asymptotic level of susceptibles S^* on the `duration` τ for three different levels of `chance-recover` ρ. Standard deviations (indicated by error bars) indicate the magnitude of fluctuations caused by noise or oscillations. At low values of $\rho = 10\%$ the susceptible population oscillates for the whole range of durations so that the standard deviation is large. Already for $\rho = 30\%$, the standard deviation drops for τ between 20 and 50 time units which means that the variability of the average is mainly due to noise rather than oscillatory dynamics. The dynamics is further stabilised for $\rho = 90\%$ but interestingly, the mean values are very close to those for $\rho = 30\%$. These results (which are similar for other levels of `infectiousness` β fixed at 50% for this figure) suggest that ρ influences the level of deterministic or stochastic fluctuations but hardly affect the average asymptotic level of susceptibles S^*.

3.2. A bifurcation diagram depending on `duration` τ

The bifurcation diagram shown in Figure 4 indicates how the spread of the infection depends on the parameter `duration` τ, see Wiggins (2003); Kuznetsov (1995) for an introduction to bifurcation theory. For low durations the infection is unable to spread (labelled $I^* = 0$ in Figure 4) because

infected do not carry the disease long enough for transmitting it to sufficiently many susceptible hosts. For an intermediate range of the `duration` τ, the infection is endemic without reducing the size of the total population P^* (labelled $P^* = K$). The reason presumably is that recovery of infected and transition to the resistant class is fast enough that the resistant and susceptible populations produce enough offspring to balance disease-induced deaths. For high values of `duration` τ the stationary solution $P^* = K$ is destabilised (labelled $P^* < K$), i.e. the disease reduces the total population below its carrying capacity. Interestingly, only then it is possible to observe oscillations that can be distinguished from noise—the higher variability of the solutions is reflected by increased standard deviations. For very high values of τ the amplitudes of the fluctuations increase and the minima of the infected population approach zero which strongly increases the risk that the infected population goes extinct. Alternatively, the infection may also drive the total population to extinction—if the disease is transmitted to all members of the population and infected die before they become resistant (this may be the case either for a high `duration` τ or a low `chance-recover` ρ).

3.3. Dependency of S^* on `infectiousness` β and `duration` τ

Because `chance-recover` ρ has little influence on the magnitude of S^* we continue our analysis for a fixed value $\rho = 50\%$. Figure 5 shows that for increasing values of `infectiousness` β and `duration` τ the population of susceptibles S^* decreases. This indicates that the strength of the infection increases with β and τ. In the next section it will be demonstrated that Figure 5 gives us a comprehensive understanding of the parameter space of the two-strain model. For given values of `infectiousness` and `duration` for both strains, $S_1^S = S^*(\beta_1, \tau_1)$ and $S_2^S = S^*(\beta_2, \tau_2)$ can be looked up in Figure 5. Applying the results from Siekmann et al. (2010); Siekmann (2013) the behaviour of the two-strain model can then be determined.

3.4. Comparison of the IBM to an ODE model

A very useful feature of analysing model behaviour depending on stationary solutions is that it is possible to compare completely different models. This is especially valuable because all modelling methods have particular strengths and shortcomings—therefore, general insights into the behaviour of a system can be gained best by comparing a large number of different models.

In order to demonstrate this, we write down a population-based multi-strain infection model that takes into account the same processes (see Figure 1) as the IBM described in Section 2.

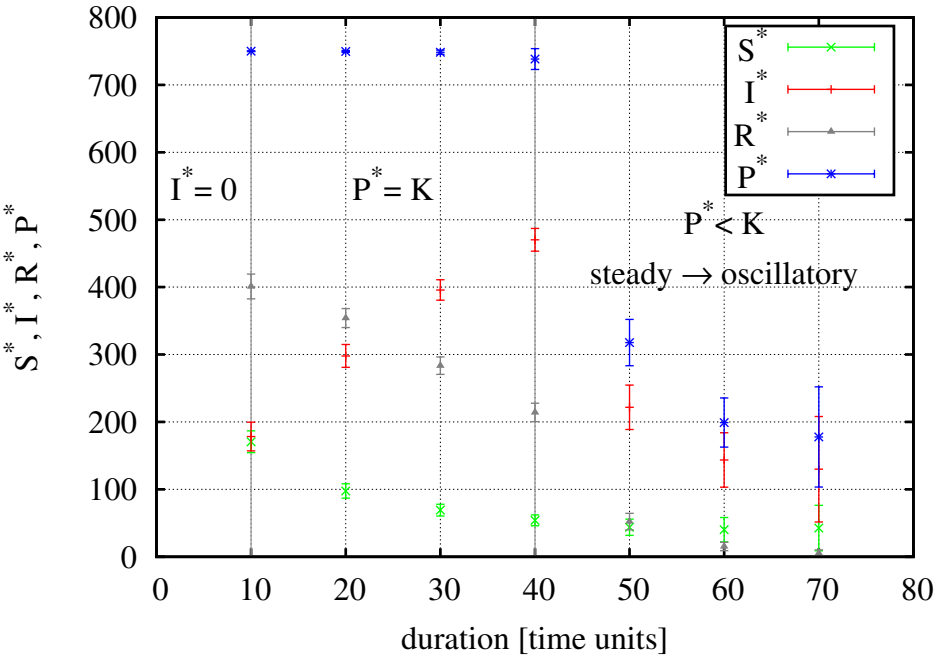
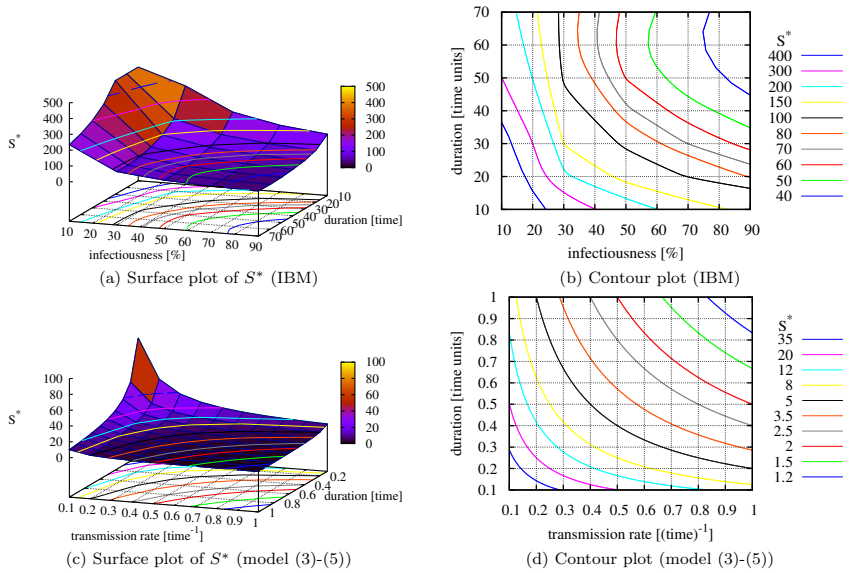


Figure 4: Mean and standard deviations of all subpopulations S^*, I^*, R^* and the total population P^* depending on the `duration` τ of the infection. The parameters chosen for this bifurcation diagram are `infectiousness` $\beta = 70\%$, `chance-recover` $\rho = 50\%$. For other choices of β and ρ the diagram may differ in that $P^* < K$ for the whole range of `duration` τ (for low values of ρ) or that there are no oscillations (for high values of ρ).



(a) Surface plot of S^* (IBM)

(b) Contour plot (IBM)

(c) Surface plot of S^* (model (3)-(5))

(d) Contour plot (model (3)-(5))

Figure 5: Comparison of the one-strain versions of the IBM described in Section 2 and the ODE model (3)-(5). (a) and (b) illustrate the dependency of the arithmetic mean of susceptibles S^* on `infectiousness` β and `duration` τ, `chance-recover` ρ was fixed at a value of 50%. (a) shows a surface plot of S^* in the IBM depending on infectiousness and duration of the disease. Contours indicating different levels of S^* are superimposed on the surface and in the β/τ plane. (b) gives a view from above of the contours in the β/τ plane. Analogous plots for the ODE model are provided in (c) and (d). Interestingly, the dependency of S^* on the transmission rate β and $\tau = (m + \alpha + \gamma)^{-1}$ (which can be interpreted as the average duration of the infection) is qualitatively very similar.

$$\frac{dS}{dt} = r(S+R)\left(1 - \frac{S+R+\sum_{k=1}^{n} I_k}{K}\right) - \sum_{k=1}^{n} \beta_k S I_k, \tag{3}$$

$$\frac{dI_k}{dt} = \beta_k S I_k - (m + \alpha_k + \gamma_k) I_k = \beta_k I_k (S - S_k^S), \tag{4}$$

$$\frac{dR}{dt} = \sum_{k=1}^{n} \gamma_k I_k - mR. \tag{5}$$

The susceptible population S grows logistically and can be infected by members of the populations I_k carrying one of the k different strains with a rate $\beta S I_k$. Infected suffer from a strain-specific mortality α_k named virulence in addition to the natural mortality m and become resistant at a rate γ_k. As in the IBM it is assumed that once resistant, an individual is resistant to all strains. Members of the resistant population (in contrast to the infected) can reproduce but immunity is not inherited, so their offspring are susceptible. Similar to the IBM this model is meant to be a simple example for a population-based multi-strain model, for an introduction to population-based multi-strain models, see, for example, the relevant chapters in Keeling and Rohani (2008) and Brauer et al. (2008).

From (4) it is easily seen that the stationary solutions for susceptible hosts are

$$S_k^S = \frac{m + \alpha_k + \gamma_k}{\beta_k} = \frac{1}{\beta_k \tau_k} \tag{6}$$

where

$$\tau_k = \frac{1}{m + \alpha_k + \gamma_k} \tag{7}$$

is the average duration of an infection with strain k. Thus, the stationary solutions for susceptibles S_k^S depend on the transmission rate β and the duration τ which have interpretations that are very similar to the corresponding parameters `infectiousness` and `duration` of the individual-based model. Figures 5a and 5c clearly demonstrate that the dependency of the stationary solution of susceptibles of the one-strain models S^* is qualitatively similar. Thus, despite the fact that the models are based on completely different modelling methods, they make analogous predictions about epidemic spread.

From (4) it also follows (using the same arguments as in Siekmann (2013)) that stable stationary coexistence of multiple strains is impossible. Moreover, analysis of the one-strain version of (3)-(5) strongly indicates that oscillatory solutions do not exist at least for small values of the recovery rate γ (this

is different from the IBM where both regular as well as chaotic oscillations can be observed). Thus, coexistence cannot be achieved either by oscillatory solutions (Armstrong and McGehee, 1976a,b). As explained in Siekmann et al. (2010); Siekmann (2013), the behaviour of (3)-(5) can in a next step be represented in a competition diagram (Figure 7a) that indicates for all possible parameter sets which of the strains is the superior competitor.

4. Analysis of the two-strain model

By taking advantage of the results from Siekmann (2013) the effort for analysing the two-strain model can be considerably reduced. But more importantly, our new approach gives additional insights into the model behaviour that cannot be obtained from traditional methods for investigating IBMs.

4.1. Additional analysis of the two-strain model is not required

From the description of the model in Section 2 it is obvious that the two strains only influence each other indirectly via depletion of a resource—in this case, the population of susceptible hosts S. Thus, the model is characterised by the "exploitative competition motif" (Siekmann, 2013). In models of this type usually one of the competitors is outcompeted by the other. Which of the two strains will drive the other to extinction can be determined by comparing $S_1^S = S^*(\beta_1, \tau_1)$ and $S_2^S = S^*(\beta_2, \tau_2)$. If

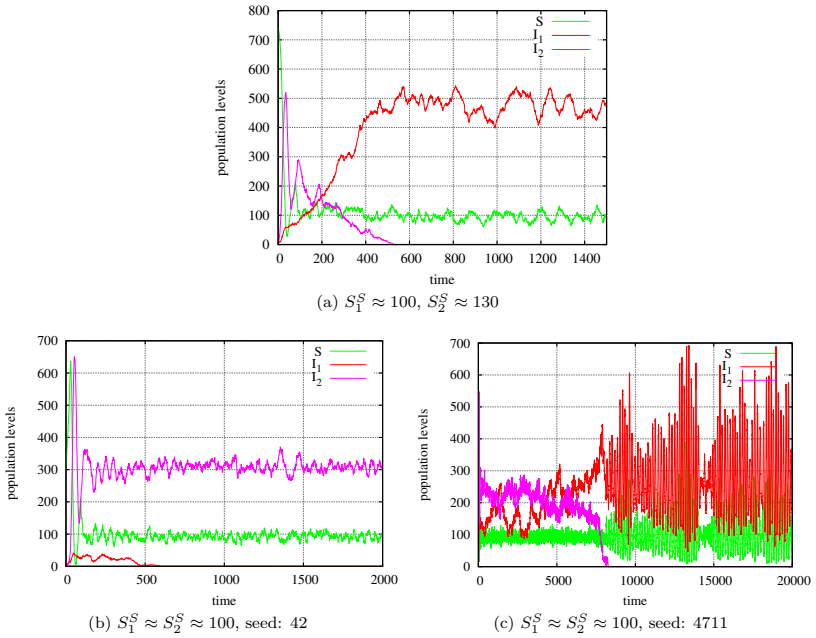
$$S_1^S < S_2^S,$$

strain 1 displaces strain 2 whereas strain 2 prevails if this inequality is reversed. From an ecological point of view, this stability condition just means that the strain that reduces the host population to the lowest level is the strongest competitor. For this reason, additional analysis of the two-strain model is not required because for a given parameter set, $S_1^S = S^*(\beta_1, \tau_1)$ and $S_2^S = S^*(\beta_2, \tau_2)$ can simply be looked up in Figure 5. Thus, rather than explicitly investigating the full model, analysis of the one-strain submodel, or in other words: half of the parameter space, is sufficient.

4.2. Extent to which parameter sets are characterised by stochasticity

In contrast to the deterministic situation of a model like (3)-(5) we cannot expect that our stochastic model will always behave as predicted by $S_1^S \lesseqgtr S_2^S$. Instead, how likely it is that, say, strain 1 outcompetes strain 2 also depends on the distance $\Delta = |S_1^S - S_2^S|$ rather than just on which of the S_i^S, $i = 1, 2$ is smaller. Figure 6a gives an example for a parameter set where $S_1^S \ll S_2^S$—thus it is expected that I_1 will nearly always outcompete I_2 independent from

the initial conditions and stochastic events. In a situation where $S_1^S \approx S_2^S$ the strength of both strains is similar, so depending on initial conditions and stochastic effects both strains have a chance to survive. Thus, due to stochastic effects in this parameter region the system is bistable, see Figures 6b,c. We have systematically confirmed these results for a large number of parameter sets.



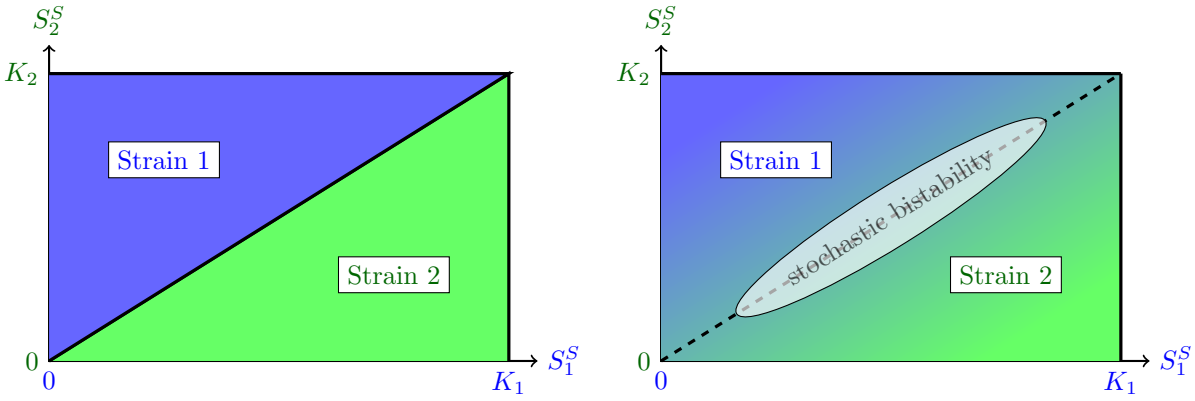
(a) $S_1^S \approx 100$, $S_2^S \approx 130$

(b) $S_1^S \approx S_2^S \approx 100$, seed: 42

(c) $S_1^S \approx S_2^S \approx 100$, seed: 4711

Figure 6: Two-strain dynamics can be predicted from the parameter study of the one-strain model. (a) For this parameter set, S_1^S and S_2^S are sufficiently different so that in most cases I_1 outcompetes I_2. Parameters: $\beta_1 = 30\%$, $\tau_1 = 50$, $\rho_1 = 50\%$, $\beta_2 = 70\%$, $\tau_2 = 15$, $\rho_2 = 50\%$, initial population: 300. seed: 815. (b) and (c) If S_1^S and S_2^S are very similar it depends on stochastic effects which of the two strains displaces the other. For both simulations, the same parameters were chosen except for the seeds of the random number generator. Parameters: $\beta_1 = 30\%$, $\tau_1 = 50$, $\rho_1 = 20\%$, $\beta_2 = 70\%$, $\tau_2 = 21$, $\rho_2 = 50\%$, initial population: 300, seeds: (b) 42, (c) 4711

The behaviour of the model depending on S_1^S and S_2^S (i.e. for any possible combination of parameters) and the extent to which the outcome is modulated by stochastic effects are graphically represented in a competition diagram (Figure 7b). Also compare with the competition diagram for the deterministic ODE model (3)-(5) (Figure 7a).



(a) model (3)-(5) (b) IBM

Figure 7: Competition diagrams for the population-based and the individual-based two-strain infection models. (a) In the deterministic model for parameter sets located above the diagonal (blue), strain 1 survives and displaces strain 2 whereas below the diagonal, strain 2 (green) prevails. (b) Unlike in the deterministic situation, the outcome of a particular realisation not only depends on the location in parameter space but also on stochastic effects. If $\Delta = |S_1^S - S_2^S|$ is large, the stronger strain will displace the weaker strain with a high probability. But if $\Delta \approx 0$ or analogously, $S_1^S \approx S_2^S$ the model behaviour strongly depends on stochastic effects. Thus, we refer to parameter sets close to the diagonal as stochastically bistable. The influence of stochasticity is accounted for in the diagram by a colour gradient where the region near the upper left and the lower right corner are plotted in blue and green respectively, whereas regions close to the diagonal are coloured blue-green indicating that both strains have a chance for survival. It has to be noted here that this graphical representation is not meant to make any quantitative statement about the actual probability which of the strains goes extinct—this would require further simulation studies.

4.3. Long-term stationary coexistence is very unlikely

Finally, let us consider the problem of coexistence. It is well-known that for exploitative competition, coexistence of all competitors is very uncommon (Hardin, 1960; Armstrong and McGehee, 1980). Stationary coexistence is impossible because it violates the stability condition $S_1^S \leqslant S_2^S$ but in models based upon deterministic differential equations we often observe oscillatory coexistence, first shown by Armstrong and McGehee (1976a,b). Due to the stochastic nature of our model it is likely that fluctuations will drive the solutions away from the coexistence attractor. In Figure 8 we show an example where the underlying deterministic dynamics is probably chaotic. The two strains of the disease initially seem to coexist until the weaker strain I_1 is eventually driven to extinction at about 3,500 time units.

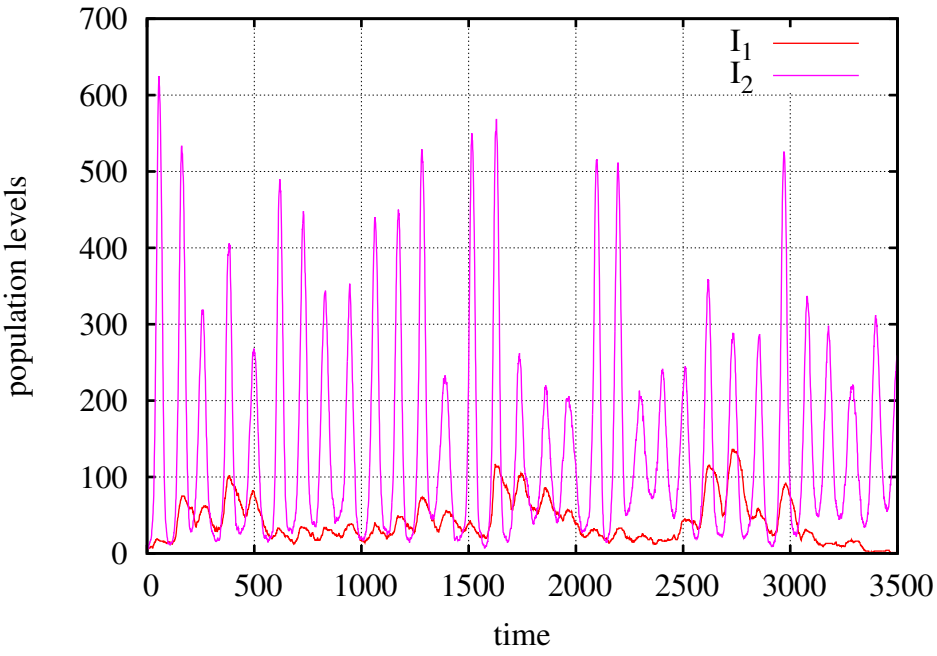


Figure 8: Two strains coexist for a long time before the weaker strain I_1 eventually goes extinct. This examples illustrates that coexistence may be less common in stochastic IBMs than in deterministic models because noise may always drive the solution off the coexistence attractor to an absorbing solution where one strain goes extinct. Parameters: $\beta_1 = 20\%$, $\tau_1 = 70\%$, $\rho_1 = 10\%$, $\beta_2 = 70\%$, $\tau_2 = 20$, $\rho_2 = 10\%$, initial population: 200. seed: 666.

5. Discussion

In this study I have shown how bifurcation theory methods that are widely and successfully used for analysing deterministic models based on differential equations can be transferred to IBMs. The new approach increases the efficiency of model analysis, has a rigorous mathematical basis and leads to results that cannot be obtained with established methods for investigating IBMs. For a simple two-strain infection model it was demonstrated that the system behaviour depends on only two aggregated parameters, the stationary population of susceptibles S_i^S, $i = 1, 2$, for strain 1 and 2 of the infection. One of the useful aspects of the representation of the model behaviour depending on S_1^S and S_2^S (Figure 7b) is that regions of parameter space that are characterised by stochastic bistability can be easily identified. In contrast, for the ODE model, the condition $S_1^S \lesseqgtr S_2^S$ holds deterministically. This result cannot be obtained with traditional methods for analysing IBMs.

The bifurcation parameters S_i^S are not only mathematically convenient but (as explained previously in Siekmann et al. (2010); Siekmann (2013)) have a clear biological interpretation. In the situation of exploitative competition for susceptible hosts, the steady state of the susceptible population simply is an aggregated measure for the strength of the infection represented by a particular parameter set. Figure 5 shows that S^* decreases monotonically for both `infectiousness` and `duration`, i.e. the strength of an infection increases with both of these parameters. Interestingly, the role of `chance-recover` is more subtle: at first sight it seems counter-intuitive that this parameter has no appreciable effect on S^*. One may expect that more susceptible offspring is produced because a larger proportion of infected becomes resistant and the resistant population produces susceptible offspring. However, most of these additional susceptibles eventually will become infected—thus, for the asymptotic dynamics, the most important effect of increasing the value of `chance-recover` seems to be that the resource pool of susceptibles available to the infected is replenished. The fact that immunity is not passed from members of the resistant population to their offspring explains why increasing the chance of recovery even stabilises the infection. In fact, the parameter range for `duration` and `infectiousness` where the total population reaches its carrying capacity K increases with `chance-recover`. As mentioned in Section 3.2, oscillations that eventually lead to extinction of the infection are only observed when the total population P^* decreases below K.

The ability to compare models as different as ODE models and IBMs has great potential for gaining deeper insight into the underlying processes of many ecological and epidemiological systems. Few modellers would argue

that the goal of modelling is not to find the perfect model of a particular system but instead to obtain a comprehensive understanding by comparing many imperfect models that all come with their own sets of simplifying assumptions and shortcomings. In this study, it has been demonstrated that choosing stationary solutions as bifurcation parameters enables us to compare models that were built using completely different modelling frameworks such as ODE models and IBMs. Grimm and Railsback (2005) are pessimistic that such a comparison was possible in a meaningful way—they make a good point that in many cases when this has been done, an IBM was specifically constructed for such a comparison and that this restricts the flexibility of representing processes at the individual level in detail. However, the method outlined here does not require us to vary "similar" parameters in the IBM and the ODE model and compare the resulting behaviour in both models. Instead (as shown in Figure 5) we can compare the effects of *independently* varying all parameters that appear in the stationary solutions. For the one-strain versions of the ODE model (3)-(5) and the IBM described in Section 2 this comparison clearly reveals which parameters of the two models correspond to each other rather than forcing us to start from parameters with similar significance in the first place. Interestingly, for both models the stationary solution S^* exhibits a similar dependency on infectiousness and duration of the infection. A clear difference between ODE model and IBM is that the coexistence steady state of the former cannot be destabilised by a Hopf bifurcation whereas the population dynamics of the IBM shows periodic and probably also chaotic fluctuations. The predictions from the two-strain versions of both models are also similar—both exhibit the exploitative competition motif where long-term coexistence of both infection strains is impossible.

Whereas we have demonstrated how bifurcation analysis can be implemented for a simple IBM, the ultimate goal is to apply this approach to more detailed models. For more complex IBMs that take into account interactions of individuals via transport and social networks, see for example Eubank et al. (2004); Ferguson et al. (2005), it is common practice to study the system behaviour for reference parameter sets for which many realisations are generated by simulating the models starting from different initial conditions. The accuracy of results obtained in this way is unclear because exploring the sensitivity of these models by exhaustive simulations even for a subset of the parameter space is not only inefficient but for the most detailed models simply infeasible. Here, it is again valuable to consider how this problem is dealt with for models that are based upon differential equations. For systems of differential equations where the symbolic calculations required for bifurcation analysis are intractable or cumbersome, numerical continuation techniques are now widely used and software tools, for example, AUTO (Doedel, 2007)

and MATCONT (Dhooge et al., 2006a,b), are available. Continuation allows us to track a stationary solution as it varies depending on model parameters. This is obviously more efficient than simulating a model until it has reached a steady state but, more importantly, continuation can also track unstable stationary solutions. Continuation techniques cannot be directly applied to simulation models like IBMs but in the last 15 years—starting from Theodoropoulos et al. (2000)— Kevrekidis and co-workers have developed their framework of "equation-free" modelling and "coarse" bifurcation analysis, see Kevrekidis and Samaey (2009) for a review. The approach has mostly been applied to complex physical and chemical systems in engineering applications but there are a few recent examples where equation-free modelling has been used for analysing IBMs, see Cisternas et al. (2004); Tsoumanis et al. (2010) and Siettos (2011). In future work we will combine these techniques with our approach in order to systematically analyse more detailed IBMs.

Acknowledgements

This work was supported by Australian Federal and Victoria State Governments and the Australian Research Council through the ICT Centre of Excellence program, National ICT Australia (NICTA). I thank Samantha Ireland who helped implementing the two-strain version of the `Virus` model during her industry placement at NICTA in the High ICT achievers program of the Victorian Information Technology Teachers Association (VITTA). Helpful comments of two reviewers are gratefully acknowledged.

Armstrong, R. A., McGehee, R., 1976a. Coexistence of two competitors on one resource. Journal of Theoretical Biology 56 (2), 499–502.

Armstrong, R. A., McGehee, R., 1976b. Coexistence of two competitors on one resource. Theoretical Population Biology 9 (3), 317–328.

Armstrong, R. A., McGehee, R., 1980. Competitive exclusion. The American Naturalist 115 (2), 151–170.

Brauer, F., van den Driessche, P., Wu, J. (Eds.), 2008. Mathematical Epidemiology. Vol. 1945 of Lecture Notes in Mathematics. Springer, Berlin.

Cisternas, J., Gear, C. W., Levin, S., Kevrekidis, I. G., 2004. Equation-free modelling of evolving diseases: coarse-grained computations with individual-based models. Proceedings of the Royal Society A 460, 27612779.

Dhooge, A., Govaerts, W., Kuznetsov, Y., Mestrom, W., Riet, A., Sautois, B., 2006a. Matcont and cl_matcont: Continuation toolboxes in matlab. URL http://www.matcont.ugent.be/manual.pdf

Dhooge, A., Govaerts, W., Kuznetsov, Y., Sautois, B., 2006b. Matcont: A matlab package for dynamical systems with applications to neural activity. URL http://www.matcont.ugent.be/manual.pdf

Doedel, E. J., 2007. Lecture Notes on Numerical Analysis of Nonlinear Equations. In: Krauskopf, B., Osinga, H., Galán-Vioque, J. (Eds.), Numerical Continuation Methods for Dynamical Systems: Path following and boundary value problems. Springer Complexity. Springer, Dordrecht, The Netherlands, Ch. 1, pp. 1–49.

Eubank, S., Guclu, H., Kumar, V. S. A., Marathe, M. V., Srinivasan, A., Toroczkai, Z., Wang, N., 2004. Modelling disease outbreaks in realistic urban social networks. Nature 429 (6988), 180–184, cited By (since 1996)590. URL http://dx.doi.org/10.1038/nature02541

Ferguson, N. M., Cummings, D. A. T., Cauchemez, S., Fraser, C., Riley, S., Meeyai, A., Iamsirithaworn, S., Burke, D. S., 2005. Strategies for containing an emerging influenza pandemic in southeast asia. Nature 437 (7056), 209–214, cited By (since 1996)814.
URL http://dx.doi.org/10.1038/nature04017

Grimm, V., Railsback, S. F., 2005. Individual-based Modeling and Ecology. Princeton Series in Theoretical and Computational Biology. Princeton University Press, Princeton and Oxford.

Hardin, G., 1960. The competitive exclusion principle. Science 131, 1292–1298.

Keeling, M. J., Rohani, P., 2008. Modeling infectious diseases in humans and animals. Princeton University Press, Princeton.

Kevrekidis, I. G., Samaey, G., 2009. Equation-Free Multiscale Computation: Algorithms and Applications. Annual Reviews of Physical Chemistry 60, 321344.
URL http://dx.doi.org/10.1146/annurev.physchem.59.032607.093610

Kuznetsov, Y. A., 1995. Elements of applied bifurcation theory. Vol. 112 of Applied Mathematical Sciences. Springer, Berlin.

Lotka, A. J., 1925. Elements of physical biology. Williams and Wilkins, Baltimore.

Roche, B., Drake, J., Rohani, P., 2011. An Agent-Based Model to study the epidemiological and evolutionary dynamics of Influenza viruses. BMC Bioinformatics 12 (1), 87.
URL http://www.biomedcentral.com/1471-2105/12/87

Rolls, D. A., Daraganova, G., Sacks-Davis, R., Hellard, M., Jenkinson, R., McBryde, E., Pattison, P. E., Robins, G., 2012. Modelling hepatitis c transmission over a social network of injecting drug users. Journal of Theoretical Biology 297, 73–87.
URL http://www.sciencedirect.com/science/article/pii/S0022519311006175

Siekmann, I., 2013. On competition in ecology, epidemiology, and eco-epidemiology. Ecological Complexity 14 (1), 166–179.
URL http://dx.doi.org/10.1016/j.ecocom.2013.01.003

Siekmann, I., Malchow, H., Venturino, E., 2010. On competition of predators and prey infection. Ecological Complexity 7, 446–457.
URL http://dx.doi.org/10.1016/j.ecocom.2009.10.005

Siettos, C. I., 2011. Equation-free multiscale computational analysis of individual-based epidemic dynamics on networks. Applied Mathematics and Computation 218 (2), 324 – 336.
URL http://www.sciencedirect.com/science/article/pii/S009630031100748X

Smieszek, T., Balmer, M., Hattendorf, J., Axhausen, K., Zinsstag, J., Scholz, R., 2011. Reconstructing the 2003/2004 H3N2 influenza epidemic in Switzerland with a spatially explicit, individual-based model. BMC Infectious Diseases 11 (1), 115.
URL http://www.biomedcentral.com/1471-2334/11/115

Theodoropoulos, C., Qian, Y.-H., Kevrekidis, I. G., 2000. "Coarse" stability and bifurcation analysis using time-steppers: A reaction-diffusion example. PNAS 97 (18), 98409843.

Tsoumanis, A. C., Siettos, C. I., Bafas, G. V., Kevrekidis, I. G., 2010. Equation-free multiscale computations in social networks: from agent-based modeling to coarse-grained stability and bifurcation analysis. International Journal of Bifurcation and Chaos 20 (11), 3673–3688.

URL http://www.worldscientific.com/doi/abs/10.1142/S0218127410027945

Volterra, V., 1926. Fluctuations in the abundance of a species considered mathematically. Nature 118, 558–560.

Wiggins, S., 2003. Introduction to applied nonlinear dynamical systems and chaos, 2nd Edition. Vol. 2 of Texts in Applied Mathematics. Springer, New York.

Wilensky, U., 1998. NetLogo Virus model. Center for Connected Learning and Computer-Based Modeling, Northwestern University. Evanston, IL. USA.
URL http://ccl.northwestern.edu/netlogo/

Wilensky, U., 1999. NetLogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University. Evanston, IL. USA.
URL http://ccl.northwestern.edu/netlogo/