Biological Modeling: Research projects

At this time point in our course you have learned enough to make your first steps as a modeler in biology. You have been trained to read and formulate mathematical models, and to analyze these graphically, mathematically, and numerically. You now have a few weeks to work on a somewhat larger project by yourself. The main goal of this project is that you apply your newly acquired skills to do a piece of research of your own. Most of the projects that we propose here start with a recent paper, but we also welcome proposals to work on a project of your own (please contact us to discuss this).

For the projects starting with a paper, you should first repeat its results by coding its model(s) into Grind. Next, find other papers, and define extensions of the research yourself. It is your own responsibility to study exciting new extensions of the work, and to make this a truly interesting project.

Please work in small groups of no more than four. Every project will be assigned an assistant whom you can contact for help: make an appointment by email to meet on MS-Teams (or another online platform). Make a concise “lab-journal” in which you shortly describe your progress every day that you worked on the project. This journal has to be handed in together with your written report.

Thursday 5 November every group will present their work using an online electronic slide show. Each of the participants should speak (in English) for about 5 minutes and the audience will be composed of your personal assistant plus at least one other assistant (probably Rob or Peter). The oral presentation should be exciting for the audience. You have to make clear what questions you addressed, how this was approached, why it is interesting, and what results you obtained. Too technical details should be avoided (as these will be explained in the written report). Subdivide your presentation into natural parts such that all members of your research team get to speak! Your presentation should be enthusiastic and strongly focus on your main line of research.

The written report has to be delivered in the last week of the course (electronic submissions of PDF-files by email to r.j.deboer@uu.nl are accepted until Sunday the 8th November). The report should have a summary, and start with an Introduction explaining the project, its context, and have a short review of the relevant literature. In the Methods section you can define and explain the mathematical model. In the Results section you can mix the results of the original paper with your own extensions of the research, and you can provide your extensions of the mathematical model. Provide the interpretation of your results in the Results section, i.e., write in a Nature or Science style. In the Discussion you describe possible problems/shortcomings, other extensions, and you provide further context of your work. Use the instructions on writing reports that you received in earlier courses, and carefully read the short tutorial on writing scientific reports that we provide: tbb.bio.uu.nl/rdb/bm/Projects/GuidelinesReport.pdf.

We will give written feedback on your oral and written presentation later by email. Please make an appointment to discuss that feedback with us. With the project we hope to increase your experience with the techniques you have encountered so far, and to show that you have arrived at a stage where you can critically continue the work of recent papers in this field.
1 The carbon cycle

We all know that atmospheric CO2 levels are increasing and to prevent global warming huge investments are being made to reduce the release of CO2 by burning fossil fuels. Estimates the human CO2 production, and of the natural flux of CO2 between the atmosphere, the upper and deep layers of the ocean, and the vegetation are published on various websites (e.g., earthobservatory.nasa.gov/features/CarbonCycle). The current yearly human production of about 9 giga ton per year is much smaller than the natural flux from the ocean (90), vegetation (60), and soil (60) into the atmosphere. So how can such a small increase give rise to such a large increase in atmospheric CO2 levels? This is being discussed on various websites (e.g., skepticscience.com/co2-residence-time.htm).

Find papers on this topic and find out where the data in this NASA diagram is based upon. Study this question by making a compartment model, for the main CO2 compartments (atmosphere, ocean, deep ocean, and vegetation (first including the soil?). For instance, for the amount of CO2 in the atmosphere one would write something like,

\[
\frac{dA}{dt} = f_{\text{OA}}O - f_{\text{AO}}A + f_{\text{VA}}V - f_{\text{AV}}A + h ,
\]

where we have four flux parameters (to and from the ocean O, and to a from the vegetation V), and the human production h (in kilo tons per year). Note that the terms like \(f_{\text{OA}}O\) correspond to the total fluxes depicted in the diagrams, whereas the parameter \(f_{\text{OA}}\) of this model is therefore defined as a relative flux. Write equations for the other compartments, fill in parameters, and study what increase in atmospheric CO2 levels one would predict if the parameter \(h\) goes from 0 to 9 kilo tons per year.

This project requires a literature search, model development, and numerical simulations.

2 COVID-19 herd immunity in the Brazilian Amazon

During the exam you made a question about a COVID-19 epidemic in Manaus (Brazil). The data used in that question came from a preprint paper that is currently available on medRXiv: www.medrxiv.org/content/10.1101/2020.09.16.20194787v1. On the exam we argued that this data can be modeled with a classic SIR model,

\[
\frac{dS}{dt} = -\beta \frac{SI}{N} , \quad \frac{dI}{dt} = \beta \frac{SI}{N} - (d + r)I \quad \text{and} \quad \frac{dR}{dt} = rI ,
\]

with susceptible individuals, \(S\), infected individuals, \(I\), and recovered individuals, \(R\). Here \(N = S + I + R\) is the total number of individuals, \(\beta\) the infection rate, \(r\) the recovery rate, and \(d\) the death rate due to COVID-19. In this project will we test whether or not this data can indeed be described with this model. We provide the data and the R-script on the subdirectory [models/manaus] on the website of the course. Because the model may have more parameters than can be estimated from the data we start with using their estimate on the \(R_0 = \frac{\beta}{r+d} \simeq \frac{\beta}{r} = 2.5\), and estimated time to recovery of 10 days, and we fix the death rate using their estimate that the disease fatality is 0.2%. Do you obtain a good description of the data, and do you obtain realistic parameter values? In the preprint they also provide data on the percentages of seroconverted individuals: what does your model predict for this? Change the number of free parameters to study which parameters are identifiable.

Since the number of seroconverted individuals is slowly declining, and the number of infections seems to be continuing at a very low level, you should extend your model allowing recovered individuals to seroconvert and become susceptible again. Does this extended model describe the data better? Read commentary on [www.reuters.com/article/us-health-coronavirus-brazil-manaus/in-brazils-amazon-a-covid-19-resurgence-dashes-herd-immunity-hopes-idUSKBN26I0I4] to put your results in perspective.
The fitting of this model to the data is complicated and takes a lot of time.

3 Transient phenomena in ecology

A recent paper published in Science, Hastings et al. (2018), reviews evidence suggesting that the dynamics of many ecosystems are typically long-term transients, and discusses how this relates to predicting sudden changes in the dynamics in the absence of changes in the environmental conditions. Repeat the examples that they provide, and explain their classification of dynamical regimes to your audience. Don’t forget to include the effect of noise.

Discuss the differences between the classical saddle-node approaches (Scheffer et al., 2009), and the examples discussed in this paper.

Somewhat technical, good for those interested in attractors and bifurcations in non-linear dynamical systems.

4 Co-existence by trade-offs?

Posfai et al. (2017) study the “Paradox of the plankton” by modeling resource competition between a large number of consumers. Their major idea is that consumers are expected to specialize on a subset of the resources, and therefore they introduce trade-offs among the consumption rates when parametrizing their model. Surprisingly, they find that an unlimited number of species can coexist, and that their model reproduces several features of natural ecosystems, including keystone species and population dynamics characteristic of neutral theory. The consumer equation of their model takes the following form:

$$\frac{dN_i}{dt} = \left( \sum_j \beta_{ij} c_{ij} R_j - \delta_i \right) N_i,$$

where each additional resource increases the maximum birth rate, $\beta_i = \sum_j \beta_{ij}$, that is approached when all resources are available at large densities, i.e., when $R_j \gg h_{ij}$ for all $j$ (which would be the natural situation when all consumer densities are low).

This project appears in the book as Question 9.8: to get started on the project make sure you first complete this question. Discuss whether or not you find this a proper model for substitutable resources or for essential resources. Second, use Grind to study the idea of a trade-off in our own models for competition for substitutable resources. Simplify their analysis by considering just two resources. Do you find similar results, and—if so— what is actually required to repeat these results? Do you expect different results when there are three (or more) resources? Third, try the same trade-off in our models for essential resources. Do obtain the same results? What do you think of this paper: is this indeed resolving the Paradox of the plankton?

Note that the same group has a novel manuscript about this model (Erez et al. (2020)), which is praised on the Princeton new bulletin: [www.princeton.edu/news/2020/09/18/princeton-scientists-explain-how-diverse-species-coexist-microbial-communities](http://www.princeton.edu/news/2020/09/18/princeton-scientists-explain-how-diverse-species-coexist-microbial-communities).

Good for students liking Tilman diagrams. Parts of the papers are somewhat mathematical, the main idea is explained above, however.
5 Temperate phages

Generally, bacteriophages infect bacterial cells by injecting their DNA into them. The viral DNA is transcribed to make many copies of viral proteins, such that numerous copies of the infecting particle are assembled. After a while the bacterial cell bursts and releases viral particles. This is the lytic life-cycle of the bacteriophage. Temperate bacteriophages can choose a different life-cycle, where they integrate into the DNA of the host cell. This “prophage” remains dormant and is passed on with every bacterial division. This is called the lysogenic life cycle. At some point the prophage can be induced to resume the lytic form of the life cycle, and start to produce proteins and burst the cell. The evolutionary difference between these two life cycles has been studied extensively (Gandon, 2016; Berngruber et al., 2013). A recent twist to the decision between lysis and lysogeny is a paper by Erez et al. (2017) demonstrating that \( \phi \)T phages phages “communicate” by producing a signal peptide upon infecting a cell (see Hynes & Moineau (2017) for a commentary). The concentration of this peptide increases the propensity for the lysogenic life-cycle of subsequent infections. A very recent paper Maldonado-Barragan & West (in press) studies the evolutionary costs of this.

In the course we have used chemostat equations to model bacterial growth. To simplify we first follow Berngruber et al. (2013) and use Logistic growth to formulate a simple mathematical model for temperate phages:

\[
\frac{dS}{dt} = rS(1 - S - L) - mS - \beta SV ,
\]

\[
\frac{dL}{dt} = rL(1 - S - L) - mL + \phi \beta SV - \alpha L ,
\]

\[
\frac{dV}{dt} = \alpha L + (1 - \phi) \beta SV - mV ,
\]

where \( S \) is the density of uninfected bacteria, \( L \) the density of lysogens (i.e., bacteria with a prophage), and \( V \) the density of the phages. The parameter \( r \) denotes the replication rate of the bacteria, and \( m \) (for mortality) is the loss of bacteria and phages (that is largely due to wash-out from the chemostat). We have simplified the model by ignoring the loss of phages by absorption to bacteria (you may want to put this back). Thus, \( \beta \) is the mass-action infection rate, a fraction \( \phi \) of the infections is lysogenic, and \( \alpha \) is the induction rate. A second simplification is that we have scaled the density of the phages by their burst rate.

First use this model to study the lysis-lysogeny decision. Can you find an “optimal” value of \( \phi \)? One approach to study this is to write a model for two strains of the virus, that differ in \( \phi \), while assuming that lysogens cannot be super-infected (see Berngruber et al. (2013)). Second, extend the model with an equation for the peptide and make \( \phi \) a function depending on its concentration. Do phages responding to this peptide outcompete phages that don’t? Would it make a difference to model this as a chemostat (rather than by logistic growth)?

This is a fairly numerical project, Grinding various different models.

6 Density dependent predation

In a recent review Terborgh (2015) summarized the importance of predation and disease in the maintenance of diversity of ecosystems. His article reviews several theories of species diversity, and his main take home message is that “keystone” predators killing the most abundant species play an essential role in ecosystem diversity. He calls this a “top-down” forcing of the food web (as opposed to the “bottom-up” theories on competition).
Read the paper and define for yourself what he means by density dependent predation. Study simple ODE models implementing various forms of density dependent predation, and use Grind to explore the effects of density dependent predation on competitive exclusion.

For instance, how many prey species can co-exist by top-down control of one predator? How many predators can co-exist on a few resources, and could there be a feedback where an increase of prey diversity allows for an increase in predator diversity, which in turn increases the prey diversity?

This is an ecological project, somewhat related to the next one on the effect of pathogens.

### 7 Competitive exclusion and parasitism

Many species suffer from a heavy burden with pathogens (Dobson et al., 2008). If pathogens truly control population densities, this may increase ecosystem diversity by reducing competitive exclusion. This is often referred to as the Janzen-Connell hypothesis, e.g., in Sedio & Ostling (2013) and Bagchi et al. (2014). In the course we considered several populations of bird species with a birth rate declining linearly with the population size, and with a death rate that is independent of the population density. We let the individuals be susceptible to an infection with a parasite that increases the death rate somewhat, but hardly affects the birth rate. We assumed that transmission of parasites occurs upon contacts between infected and susceptible individuals of the same species, and obeys mass action kinetics. Further there was no vertical transmission, i.e., the parasite is not transmitted to eggs. Thus, we let $N_j = S_j + I_j$ be the total number of birds, $S_j$ be the susceptible non-infected birds, and $I_j$ be the infected birds of the $j^{th}$ species:

$$\frac{dS_j}{dt} = bN_j(1 - N_j/k) - d_jS_j - \beta S_jI_j$$

and

$$\frac{dI_j}{dt} = \beta S_jI_j - (d_j + \delta)I_j,$$

where $\delta$ reflects the deleterious effect of the infection.

First analyze a 2-dimensional system, i.e., let $j = 1$ and consider one species. Second, study how many new species you can add to this one-species ecosystem assuming that (1) all bird species occupy the same niche, and (2) every new species has a faster death rate, i.e., a lower fitness, than the previous one ($d_{j+1} > d_j$). Make a simple function describing how $d_j$ depends on $j$. Note that you can define vectors of equations in Grind (see the tutorial).

The Janzen-Connell hypothesis typically states that pathogens are expected to evolve towards infecting the most abundant species. This is called negative density dependence (Bagchi et al., 2014). Can you modify this model to study such effects of pathogen evolution? Other studies suggest that host and pathogen diversity in a community may also affect the infection rates (Johnson et al., 2013). How would that affect these results?

The ‘top-down’ control by pathogens in this project makes it overlap with the previous one. Here we need not have density dependent predation, however.

### 8 Ontogenetic development for dummies

Persson & De Roos (2013) and De Roos & Persson (2013) summarize their extensive work on the effects of having juveniles and adults with different energetic requirements. These surprising effects include increases of the population size when the death rate increases, implicit Allee effects, and several
more. The use both ODEs and PDEs for the modeling of the age dependent growth of the biomass of adults and juveniles, and these models are fairly complicated.

The aim of this project is to see whether their interesting effects can also be found in more simple (phenomenological) models, e.g.,

\[
\begin{align*}
R &= K - c_1 J - c_2 A, \\
\frac{dJ}{dt} &= \frac{eAR}{h_2 + R} - \frac{mJR}{h_1 + R} - \mu d_1 J \\
\text{and} \quad \frac{dA}{dt} &= \frac{mJR}{h_1 + R} - \mu d_2 A
\end{align*}
\]

where \( R \) is the available amount of resource, \( K \) the total, and \( c_1 \) and \( c_2 \) determine how much stored in juveniles, \( J \), and adults, \( A \). The rates at which juveniles mature, and the rate at which adults produce juveniles, depend on the availability of the resource. With the two \( h_i \) parameters one can change the symmetry of this dependence on the resource (\( h_1 = h_2 \) would be a conventional symmetric system). With the parameter \( \mu \) one can increase the death rate of both juveniles and adults simultaneously.

Read their paper and try to repeat as much of their results with this toy model. You may also enjoy watching these lectures: staff.fnwi.uva.nl/a.m.deroos/Research/Webinars/.

This is a somewhat numerical project, introducing you to quite a different class of models for population dynamics.

9 Long term effects of vaccination

Holdo et al. (2009) investigate the limiting factors determining the wildebeest population size in the Serengeti ecosystem in East Africa (see also the primer by Getz (2009)). Possible factors are the tree cover, which is related to rainfall and frequent fires, disease outbreaks, and competing herbivores like elephants. They study this by analyzing long time series (1960-2003) by fitting statistical models. At the start of this period the wildebeest were vaccinated to rinderpest, and as a consequence the wildebeest population increased. Rinderpest was eradicated in 2012 and is the second pathogen that went extinct due to our vaccination efforts.

A similar project can be made on the basis of a review paper by Peterson et al. (2014) discussing several long term cascades in Yellowstone national park (in a systems with bears and wolves). The advantage of this paper is that they provide several good time series of the population densities.

See if you can describe the outcome of all these interactions with simple ODE models (there are no models in these papers). Note that you can introduce environmental variation, like fires, by allowing for noise on some of the parameters (see the Grind tutorial).

This project requires the development of a relatively large model, with a lot of parameters. Implement fires is not trivial, and can be done by introducing a stochastic term.

10 Early warning signals in the Sahel

The notion that we might be able to observe “early warning” signals in time series data of systems that are about to collapse is receiving a lot of attention recently (Scheffer et al., 2009; Veraart et al., 2012; Scheffer et al., 2012). This theory is based upon the simple fact that when a system approaches a catastrophic bifurcation (like a saddle-node bifurcation) the dominant eigenvalue is approaching zero, implying that the return time of the system to its steady state is becoming very long. Thus, an increasing return time of a system under slowly changing environmental conditions could provide
a warning signal for an upcoming catastrophe. It would be extremely important if one could indeed detect such early warning signals in the time series of any particular system, because one could change the environmental conditions to prevent a future disaster. Long return times should be associated with more variation in the data, and to a better correlation between subsequent data points. The review paper by Scheffer et al. (2009) provides interesting examples of early warning signals in biological data, and clearly explains the underlying theory in several boxes. Read this paper before you embark on this exercise. In the book this project was introduced as Question 12.8, where you may have repeated their analysis using a logistically growing resource harvested according to a sigmoid functional response, i.e., $\frac{dX}{dt} = X(1 - X/K) - cX^2/(1 + X^2)$.

Now we ask you to consider a more real-world example. Start a model describing the dynamics of water uptake in arid zones (Rietkerk & Van de Koppel, 1997; HilleRisLambers et al., 2001), that you may have seen in the first year course. Overgrazing by cattle in arid areas is known to lead to desertification. In the Sahel zone one may find both barren areas and vegetated areas in the same region. This bistability has been studied with models having saddle-node bifurcations (Noy-Meir, 1975; Rietkerk & Van de Koppel, 1997; HilleRisLambers et al., 2001). The main idea is that in areas with little vegetation coverage most of the (sometimes heavy) rainfall fails to penetrate into the soil, and is rapidly washed off into rivers and disappears. In areas with somewhat more vegetation water is better captured, but vegetation also consumes water by growth. Models for vegetation growth in arid areas can remain simple because the availability of water in the soil is typically the major limiting factor. This main idea is translated into the following model:

$$\frac{dW}{dt} = R\left(w_0 + \frac{V}{k_2 + V}\right) - r_W W - \frac{gVW}{k_1 + W},$$

$$\frac{dV}{dt} = i + \frac{cgVW}{k_1 + W} - dV - hV,$$

where $R$ is the rainfall (in mm d$^{-1}$), $V$ is the vegetation biomass (in g m$^{-2}$), and $W$ is the amount of water in the soil (mm). The model has two saturation constants, $k_1$ (mm) defines the amount of water at which the vegetation grows at half its maximal rate, and $k_2$ (g m$^{-2}$) is the vegetation cover at which the penetration of water into the soil is $R(w_0 + 1/2) = 1.4$ mm d$^{-1}$. The parameter $h$ denotes the grazing by cattle, which depends on the herd size that is set by people buying and selling cattle. Parameters of this model have been estimated by Rietkerk & Van de Koppel (1997) and HilleRisLambers et al. (2001):

<table>
<thead>
<tr>
<th>Name</th>
<th>interpretation</th>
<th>value</th>
<th>dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R$</td>
<td>rainfall</td>
<td>2</td>
<td>mm d$^{-1}$</td>
</tr>
<tr>
<td>$c$</td>
<td>conversion of water to plant biomass</td>
<td>10</td>
<td>g mm$^{-1}$m$^{-2}$</td>
</tr>
<tr>
<td>$d$</td>
<td>death rate of vegetation</td>
<td>0.05</td>
<td>d$^{-1}$</td>
</tr>
<tr>
<td>$g$</td>
<td>maximum water uptake</td>
<td>0.05</td>
<td>mm g$^{-1}$m$^2$ d$^{-1}$</td>
</tr>
<tr>
<td>$k_1$</td>
<td>half saturation constant</td>
<td>5</td>
<td>mm</td>
</tr>
<tr>
<td>$k_2$</td>
<td>half saturation constant</td>
<td>5</td>
<td>g m$^{-2}$</td>
</tr>
<tr>
<td>$r_W$</td>
<td>soil water loss due to evaporation</td>
<td>0.2</td>
<td>d$^{-1}$</td>
</tr>
<tr>
<td>$w_0$</td>
<td>water infiltration in absence of vegetation</td>
<td>0.2</td>
<td>—</td>
</tr>
<tr>
<td>$i$</td>
<td>immigration of seeds/plants</td>
<td>0.01</td>
<td>g m$^{-2}$ d$^{-1}$</td>
</tr>
</tbody>
</table>

In the absence of a vegetation there is about 2 mm water in the soil. The death rate of the vegetation is partly due to normal turnover and partly due to grazing by cattle. Assume that in the absence of cattle the vegetation turnover is about $d = 0.05$ d$^{-1}$, resulting in a carrying capacity of approximately 450 g m$^{-2}$. Study the effect of increasing the herd size, $c$, starting with just a small herd. The model is available on the website under the name sahel.R: Suppose the local community has decided on a safe and stable herd size, $c$, and that the environment is faced with a few years of declining rainfall. Can you predict when they should start decreasing the herd size?
Finally, compare your results to those of Boerlijst et al. (2013): do the variables in this model differ in the early warning signals they provide?

11 Influenza infections

Every year many people become vaccinated with a current influenza vaccine. It is a major challenge to design such a vaccine because influenza is evolving, and differs from year to year. It has been described that if a major fraction of the population is immune to the currently dominant strain, a new strain may evolve during the season for which most of the population has no immunity. This is called “strain replacement”. People may even be infected with both strains during a season because the crossreactive immunity for both strains is short-lived. See Furuse & Oshitani (2016) for a recent paper on this topic.

Zarnitsyna et al. (2018) write an SIR model allowing for strain replacement and surprisingly conclude that intermediate levels of vaccination coverage may minimize seasonal influenza outbreaks. Repeat their results and study how the size of the total epidemic depends on the fraction of people that are vaccinated before the season starts. A cool paper to discuss during your presentation is Smith et al. (2004), who depict the evolution of influenza in a 2-dimensional antigenic map.

This is an interesting project given the current COVID-19 pandemic.

12 Stem cell renewal

Many tissues and populations of cells are maintained by a subpopulation of stem cells. A classic example is the formation of several populations of circulating cells in the blood by a relatively small population of hematopoietic stem cells (HSCs) in the bone marrow, or the stem cells located deep in the crypts of the epithelial layer lining the gut. Stem cells are self-renewing cells that (at least sometimes) divide into two different daughter cells, which is called an asymmetric division. When a sufficient fraction of their daughter cells remains as a stem cell, the stem cell population can maintain itself, and provide progeny to the populations of differentiated cells that depend on it. It is unclear how stem cells regulate the fraction of asymmetric divisions, as at least half of their cell divisions should deliver a daughter with stem cell properties. Otherwise the stem cell population declines. The first question of this exercise is to write a model for the simplest situation where on average half of the daughter cells remains a stem cell while the other half differentiates. You will see that this fraction should be more than one half to compensate for the death of stem cells, and hence that it is unclear how the fraction of asymmetric divisions is regulated.

Lander et al. (2009) developed an interesting model for this problem by arguing that the fraction of renewal divisions delivering a daughter with stem cell properties should depend on the density of the population. This could either be the total density, i.e., stem cells plus differentiated cells, the density of differentiated cells, or the density of stem cells. This is interesting because this entails the population with at least two density dependent mechanisms, one regulating the fraction of self-renewal divisions, and another regulating the rate of cell division. They develop a chain of equations where at every level cells may divide asymmetrically, and Lander et al. (2009) show that the parameters of that system determine which population in the chain will function as stem cells for the entire chain. Here we simplify their model by considering just two populations: stem cells, $S$, and differentiated cells, $D$. In the questions below we ask you to devise a model where both the fraction of asymmetric divisions, $0 < f(D) \leq 1$, and the division rate of the stem cells, $g(D)$ for growth, depends on the density of differentiated cells. Following Lander et al. (2009) suggestions it would be natural to allow
for a larger fraction of asymmetric divisions, and a higher division rate, when the population is small and the tissue should be regenerated.

This project is introduced in the book as Question 12.6, and a good way to get started is to make questions 12.6a–e. Next, the main question is to address what the stem cells obtain by having two homeostatic mechanisms (one for the fraction of asymmetric divisions and one for the division rate). Study how rapidly a damaged tissue would be repopulated in the absence and presence of the two mechanisms. Lander et al. (2009) also consider more complicated models where the differentiated cells can also undergo asymmetric divisions. It then depends on the parameters of the various homeostatic functions which of the two populations becomes the stem cell. Implement this and study if and how this affects the results.

This project requires a good combination of understanding equations and performing numerical analysis.

13 Cryptic oscillations

In Question 5.7 of the book we depict the classic experiments of Bohannan & Lenski (1997) demonstrating that when E. coli is cultured with its bacteriophage T4, the system can initially display predator-prey like oscillations, and later develop so-called “cryptic” oscillations where the phage densities continue to oscillate, but the E. coli densities become stable. They study this in more detail and put this in a more general perspective in their later paper Bohannan & Lenski (1999). This data is also discussed and modeled in Figure 4.6 of the excellent book of Weitz (2015).

Repeat the Bohannan & Lenski (1999) paper to see if you can get cryptic oscillations. They employ the Levin et al. (1977) model that defines the eclipse phase by a time delay. Do you obtain similar results in models without this time delay (i.e., is a DDE required to obtain cryptic oscillations)? Discuss whether or not you agree that this happens only when ecological and evolutionary processes occur on a similar time scale.

This is a fairly numerical project, finding parameters and equations to obtain the desired results.

14 Lymphocyte migration

Naive T cells in the immune system circulate via the blood between various lymphoid organs, such as the spleen and many different lymph nodes. The residence time in the blood is short, and under normal conditions only a small percentage of the naive T cells reside in the blood. In the spleen the residence time is about 6 hours and a typical residence time for a peripheral lymph node is about 13.5 hours (Textor et al., 2014). A simple model considering the number, $N$, of naive T cells specific for a particular antigen (e.g., a peptide derived from a virus) would be

$$B = N - S - L, \quad \frac{dS}{dt} = i_S B - e_S S \quad \text{and} \quad \frac{dL}{dt} = n_i L - e_L L,$$

where $B$ is the number of cognate naive T cells in the blood, $S$ is the number of cells in the spleen, $L$ is the total number of cells in all $n$ lymph nodes, and the $i$ and $e$ parameters are influx rates and efflux rates into the lymphoid organs. Textor et al. (2014) estimate that $i_S = 1h^{-1}$ and that $ni_L = 1.5h^{-1}$, i.e., most cells leave the blood by entering a lymph node, but the spleen is the organ receiving the vast majority of naive T cells from the blood. Estimating $n$ is difficult because lymph nodes have different volumes, and there are many small “nodes” like Peyers patches in the gut. Textor et al. (2014) estimate that in a mouse there are about $n = 39$ major lymph nodes.
When an organism is challenged by a pathogen in one of its tissues, proteins from the pathogen are transported to the lymph node(s) draining this tissue, and this triggers an immune response by the cognate naive T cells that are present in this lymph node(s). Since there are so many lymph nodes, only a small fraction, i.e., about $1/n$, of the naive T cells is expected be present in the draining lymph node. However, various experiments have demonstrated that during a localized infection almost all cognate naive T cells are recruited into the immune response on a time scale of a few days. Since cognate naive T cells are trapped in the draining lymph node when they find their antigen there, one would predict that they will slowly accumulate there. This accumulation is expected to be too slow because naive T cells exiting from another lymph node have only a $1/n$ probability to arrive in the draining lymph node, and therefore probably end up in another node, where they are expected to spend another 13.5 hours (Textor et al., 2014). In one of the questions below you will study how long this would take. The real solution for recruiting most naive T cells is to enlarge the influx into the draining lymph node, and it is quite spectacular how inflamed lymph nodes can increase their blood supply by angiogenesis. To study this we extend the model by assigning one of the lymph nodes as the draining lymph node, and provide this one lymph node with parameters $f_i$ and $f_e$ to modify its influx and efflux upon infection,

$$\frac{dS}{dt} = iSB - eSS,$$
$$\frac{dL}{dt} = (n - 1)iLB - eLL$$
$$\frac{dD}{dt} = f_iLB - f_eeLD,$$

(2)

where $B = N - S - L - D$ and $f_i = f_e = 1$ when there is no infection. This model is available as the file circulation.R and is presented in the book as Question 12.5.

Make Question 12.5 to make sure you fully understand the model and the Textor et al. (2014) paper. Since the total number of lymph nodes differs between mice and men, and remains somewhat difficult to define as there are many small “lymph nodes” like Peyers patches in the gut, it would be good to generalize these results and study how rapidly or how much the influx to the relevant draining lymph node should increase as a function of the total number of lymph nodes, $n$.

15 Evolution of virulence

The evolution of infectious diseases has attracted the attention of evolutionary biologists and epidemiologists for decades. The predominant model is based upon the trade-off between virulence and transmissibility, i.e., virulent pathogens typically transmit better per encounter, but because they shorten the lifespan on their host, they have fewer encounters with susceptible hosts. See Fraser et al. (2014) and Ebert & Bull (2003) for a reviews on this topic.

Study the evolution of virulence in a SI model where you make the infection rate, $\beta$, and saturation function of the virulence, $v$:

$$\frac{dS}{dt} = s - dS - \beta SI$$
$$\frac{dI}{dt} = \beta SI - (d + v)I$$

where $\beta = \frac{cv}{h + v}$.

Consider strains, $I_j$, differing in virulence, $v_j$, and hence $\beta_j$ to investigate what to predict on the long run.

Study steady state situations and long-term transients with re-merging epidemics. Under what circumstances do pathogens evolve to become more benign to their host?

References


