# Population Dynamics

http://www.maths.warwick.ac.uk/~keeling/Teaching_index.html
m.j.keeling@warwick.ac.uk

**LAYOUT OF THE COURSE**

This lecture course is divided into eleven sections. Each section generally comes in two parts. In the first part I'll describe the basic mechanisms, concepts and results, in the second part (generally on Thursdays) we will focus our attention on the recent literature to understand how these ideas are used in practice. Most publications listed in this section are available on my web page, any that are not are marked with an asterisk – you will have to photocopy these from the library. Those papers cited within the main text are for reference only.

• **Introduction** Lecture 1

In this first lecture I'll explain how the course will work, and hopefully provide sufficient motivation for you to understand why mathematical modelling has a vital role to play in population dynamics and epidemiology.

• **Population Models** Lectures 2 & 3

These basic models of the behaviour of simple populations utilise the standard tools of dynamical systems (eg stability etc). Particular attention is paid to what biological understanding can be gained and the assumptions that underlie the models. Major topics include the Logistic growth model, Predator-Prey models, and Lokta-Volterra competition.

• **Disease Models** Lectures 4-6

The basic models of infectious disease dynamics (the SIS and SIR models) are reviewed. Again, attention is paid to what biological understanding can be gained and the assumptions that underlie the models. The heterogeneities that arise when these assumptions are relaxed are also discussed.

• **Temporal Forcing** Lectures 7-9

The vast majority of organisms in temporal climates breed at fixed times of the year. Others (such as many insects) may die before next years young become adults. In such situations, discrete time models of their dynamics may offer the best description, the most famous example of this is the Nicholson-Bailey model of host-parasitoid interactions. Many disease models are affected by seasonal variations. For example colds and 'flu are most common in the winter, whereas disease of animals may be influenced by patterns of seasonal births. The most commonly studied time of seasonal forcing is the effects on childhood diseases of the opening and closing of schools.

Keeling, Rohani and Grenfell (2001) Seasonally-forced Disease Dynamics Explored as Switching Between Attractors *Physica D* **148** 317-335

Finkenstädt, B. and Grenfell, B. (2000) Time series modelling of childhood diseases: a dynamical systems approach. *J. Roy. Stat. Soc. C* **49** 187-205.

Grenfell, B.T., Wilson, K., Finkenstädt, B.F., Coulson, T.N., Murray, S., Albon, S.D., Pemberton, J.M., Clutton-Brock, T.H. and Crawley, M.J. 1998 Noise and determinism in synchronized sheep dynamics *Nature* **394** 674-677

Hassell, M.P. and Pacala, S.W. 1990 Heterogeneity and the dynamics of host-parasitoid interactions. *Phil. Trans. Roy. Soc. Lond. B* **330** 203-220

• **Age/Risk Structure** Lectures 10-12

Initially we concentrate on sexually transmitted diseases, where the number of different sexual

partners determines the risk of infection and the risk of transmission. In such risk structured models the basic relationships no-longer hold, and we consider $R_0$ in detail. We also look at adding age-structure to the standard models, and consider why measles and whooping cough are childhood diseases. The implications of non-random mixing and the who acquires infection from whom matrix are discussed. We consider how the average age of infection can change and what effects this may have on disease severity, including the idea of "endemic stability". From an ecological perspective we consider the effects of size and age classification on the interaction of organisms. In particular the cannibalism of large adult fish on much smaller fry will be considered.

***** Anderson, R.M., Medley, G.F., May, R.M. and Johnson, A.M. (1986) A Preliminary Study of the Transmission Dynamics of the Human Immunodeficiency Virus (HIV), the Causative Agent of AIDS *IMA J. Math. App. Med. Biol.* **3** 229-263.

***** Schenzle, D. (1984) An age-structured model of pre- and post-vaccination measles transmission. *IMA. J. Math. App. Med. Biol.* **1** 169-191.

De Roos, A.M. and Persson, L. 2002 Size-dependent life-history traits promote catastrophic collapses of top predators. *P.N.A.S.* **99** 12907-12912.

De Roos, A.M., Persson, L. and Thieme, H.R. 2003 Emergent Allee effects in top predators feeding on structured prey populations. *Proc. Roy. Soc. Lond. B* **270** 611-618.

• **Multi-Host / Multi-Strain** Lectures 13-15

So far all population models have considered the interaction of one or two species. Here we consider the interaction of three or more species, focusing on examples of plankton dynamics and lynx-hare cycles to illustrate the range of complex behaviour. Diseases and hosts do not exist in isolation. Many diseases can infect multiple hosts, all hosts have many disease and strains of diseases that can infect them. Here we review the models necessary to understand such situations, concentrating primarily on the dynamics of strain structure.

So far we have only considered micro-parasites (diseases caused by viruses and bacteria). As the name suggests macro-parasites are much larger and generally have more complex life-cycles; this necessitates more complex models. In this lecture we review the basic elements of macro-parasite models.

Gupta S, Trenholme K, Anderson RM, and Day, K.P. (1994) Antigenic Diversity and the Transmission Dynamics of Plasmodium-Falciparum. *Science* **263** 961-963

Gog J.R., and Swinton J. (2002) A status-based approach to multiple strain dynamics. *J. Math. Biol.* **44** 169-184.

Rohani, P., Earn, D.J., Finkenstadt, B. and Grenfell, B.T. 1998 Population dynamic interference among childhood diseases *Proc. Roy. Soc. B* **265** 2033-2041

Blasius, B., Huppert, A. and Stone, L. 1999 Complex dynamics and phase synchronization in spatially extended ecological systems *Nature* **399** 354-359

• **Stochasticity** Lectures 16-18

All the models we have considered so far are deterministic, clock-work models with no variability. In practice, populations are governed by chance with their dynamics described by a series of random events. We will consider a range of models that mimic this behaviour including event-driven models. We will see how randomness can influence the mean and variance of the dynamics, as well as leading to the risk of extinction for both animal and disease populations.

Bartlett, M.S. (1956) Deterministic and Stochastic Models for recurrent epidemics. *Proc. of the Third Berkley Symp. on Math. Stats. and Prob.* **4** 81-108.

Nasell, I. (2002) Stochastic models of some endemic infections *Math. Biosci.* **179** 1-19.

Nasell, I. (2003) Moment closure and the stochastic logistic model *Theo. Pop. Biol.* **63** 159-168.

• **Spatial Heterogeneity** Lectures 19-24

The populations are aggregated and segregated at a variety of different scales. At the most local level, the interaction of individuals or the transmission of infection is generally by close contact – we examine the travelling waves that arise in this situation. At a larger scale, humans (and many animals) form patchy distributions such that there are areas of high density (which are prone to disease) separated by regions of low density (which act as barriers to disease spread). A wide variety of modelling techniques and tools will be considered.

Murray, J.D., Stanley, E.A. and Brown, D.L. 1986 On the spatial spread of rabies by foxes. *Proceedings of the Royal Society of London B* **229** 111-150

Rhodes, C.J. and Anderson, R.M. 1996 Evaluation of Epidemic thresholds in a lattice model of disease spread *Phys. Letters A* **210** 183-188

Keeling, M.J. 1999 The Effects of Local Spatial Structure on Epidemiological Invasions *Proc. Roy. Soc. Lond. B* **266** 859-869

Keeling, M.J. *et. al.* 2001 Dynamics of the 2001 UK Foot and Mouth Epidemic: Stochastic Dispersal in a Heterogeneous Landscape *Science* **294** 813-817

***** Hassell, M.P., Comins, H. and May, R.M. 1991 Spatial Structure and Chaos in Insect Population Dynamics. *Nature* **353** 255-258

Hanski, I. 1998 Metapopulation dynamics *Nature* **396** 41-49

• **Control** Lecture 25-27

In this section we focus on human interventions and, in particular, how diseases can be effectively and efficiently controlled or even eradicated. This is a very applied subject area and we will consider several of the most popular methods of disease control: vaccination, quarantining, contract-tracing, ring-vaccination and ring-culling. These will be addressed in terms of efficiency and logistics, with reference to real situations. We will also consider methods to prevent the evolution of the multi-resistant "super-bugs".

Müller, J., Kretzschmar, M. and Dietz, K. 2000 Contact tracing in stochastic and deterministic epidemic models. *Math. Biosci.* **164** 39-64

Eames, K.T.D. and Keeling, M.J. 2003 Contact tracing and disease control. *Proc. Roy. Soc. Lond. B* **270** 2554-2560

Ferguson, N.M., Keeling, M.J., Edmunds, W.J., Gani, R., Grenfell, B.T., Anderson, R.M. and Leach, S. 2003 Planning for smallpox outbreaks *Nature* **425** 681-685

Matthews, L., Haydon, D.T., Shaw, D.J., Chase-Topping, M.E., Keeling, M.J. and Woolhouse, M.E.J. (2003) Neighbourhood control policies and the spread of infectious diseases *Proc. Roy. Soc. Lond. B* **270** 1659-1666

Müller, J., Schönfisch, B. and Kirkilionis, M 2000 Ring Vaccination *J. Math. Biol.* **41** 143-171

• **MacroParasites** Lecture 28

All infectious diseases that have so far been considered are *micro*parasites such as viruses and bacteria. Another class of parasites also exist, the *macro*parasites. As the name suggests these are far larger and consequently have more complex dynamics. In particular, the burden (of number) of such parasites per host plays a vital role.

• **Biological Data Sources** Lectures 29-30

While theoretical models give us an excellent understanding of the expected behaviour of ecological and epidemiological systems, good data sources are needed if our models are to be accurately parameterised and comparisons to observations made. Long temporal data sources are somewhat rare, due to the time and expense of collecting them. However, a few notable examples exist and we will investigate the modelling work that is based on them.

Hudson, P.J., Dobson, A.P. and Newborn, D. 1998 Prevention of population cycles by parasite removal

*Science* **282** 2256-2258
Bjørnstad, O.N. and Grenfell, B.T. 2001 Noisy clockwork: time series analysis of population fluctuations in animals. *Science* **293** 638-643
Turchin, P., Oksanen, L., Ekerholm, P., and Henttonen, H. 2000 Are lemmings prey or predators? *Nature* **405** 562-565.
Blasius, B., Huppert, A. and Stone, L. 1999 Complex dynamics and phase synchronization in spatially extended ecological systems *Nature* **399** 354-359
May, R.M. 1998 Population biology - The voles of Hokkaido *Nature* **396** 409-410
Ranta, E. and Kaitala, V. 1997 Travelling waves in vole population dynamics *Nature* **390** 456
Ranta, E., Kaitala, V., and Lundberg, P. 1997 The spatial dimension in population fluctuations *Science* **278** 1621-1623

## FINDING THE RIGHT MODEL.

First – "*there are no right model, but there are certainly lots of wrong ones.*". Picking the right model for the job is a trade-off between simplicity, accuracy and generality. An inaccurate model is no good to anyone, we need a model that approximates what's going on in the real world. A complex model might be more accurate (in general), but it may be too complex to understand or parameterise. Finally, we need a model that is general enough that it can be adapted to suit our purposes. We can broadly classify models into three groups.

- **Simulations: complex but accurate** Such models include all the gory details. They attempt to describe every process and hence paint as accurate a picture as possible of the real system. The prime example of a simulation is the large computer models used for weather prediction - an example from the ecological literature is the forest competition model SORTIE and the ecosystem model ATLSS. Another large scale simulation comes from general circulation models for ocean productivity, here meteorological data and ocean flows are combined with plankton models to create detailed spatial predictions. Simulations such as these should only be used when an exact answer is required as they usually provide little intuitive understanding of the problem. However they may be used to perform 'experiments' which could not be attempted on the natural system.

- **Models: simple and general** Probably the most common way of analysing a biological system; they attempt to capture the main features and are parameterised with a specific problem in mind. In general they do not provide analytical solutions, but many of the standard mathematical techniques can be applied numerically.

- **Caricatures: simple** These types of model are highly generic (so the results hold for many biological systems) and they often give analytically tractable results. Caricatures are most commonly used to gain an intuitive understanding of an additional feature which is not commonly included, such as spatial heterogeneity, genetic variability or stochasticity.

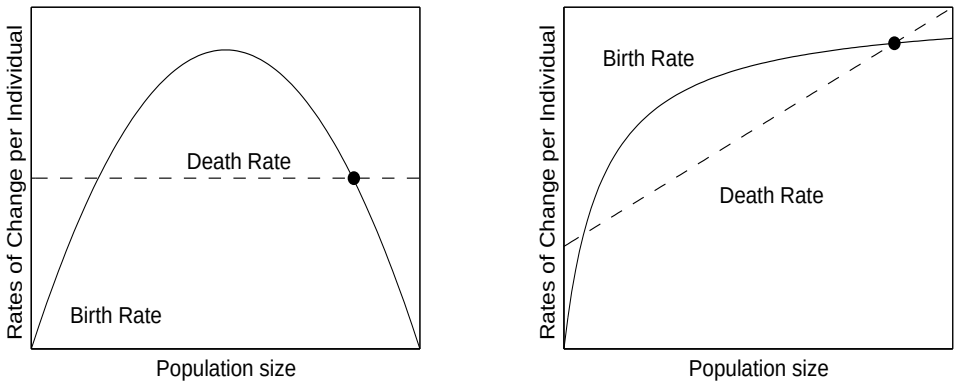
Most of the early quantitative work on population levels was made using continuous time differential equations - the solution or properties of which can be found using standard techniques from dynamical systems. The simplest form of model considers the behaviour of a single species whose population level $x$ is regulated by **intra-specific** competition.

**Intra-specific Competition**.
The simplest model which is commonly used is **logistic growth**,

$$\frac{dx}{dt} \;=\; f(x) \;=\; rx\left(1-\frac{x}{K}\right)$$

where $r$ is the **basic reproductive rate** and $K$ is the **carrying capacity**. This equation can either be used to describe the growth of a population or the growth of a single organism and leads to a sigmoidal growth curve. The quadratic term leads to **density dependent** behaviour.

In reality the change in a population is composed of a reproductive rate $b(x)$ and a death rate $d(x)$ per individual. Hence, $f = x[b(x) - d(x)]$ is the expected rate of change to the population from this birth-death process (Kretzschmar 1989); therefore the population level is described by a Markov chain with non-uniform probabilities. If $b$ and $d$ are constants then the Markov chain is a Yule-Furry process, which has zero as an attracting fixed point.




The *Allee effect* is when the non-trivial fixed point is NOT globally attracting. The graphs show two ways in which this can be achieved

For many species $b$ has been found to be unimodal, such that reproduction is low whenever the population is at extremes; in contrast the death rate $d$ is often independent of the population density $x$ (left-hand graph). Combining these two rates of change,

$$\frac{dx}{dt} \;=\; [b(x) - d]x$$

For such equations we can encounter the **Allee effect** (Courchamp *et al* 1999, Stephens and Sutherland 1999), this occurs when there exists a stable fixed point at $x^* > 0$, but $d(0) > b(0)$ such that $x = 0$ is also stable. Such populations may be stable at non-zero population levels for a long time, but are unable to recover once the population density drops below a threshold.

The allele effect is can also be seen when we consider **harvesting** (or hunting) of the population. If we take the standard logistic model, and assume that individuals are 'harvested' at a constant rate,

$$\frac{dx}{dt} \;=\; rx\left(1-\frac{x}{K}\right) \;-\; H\,.$$

From this equation we see that there are two fixed points for all $0 < H < \frac{rK}{4}$ and that as the harvesting increases, so does the risk of extinction due to external perturbations. In general we wish to have the **maximum sustainable yield** which occurs at $H = \frac{rk}{4}$. Note that an alternative assumption is hunting with a constant effort,

$$\frac{dx}{dt} = rx\left(1 - \frac{x}{K}\right) - Hx$$

does not lead to the allele effect. For this case we wish to find the **maximum economic yield**, maximising the harvest for the minimum amount of effort. In most real situations, we observe a mixture of these two scenarios.

To fully understand and predict trends in hunting/harvesting we need to combine ecological models with economic theory and sociological understanding. For some slowly reproducing species it may be the best economic policy to hunt the species to extinction (and place the profits in some other investment) than to harvest the species in a sustainable manner. Another problem is in enforcing restrictive hunting practices, this is the so-called *tragedy of the commons*, where resources accessible to all will be over exploited.

More complex dynamics can obviously be found. The Spruce budworm from Canada is thought to be described by,

$$\frac{dx}{dt} = rx\left(1 - \frac{x}{K}\right) - \frac{x^2}{1+x^2}$$

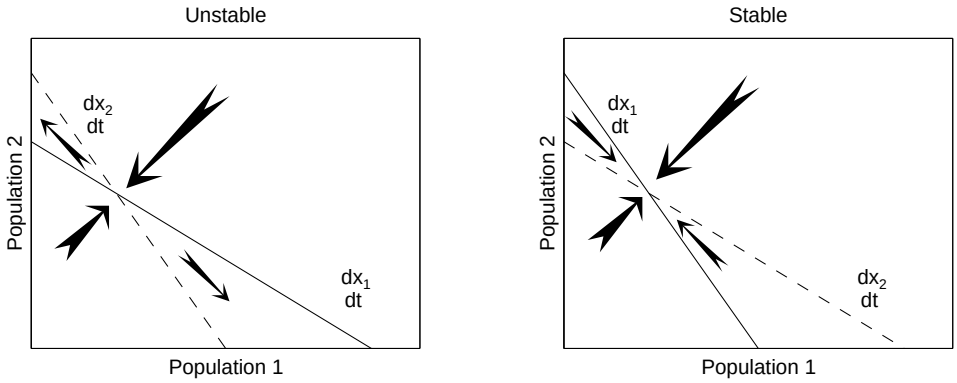
that is logistic growth together with type 3 predation (see functional responses). Changes in $r$ can change the number of fixed points (cusp catastrophe), which can cause dramatic jumps in the population. (See Murray 1989).

Of course, as these are one-dimensional continuous time systems, the dynamics are very simple, usually with a single globally attracting fixed point. (Note: we only consider *real positive* values of $x$ as we are dealing with populations.) The next step is to include another species and let the populations be limited by **inter-specific** competition or predation.

**Inter-specific Competition**.
We can now modify the logistic equation to model two species competing for some limiting resource - this could be food, space, nest sites or even avoiding the effects of predation - the **niches** of the two species overlap.

$$\begin{aligned}\frac{dx_1}{dt} &= r_1 x_1\left(1 - \frac{x_1 + \alpha_{12}x_2}{K_1}\right)\\ \frac{dx_2}{dt} &= r_2 x_2\left(1 - \frac{x_2 + \alpha_{21}x_1}{K_2}\right)\end{aligned}$$

If the two species are equal in every respect other than their competitive ability ($K_1 = K_2$ and $r_1 = r_2$) then we predict that only one species will survive - this is known as the **Competitive Exclusion Principle**. It was the ecologist Garrett Hardin who coined the maxim "Complete Competitors Cannot Coexist" (CCCC). [The above equations and therefore the competitive exclusion principle rely on the species being homogeneously distributed, if the species aggregate then the weaker species can still persist (Hanski 1981). Forms of heterogeneity may explain the 'paradox of plankton' – a great number of species with similar niches coexist.] When there is a trade off between competitive advantage $\alpha$ and carrying capacity $K$ the two competing species can coexist. This is frequently explained by considering the null clines of the equations, or we can look at the Jacobian.




For stability we require a trade-off between competitive ability and carrying capacity.

Hence coexistence can only occur when

$$\alpha_{12}K_2 < K_1 \qquad \text{and} \qquad \alpha_{21}K_1 < K_2$$

Note that the values of $r_1$ and $r_2$ do not play a role in determining the stability of the system. When the coexistence conditions are reversed,

$$\alpha_{12}K_2 > K_1 \qquad \text{and} \qquad \alpha_{21}K_1 > K_2$$

the fixed point is unstable and we observe **founder control**, where the initial density of the two species determines the final outcome. This type of behaviour is often seen in plants when, due to shading, it is difficult for a new species to invade.

**Predation**

Predator-prey models date back to 1925 when they were derived separately by Lotka and Volterra. For the predator $P$ and the victim $V$ the Lotka-Volterra equations are

$$\begin{aligned} \frac{dV}{dt} &= BV - CVP \\ \frac{dP}{dt} &= -DP + AVP \end{aligned}$$

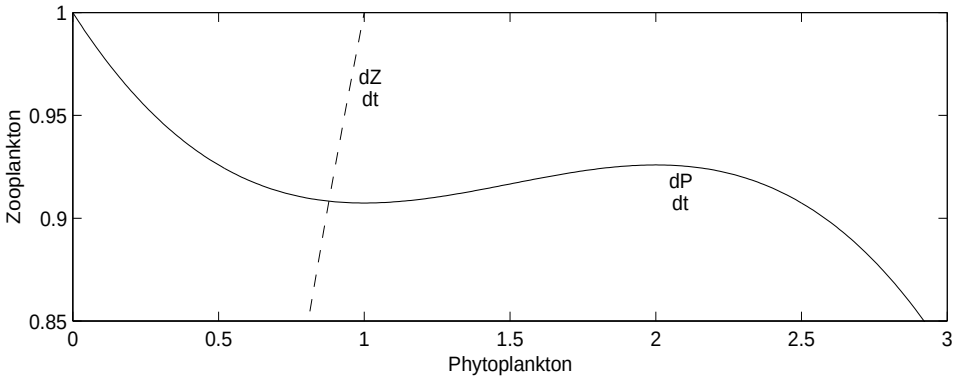
Such a generic model can be used to describe the behaviour of predator-prey, plant-herbivore, host-parasite or host-parasitoid species. Where does this set of equations come from? We want to think about $V$ and $P$ as being the *densities* of the two species within a given region. In the absence of interactions the prey are born at rate $B$ and the predators die at rate $D$. The interactions take the form of a **mass action** term $V \times P$, so as the density of predators and prey increases so does the number of interactions. This views organisms like the molecules of a gas, bouncing around randomly and colliding with each other. In truth ecology isn't this simple, but its a reasonable first approximation.

This system of equations has a fixed point ($V^* = D/A$, $P^* = B/C$) which is neutrally stable, so the dynamics is a closed orbit around the fixed point defined by the initial conditions.

Due to the simple non-linearities in the equations, the prey null cline ($\frac{dV}{dt} = 0$) is parallel to the $V$axis - it is this that leads to the **paradox of enrichment**. This is the name given to the phenomena that if conditions improve for the prey (so that $B$ is increased) this is reflected in an increase in the predators, while the prey density is unchanged.

More complex interactions than the mass action assumption are often incorporated into ecological models. For example the interaction between zooplankton (predators) and phytoplankton (prey) can be better described by a rational function which together with slower dynamics for the zooplankton can lead to excitable behaviour.

$$\begin{aligned} \frac{dP}{dt} &= 9P - 2P^2 - 54\frac{ZP}{6+P^2} \\ \frac{dZ}{dt} &= -\varepsilon Z^2 + 7\varepsilon\frac{ZP}{6+P^2} \end{aligned}$$




The null clines from the above plankton model. Due to the non-monotonic behaviour of the $P$ null cline the system can be excited by a small perturbation.

More complex equations, including age structure, functional responses, aggregation of individuals and time delays will be discussed later. However some other complications/ideas will now be mentioned briefly.

**Delay Models**
Very complex behaviour can be observed from even single species models with delays, eg.

$$\frac{dx}{dt} = rx(t)\left(1 - \frac{x(t-T)}{K}\right)$$

which is logistic growth but with a lag in the carrying capacity term. These delays are usually seen as some representation of the effects of either another species or age-structure. Equations with delays can be very hard to analyse or numerically simulate.

**Nutrient Models**
Rather than modelling the number of individuals (or density) for each species, some researchers consider the dynamics in terms of the flow of nutrients/resources through the ecosystem. These models generally consider the flow of carbon, nitrogen or energy through the system and are frequently coupled to complex environmental models of the atmosphere or oceans.

**Multi-species food webs**
With just two interacting species, we can only get fixed points and limit-cycles. As the number of species increases so does the likelihood of finding more complex behaviour and chaotic dynamics. Despite the vast number of interacting species in any one environment, chaotic dynamics seems to be rare - or simply swamped by noise. In general, applicable models have tended to ignore more than three interacting species, due to the difficulties with parameter estimation and

fully exploring the entire range of dynamical behaviour.

---

Begon, M., Mortimer, M. and Thompson, D.J. 1996 *Population Ecology: A Unified Study of Animals and Plants* Blackwell Science.

Courchamp, F., Clutton-Brock, T. and Grenfell, B. 1999 Inverse density dependence and the Allee effect *T.R.E.E.* **14** 405-410

Gillman, M. and Hails, R. 1997 *An Introduction to Ecological Modelling* Blackwell Science.

Hanski, I. 1981 Coexistence of competitors in patchy environments with and without predation. *Oikos* **37** 306-312

Kretzschmar, M. 1989 A Renewal Equation with a Birth Death Process as a Model for Parasitic Infections. *Journal of Mathematical Biology* **27** 191-221

Lotka, A.J. 1925 *The Elements of Physical Biology.* Williams and Williams Co., Baltimore.

Stephens, P.A. and Sutherland W.J. 1999 Consequences of the Allee effect for behaviour, ecology and conservation *T.R.E.E.* **14** 401-405

Volterra, V. 1926 Fluctuations in the Abundance of a Species Considered Mathematically. *Nature* **118** 558-560

**The Notation and Classification**

Most diseases we shall be considering are caused by either viruses or bacteria. However, we shall ignore the dynamics within the human body (the interaction between diseases and the immune system is a growing field that is yet to be fully explored). Instead we classify individuals according to their status with respect to the disease,

• *Maternal Immunity* For around 6-12 months after birth, new-born infants may be protected by maternal immunity. For modelling purposes this is often ignored, individuals are only assumed to by "born" once maternal immunity has waned.

• *Susceptible* Individuals in this class can catch the disease if they are exposed to it. The number of susceptible individuals is usually labelled $S$.

• *Exposed* This class, usually labelled $E$, covers those people that have caught the disease, but that are not yet infectious. They are incubating the disease, with the number of viral particles or bacteria increasing rapidly.

• *Infectious* Infectious individuals, $I$, can spread the disease to any susceptibles that they come into contact with. (Note that *Infected* individuals are those that are either exposed or infectious)

• *Recovered* These people have recovered from the disease. Often (as is the case for measles) they maintain a life-long immunity after infection, although for other diseases with multiple strains (such as influenza) this is more complex. Not surprisingly these are labelled $R$.

Not all disease models will include all of these classes, and some models will include more. For example, there may be a *vaccinated* class, there may be waning immunity such that those individuals who have recovered steadily decay to a *partially-susceptible* class. We may also wish to sub-divide the classes further, such that the population is age-structured or sexually-structured. In the next few lectures we shall consider in detail a variety of models which use these classes in different combinations.

**The Basic Reproductive Ratio $R_0$**

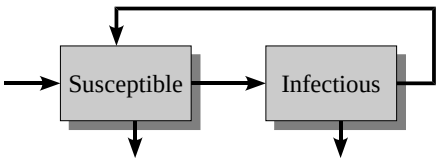
This is without doubt the most important quantity in the whole of epidemiology. It is defined as ....

**The *average* number of secondary cases produced by an *average* infectious individual in a totally susceptible population.**

As such it tells us whether a disease can invade a naive population. In practice we calculate $R_0$ either from first principles or we look at the dominant eigenvalue of the disease free equilibrium. If $R_0 > 1$, then each infectious individual creates more than one new case and hence the disease spreads. When $R_0 < 1$ the disease dies out. Thus there is a relationship between $R_0$ and the stability of the disease-free state.

The following table shows the values of $R_0$ for a range of well-known diseases. Note the wide range of values, and the fact that $R_0$ depends on both the disease and the environment.

| Disease | $R_0$ (approx) |
|---|---|
| Small pox | 4 |
| Measles | 17 |
| Chicken pox | 11 |
| HIV (male homosexuals in England and Wales) | 4 |
| HIV (female prostitutes in Kenya) | 11 |
| Malaria | $\approx 100$ |




The SIS model, showing the routes between the compartments.

The Susceptible-Infectious-Susceptible model is by far the simplest of all epidemiological models, but it serves to illustrate the main features of all disease models. It is commonly used in the study of sexually transmitted disease, where due to the vast number of strains there is no such thing as a recovered class. The dynamics of this model are described by,

$$\begin{aligned} \frac{dS}{dt} &= BN - \beta SI/N + gI - dS \\ \frac{dI}{dt} &= \beta SI/N - gI - dI \\ & \text{where } N = S + I \end{aligned}$$

Where $B$ is the birth rate per individual, $d$ is the natural death rate, $g$ is the rate that infection is 'lost' or the recovery rate, and $\beta$ is the so-called transmission or contact rate. We note that the term $-gI$ means that infectious individuals decay back to being susceptible (*cf* radioactive decay and half-lives). The rate at which susceptibles catch the disease is termed the **force of infection**, $\lambda$,

$$\lambda = \beta I/N$$

We note that the interaction between susceptibles $S$ and infectious $I$, is divided by the total population size $N$. This is termed **pseudo mass-action**. True mass-action, which comes from thinking about randomly moving particles, would not be divided by $N$. For true mass-action we would be envisaging a situation where as the population increased the individuals became more and more tightly packed and hence interacted more often. For most human populations, the density of individuals is fairly independent of population size – in fact it is the number of social contacts that is the determining factor. It appears that **pseudo mass-action** is the best description of human diseases – the evidence comes from the fact that the same parameters can be used to produce accurate models of measles and other diseases across a wide range of population sizes.

We can make this model even simpler by restricting our attention to a closed population without births or deaths ($B = d = 0$).

$$\begin{aligned} \frac{dS}{dt} &= -\beta SI + gI \\ \frac{dI}{dt} &= \beta SI - gI \end{aligned}$$

where $S$ and $I$ are now the proportion of susceptible and infectious individuals respectively. From this most basic of models we can calculate $R_0$.

$$\begin{aligned} R_0 &= \text{(rate at which secondary cases are produced)} \times \text{(average infectious period)} \\ &= (\beta S) \times (1/g) = \frac{\beta}{g} \end{aligned}$$

So $R_0$ is increased by a long infectious period or a high rate of transmission.

Similarly, if we look at the disease-free equilibrium ($S = 1$, $I = 0$), and examine the stability,

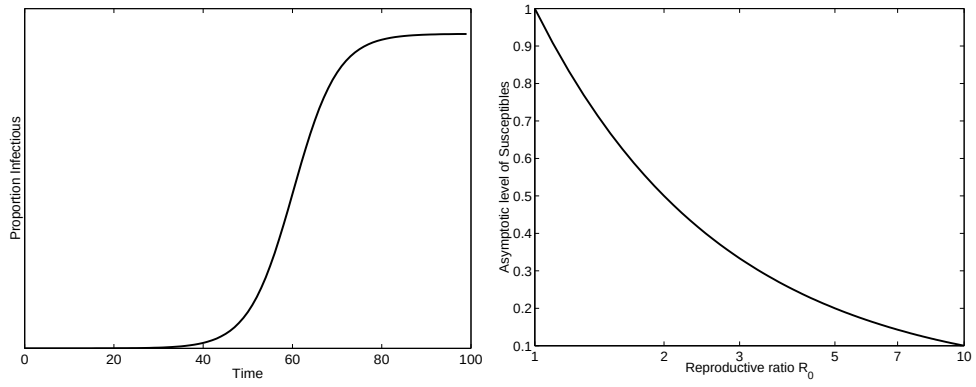
$$\frac{dI}{dt} \;=\; \beta I \;-\; gI$$

Thus the dominant eigenvalue is $\Lambda = \beta - g$, and $R_0$ is given by $R_0 = \frac{\Lambda/g}{+}1$. Clearly the disease-free state is unstable if $R_0 > 1$.

We now want to look in more detail at the behaviour close to the fixed points. First we notice that we can simplify the dynamics by setting $S = 1 - I$, we then get the logistic growth model,

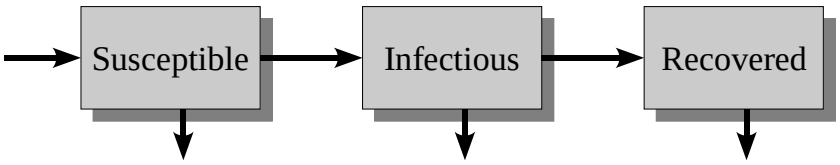
$$\frac{dI}{dt} \;=\; \beta(1-I)I \;-\; gI \;=\; [\beta - g]I\left(1 \;-\; \frac{I\beta}{\beta - g}\right)$$

From this it should be obvious that the fixed points are $I^* = 0$ and $I^* = 1 - \frac{g}{\beta}$. Alternatively this could be written as $S^* = 1$ and $S^* = \frac{1}{R_0}$. This latter relationship is a classic result that we will show is generally true for a wide variety of models, **the non-trivial level of susceptibles is the inverse of $R_0$.**




The left-hand graph shows the typical dynamics of an SIS model ($\beta = 1.2$, $g = 1$), the right-hand graph shows how $S^*$ varies with $R_0$.

## The S-I-R Model




The SIR model, showing the routes between the compartments.

The simple SIS model could be reduced to one-dimension (as $S$ and $I$ are related), as such its dynamics are trivial and cannot capture the behaviour of many real human diseases - although it can give a reasonable description of some STDs. We now study the SIR model – adding a

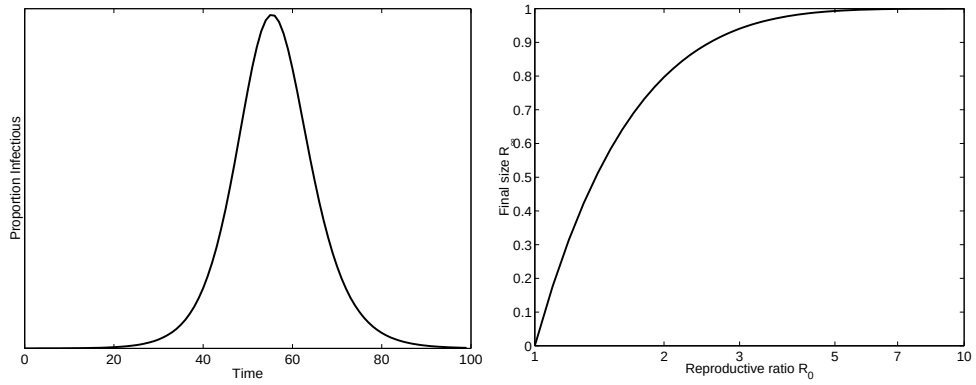
recovered class which cannot catch the disease.

**The Simple Epidemic**

Initially, we shall ignore all demography (births and deaths) by assuming that the progress of the disease is much more rapid than the natural birth and death rate. This model is called **the Simple Epidemic** and may provide a good description of a 'one-shot' disease such as influenza.

$$\begin{aligned} \frac{dS}{dt} &= -\beta SI \\ \frac{dI}{dt} &= \beta SI - gI \\ \frac{dR}{dt} &= gI \end{aligned}$$

The initial condition for this system is $S(0) = 1 - \varepsilon$, $I(0) = \varepsilon$, $R(0) = 0$, with $\varepsilon \ll 1$. Hence we are again dealing with proportions of the population, and we have assumed that all individuals are initially susceptible and then introduce a very small quantity of infection. We can find a value for $R_0$, which again is $\beta/g$. We note that there is still a simple relationship between $S$ and $R_0$; when $S > 1/R_0$ the disease increases whereas when $S < 1/R_0$ the disease decreases.We note that the only fixed-point of the system is when there is no disease present $I^* = 0$.




The left-hand graph shows the typical dynamics of a simple-epidemic SIR model ($\beta = 1.2$, $g = 1$), the right-hand graph shows how the final size of the epidemic varies with $R_0$.

For this simple-epidemic, the more interesting question is the final state of the system. It is clear that without births to replenish the number of susceptible individuals, the level of susceptibles must decrease through time. This in turn leads to a decrease in the amount of infection - the question is does the disease die out before all the susceptibles are exhausted. We now wish to determine the **final size** of the epidemic ($R_\infty$),

$$R_\infty \;=\; R(\infty) \;=\; 1 - S(\infty) \;=\; \text{total proportion infected}$$

We calculate this value using an important trick developed by Kermack and McKendrick (1927):

$$\frac{dS}{dR} \;=\; -R_0 S \qquad \Rightarrow \qquad S \;=\; \exp(-R_0 R).$$

We now use the fact that all the elements sum to one,

$$S + I + R \;=\; \exp(-R_0 R) + I + R \;=\; 1$$

The equilibrium solutions occur when $I = 0$, substituting this in gives:

$$R_\infty \;=\; 1 - \exp\left(-R_0 R_\infty\right).$$

Although this has no analytical solution we can easily calculate its value numerically, in which case we find that $R_\infty$ rapidly approaches 1 as $R_0$ becomes significantly larger than one.

**The full SIR model**

In reality there is no single model that can be called the SIR model - many different slight variations exist. The version given below is one of the most used, and displays all the important features common to this family of models.

$$\begin{aligned}
\frac{dS}{dt} &= B - \beta SI/N - dS \\
\frac{dI}{dt} &= \beta SI/N - gI - dI \\
\frac{dR}{dt} &= gI - dR \\
N &= S + I + R
\end{aligned}$$

In general this is a three-dimensional system, but to simplify matters we can set $B = dN$ such that the population size remains constant, and then rescale so that we are again dealing with proportions. We now look at the fundamental parameters and dynamics of this model.

Let us again look at the value of $R_0$, by comparison with the Simple Epidemic we would expect to have $R_0 = \beta/(g + d)$; we shall check this from the eigenvalue behaviour,

$$J = \begin{pmatrix} -\beta I - d & -\beta S \\ \beta I & \beta S - g - d \end{pmatrix} = \begin{pmatrix} -d & -\beta \\ 0 & \beta - g - d \end{pmatrix}$$

So the eigenvalues are $\Lambda = -d$ and $\Lambda = \beta - g - d$, taking the second value (which is almost always larger) we see that $R_0 = \Lambda \times (\text{Infectious Period}) + 1 = \beta/(g + d)$.

For the full SIR model, with births replenishing the level of susceptibles, there is a second fixed point in which the disease is present. From setting $\frac{dI}{dt} = 0$ we get,

$$S^* \;=\; \frac{g+d}{\beta} \;=\; \frac{1}{R_0}$$

which shouldn't be too surprising by now and comes from the linear interaction of susceptibles and infectious cases. And from $\frac{dS}{dt} = 0$,

$$I^* \;=\; \frac{B}{g+d} - \frac{d}{\beta}$$

It should be clear that the number of infectious individuals is limited by the birth rate – clearly it is impossible to keep infecting more people than are being born.

Looking at the Jacobian for this fixed point provides a very important insight into the dynamics of real diseases,

$$J = \begin{pmatrix} -\beta I - d & -\beta S \\ \beta I & \beta S - g - d \end{pmatrix} = \begin{pmatrix} -\frac{\beta B}{g+d} & -g - d \\ \frac{\beta B}{g+d} - d & 0 \end{pmatrix}$$

Therefore the eigenvalues are,

$$\lambda \;=\; \frac{-\beta B \pm \sqrt{\beta^2 B^2 + 4d(g+d)^3 - 4\beta B(g+d)^2}}{2(g+d)}$$

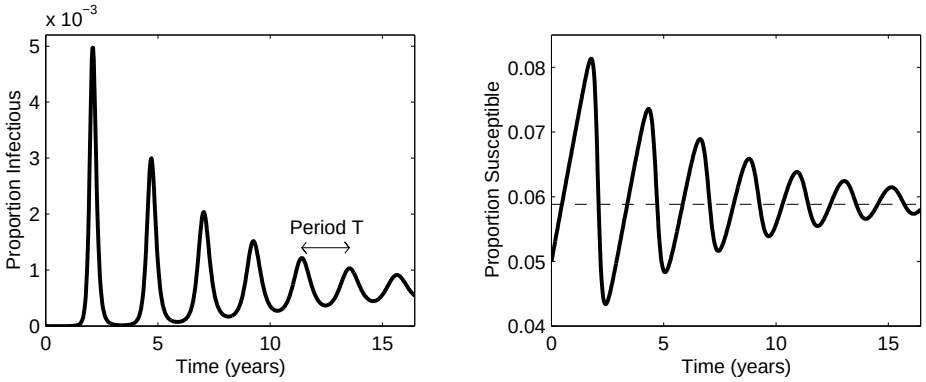
and making the assumption that $d = B$ is small compared to the other terms,

$$\lambda \;=\; \frac{-\beta d \pm 2i\sqrt{dg^2(\beta - g)}}{2g} \;=\; -\frac{dR_0}{2} \pm i\sqrt{d(\beta - g)}$$

So it is clear that if $R_0$ is greater than one then the diseased fixed point will be stable, and we converge to this point with damped oscillations.

Let us consider the period of these oscillations in more detail. The period $T$ is given by:

$$\begin{aligned} T \;=\; \frac{2\pi}{\sqrt{d(\beta - g)}} \;&=\; 2\pi\sqrt{\frac{\text{Life expectancy} \times \text{Infectious Period}}{R_0 - 1}} \\ &=\; 2\pi\sqrt{\text{Average age of infection} \times \text{Infectious Period}} \end{aligned}$$




The dynamics of the full SIR model using the parameters for measles in England and Wales ($B = d = 5.5 \times 10^{-5}$ per day, $R_0 = 17$, $g^{-1} = 13$ days). Note the period of inter-epidemic oscillations, which is longer for large epidemics due to non-linearities far from the fixed point. The susceptibles decline during an epidemic, but are replaced by births in the troughs.

We note that this approximation only holds when the oscillations are small. For larger oscillations, the period between epidemics is longer.

**Vaccination**

Successful vaccination moves individuals straight into the recovered class, so that they can no-longer catch or spread the infection. We can therefore introduce a model in which individuals are vaccinated at a rate $V$ (note that the vaccination of infected or recovered individuals is assumed to have no effect).

$$\begin{aligned} \frac{dS}{dt} &= d - \beta SI - dS - VS \\ \frac{dI}{dt} &= \beta SI - gI - dI \\ \frac{dR}{dt} &= VS + gI - dR \end{aligned}$$

Let us consider the equilibria of this system in some detail. First the endemic equilibria ($I^* > 0$), we have $S* = 1/R_0$ (surprise surprise) and $I* = \frac{d}{g+d} - \frac{d}{\beta} - \frac{V}{\beta}$. However, if the level of vaccination is sufficiently high $V > d(R_0 - 1)$, the level of infection is negative – intuitively the disease cannot survive. We shall now consider the disease free equilibrium, $S^* = \frac{d}{d+V}$, so as expected the proportion of susceptibles decreases as we vaccinate more and more. Looking at the Jacobian,

$$J = \begin{pmatrix} -\beta I - d - V & -\beta S \\ \beta I & \beta S - g - d \end{pmatrix} = \begin{pmatrix} -d - V & -\beta S^* \\ 0 & \beta S^* - g - d \end{pmatrix}$$

So the dominant eigenvalue is $\Lambda = \beta S^* - g - d$ and this is less than zero (a disease can't invade) if $V > d(R_0 - 1)$ or $S* < 1/R_0$.This leads us to the critical vaccination threshold – the proportion of the population that needs to be vaccinated if a disease is to be eradicated and prevented from returning.

$$V_C \;=\; 1 - \frac{1}{R_0}$$

Thus diseases with a low $R_0$, such as small-pox ($R_0 \approx 5 \;\Rightarrow\; V_C \approx 80\%$), can be easily erradicated by vaccination, whereas diseases with high $R_0$, such as measles ($R_0 \approx 17 \;\Rightarrow\; V_C \approx 94\%$) and malaria ($R_0 \approx 100 \;\Rightarrow\; V_C \approx 99\%$) are much more difficult.

**DISCRETE TIME MODELS** In the last sections we considered the dynamics of simple one and two species systems in continuous time. However, many animals and plants (especially in temperature regions) reproduce on an annual basis. To capture this discrete behaviour we introduce a map $f$ which takes us from the population level this year to the population next year, $x_{t+1} = f(x_t)$. In truth $f$ is acting as a Poincare return map for the continuum dynamics of a periodically forced system. The most widely know and used equation is the **logistic map**,

$$f(x) = r\,x\left(1 - \frac{x}{k}\right)$$

where the carrying capacity $k$ acts to scale $x$, and the basic reproductive rate $r$ is a bifurcation parameter. When $r < 1$ the population is doomed to extinction, as $x = 0$ is a globally attracting fixed point. As $r$ is increased we observe the usual period doubling route to chaos that is expected for a logistic map. It is interesting to note the differences between the logistic map and the logistic equation.

[ **Aside**: since such a simple (and possibly even biologically-realistic) map was shown to produce chaotic dynamics (May 1976) many researcher have looked for chaos in laboratory and natural populations. We would think that chaos should be common place - if such 1-D maps can produce complex dynamics, then surely the complexities of the real world should make chaos even more likely; secondly chaos occurs for larger values of $r$ and would would intuitively feel that a faster reproductive rate should be favoured by evolution. However, chaos is rarely, if ever detected in natural populations. It could be that it is present but is swamped by environmental noise and internal stochasticity; or it could be that the frequent excursions to low densities means that chaotic dynamics are not selected for by evolution. ]

Let us consider in more detail what we want from our map $f$. We require $f : \mathbb{R}^+ \mapsto \mathbb{R}^+$, that is $f$ takes positive real values to positive real values, so that any feasible population level leads to another feasible population. We note that the logistic map does not satisfy this criterion, as values of $x$ greater than $k$ are mapped to negative (and therefore impossible) population levels. An alternative form is,

$$f(x) = rx\exp\left(1 - \frac{x}{k}\right)$$

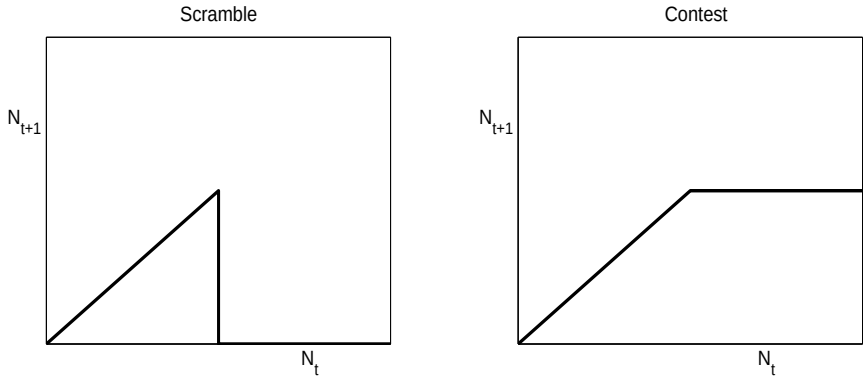
this is the **Ricker map**. It has the same period doubling properties as the logistic map (because it possesses a quadratic maximum) but is well behaved on the positive real line.

Where does the form for the map come from? As mentioned earlier these maps are Poincare return maps, which sample a continuous system periodically. Hence to understand the map we must understand the continuous system, and in particular study the effects of overcrowding (intra-specific competition) that lead to the non-linearity of the map. Two extremes of intraspecific interaction were classified by Nicholson; Scramble and Contest.

**Scramble** In this scenario all individuals scramble to get a proportion of a limited resource. All individuals get an equal share, and once this share drops below a critical threshold all the individuals die from lack of resource. We can consider this to be a good model for a large food source that is produced over a short period; when the food is abundant all individuals can eat their fill, but as soon as all the food is used up all the animals suffer from a lack of food.

**Contest** At the opposite extreme, imagine a situation where the fittest individuals take what share of the resource they require. Thus even when there is lots of competition for the resource the fit individuals still survive (Grenfell *et al* 1998). This model is used for when there is a continual supply of a limited food resource, so that the fittest individuals always eat their fill

first.




The iterative maps due to **scramble** and **contest** for a limited resource. Notice that the scramble map can never be stable.

Almost all real situations lie somewhere between these two extremes. Hassell in 1975 introduced a map, which has proved popular with ecologists ever since, that attempts to scale between these two extremes with a parameter $b$,

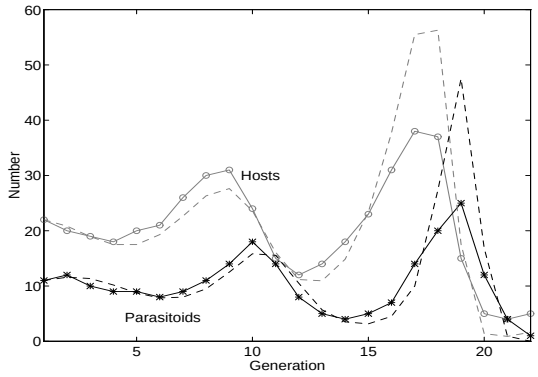
$$f(x) \;=\; \frac{rx}{(1+kx)^b}$$

This behaves more and more like scramble competition as $b \to \infty$ (with the other parameters suitably rescaled). Once again, for $b$ sufficiently large, we find that there is a period doubling route to chaos as $r$ is increased.

**The Nicholson Bailey Model**

For the discrete time maps (as for the continuous system) we can also consider multi-trophic models, the most celebrated being the host-parasitoid model of Nicholson and Bailey (1935).

$$\begin{aligned} H_{t+1} &= rH_t \exp(-aP_t) \\ P_{t+1} &= \lambda r H_t \,[1 - \exp(-aP_t)] \end{aligned}$$




Comparison between laboratory populations (solid line) and the theoretical equations (dashed line) for whitefly parasitised by a chalcid.

To understand this map we again need to consider the underlying behaviour between measurements. Let us start with $H_t$ adult hosts and $P_t$ adult parasitoids.

The adult hosts produce $Y(0) = rH_t$ young before dying, the adult parasitoids then search round and lay their eggs in any host young that they find. If parasitoids encounter host offspring randomly then the number of un-parasitised offspring $Y$ obey the differential equation

$$\frac{dY}{dT} \;=\; -A\,P_t Y(T)$$

Hence after time $\tau$ there are $Y(\tau) = rH_t \exp(-aP_t)$ young which mature into adult hosts, and the remaining $rH_t\,[1 - \exp(-aP_t)]$ parasitised young produce on average $\lambda$ parasitoids each ($a = A\tau$). Hence the standard Nicholson-Bailey model is the product of a short host reproduction phase followed by the random search for host offspring by parasitoids. Rather than study these two continuous time phases, we just look at the map which takes the populations from one year to the next.

This map possesses a single unstable fixed point, and trajectories rapidly spiral outwards with successively larger cycles of boom and bust - such that in any finite population either the parasitoids or both species will go extinct. Despite this obvious drawback the Nicholson-Bailey equations give a good approximation, in the short time, to the dynamics of laboratory populations. Obviously natural populations don't suffer these cycles leading to extinction, so several modifications have been suggested to add greater realism to the dynamics. [Note all these modifications can be made to the continuous time Lotka-Volterra models for predation.]

**Intraspecific Competition** Currently, in the absence of parasitoids the host population will grow geometrically - this is not realistic. Probably the simplest modification is to introduce intraspecific competition between hosts after the reproductive phase. Beddington *et al* (1975) proposed the following form for the number of host offspring that mature,

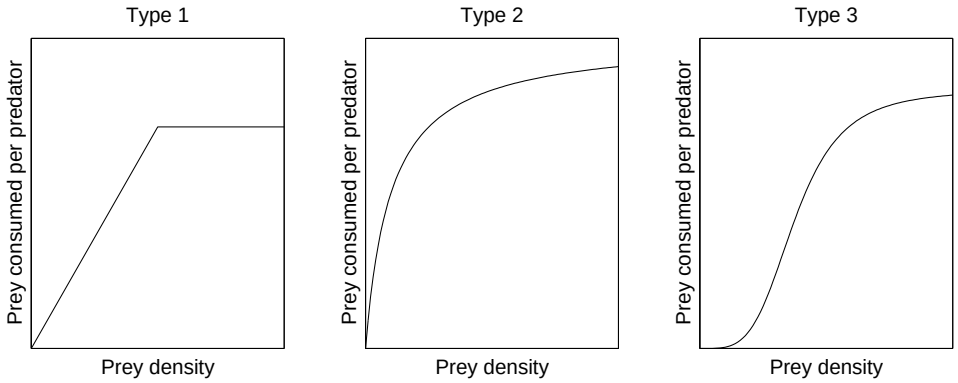
$$Y(\tau) \;=\; H_t \exp\left(r\left[1 - \frac{H_t}{K}\right] - aP_t\right)$$

This density-dependent growth of the hosts can be compared to the Ricker map.
**Non Mass-Action Searching** An alternative addition to the equations, which acts to stabilise the cycles is changes to the searching pattern. The parasitoids may suffer from intraspecific competition such that as the number of parasitoids increases so the rate of oviposition decreases, in which case the number of host offspring which mature can be given by

$$Y(\tau) \;=\; rH_t \exp(-aP_t^{1-m})$$

where $m$ measures to degree of **interference**. Interference has a stabilising effect on the dynamics. Note that there is no underlying analytical reason why this exponent is introduced, but the form of this model is backed-up by observations. An exponent is often added to mass-action terms as a means of approximating the effects of space or aggregation.
**Functional Responses**

It is unlikely that as the density of hosts increase, so the number that can be parasitised by a fixed number of parasitoids will grow linearly. A functional response curve tell us how the number of prey (or hosts) 'consumed' by a predator (or parasitoid) changes with the density of prey (or hosts). The asymptotic nature is because there is a handling time $T_h$ associated with the capture and consumption of the prey, which means that less of the time $T$ is available for searching. In general it is very difficult to determine the precise functional response curves from ecological data, due to stochastic effects and the fact that the entire range of prey densities are rarely sampled.




Three different functional forms, which illustrate how the number of prey consumed by a predator scales with prey density.

Holling (1959) classified 3 forms of functional response. **Type 1** is generally experienced by filter-feeds such as daphnia.
For **Type 2** responses, the initial rate change is linear, but eventually the consumption asymptotes to a fixed level. With Type 2 functional responses the number of host young which mature can be approximated by

$$Y(\tau) \;=\; rH_t \exp\left(\frac{-aP_t}{1 + AT_hH_t}\right)$$

This form actually makes the Nicholson-Bailey model LESS stable. Type 1 and 2 responses can also be thought of as being **Inverse Density Dependence** for the prey species, because as the number of prey increase each one stands less chance of being consumed. We can think of this functional response to be due to saturation of a predator's appetite - each predator can only consume a fixed amount of prey irrespective of how easy it is to catch.
**Type 3** responses are more complex again, and have a very low consumption level for low host densities,

$$Y(\tau) \;=\; rH_t \exp\left(\frac{-aH_tP_t}{1 + cH_t + bT_hH_t^2}\right)$$

This type of response can occur because of a predator's limited ability to find rare prey, or due to predators conserving energy when prey is scarce. Although this form is stabilising, it can never fully stabilise the fixed point.

**Heterogeneity** One final method of stabilising the system is to introduce heterogeneities, so that the hosts and parasitoids are not assumed to be uniformly distributed over all space. The original way of accomplishing this was to assume a predefined aggregation for each species, this was achieved by assuming a negative binomial form rather than a Poisson distribution,

$$Y(\tau) \;=\; rH_t \left(1 + \frac{aP_t}{k}\right)^{-k}$$

When $k$ is small the species are highly aggregated, but as $k \to \infty$ we regain the familiar exponential form.
In 1990 Hassell and Pacala developed the $CV^2$ rule. This is a general rule of thumb, that the system is stable if

$$CV^2 = \frac{\text{std. dev. of risk of parasitism per host}}{\text{mean risk of parasitism}} \;>\; 1$$

A more modern approach, which has triggered much interest, is to subdivide the population into a grid of habitats and to allow the heterogeneity to arise naturally (Hassell *et al* 1991).

Beddington, J.R., Free, C.A. and Lawson, J.H. 1975 Dynamic complexity in predator-prey models framed in difference equations. *Nature* **225** 58-60
Begon, M., Mortimer, M. and Thompson, D.J. 1996 *Population Ecology: A Unified Study of Animals and Plants* Blackwell Science.
Grenfell, B.T., Wilson, K., Finkenstädt, B.F., Coulson, T.N., Murray, S., Albon, S.D., Pemberton, J.M., Clutton-Brock, T.H. and Crawley, M.J. 1998 Noise and determinism in synchronized sheep dynamics *Nature* **394** 674-677
Hassell, M.P., 1975 Density dependence in single species populations *J. Animal Ecol.* **42** 693-726
Hassell, M.P., Comins, H. and May, R.M. 1991 Spatial Structure and Chaos in Insect Population Dynamics. *Nature* **353** 255-258
Hassell, M.P. and Pacala, S.W. 1990 Heterogeneity and the dynamics of host-parasitoid interactions. *Phil. Trans. Roy. Soc. Lond. B* **330** 203-220
May, R.M. 1976 Simple mathematical models with very complicated dynamics. *Nature* **261** 459-467

**FORCED DISEASES**

Let us consider a disease, such as influenza, which is spread more easily when everyone is packed inside in centrally heated buildings. It is plausible to assume that the transmission of such an infection varies with the outside temperature in a sinusoidal manner, $\beta = \beta_0 + \beta_1 \sin(\omega t)$. Other causes of such forcing might be climatic changes such as moisture (which affects biting insects). This leads to a modified version of the SIR equations,

$$\begin{aligned} \frac{dS}{dt} &= B - \beta(t)SI - dS \\ \frac{dI}{dt} &= \beta(t)SI - gI - dI \\ \frac{dR}{dt} &= gI - dR \end{aligned}$$

Let us consider the stability of such equations if $\beta$ undergoes small oscillations. Suppose that $S(t) = S_0 + S_1 e^{i\omega t}$ and $I(t) = I_0 + I_1 e^{i\omega t}$, then $S_0$ and $I_0$ obey similar equations to the standard (non-forced) model, whereas

$$\begin{aligned} \frac{dS_1}{dt} &= -\beta_1 S_0 I_0 - \beta_0 S_0 I_1 - \beta_0 I_0 S_1 - dS_1 - i\omega S_1 \\ \frac{dI_1}{dt} &= \beta_1 S_0 I_0 - \beta_0 S_0 I_1 - \beta_0 I_0 S_1 - dI_1 - gI_1 - i\omega I_1 \end{aligned}$$

Assuming that $d = B$ is small, this leads to solutions:

$$\begin{aligned} I_1^* &\approx -\frac{d\beta_1}{\omega^2 \beta_0^2}(\omega g + i\omega\beta_0) \\ S_1^* &\approx \frac{d\beta_1}{\omega^2 \beta_0^2}(\beta_0 g + g\omega - g^2 + i\omega\beta_0) \end{aligned}$$

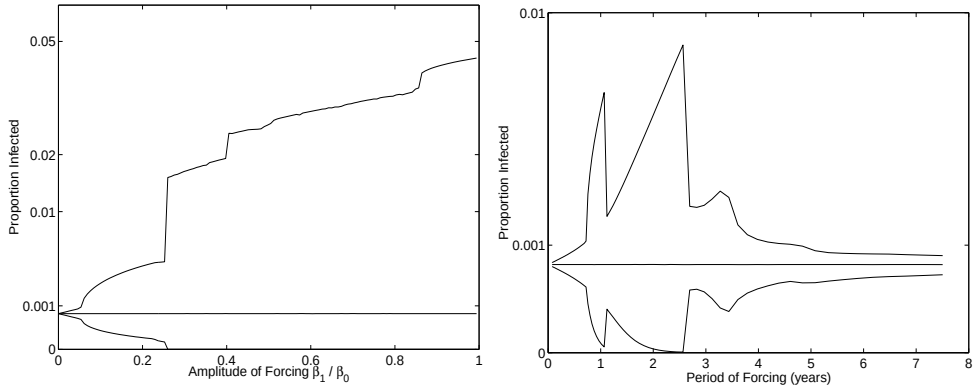
So small amounts of seasonal forcing lead to small oscillations of the same period, with the infection significantly lagging the forcing. The stability of these oscillations will be left as an exercise.

For wild-life or agricultural populations, it may be more appropriate to consider a pulsed birth rate.

$$\begin{aligned} \frac{dS}{dt} &= B(t) - \beta SI - dS \\ \frac{dI}{dt} &= \beta SI - gI - dI \\ \frac{dR}{dt} &= gI - dR \end{aligned}$$

where $B$ may be a periodic delta function. As such the susceptibles get replaced on a yearly basis, and between these pulses of births, the dynamics are that of the Simple Epidemic. Such a system has yet to be examined in detail.

The obvious next question is what happens as the level of seasonality gets bigger? We can then get interactions between the natural period of the epidemics and the seasonal forcing. Compare this to pushing a damped pendulum. The figure below shows the effects of varying the amplitude and frequency of the oscillations.



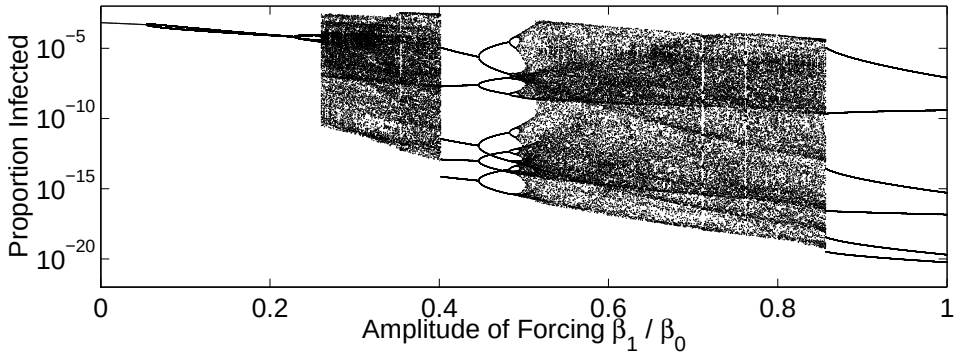

For sinusoidal forcing the average, maximum and minimum proportion infectious. The left-hand figure sets the period at one year and varies the amplitude, the right-hand figure sets $\beta_1 = 0.1\beta_0$ and varies the period. For this disease the natural period of the epidemics is just over two years.

Some classic work by Olsen and Schaffer showed that the dynamics of diseases, such as measles, undergoes a variety of complex bifurcations as the parameter $\beta_1$ is increased.

This figure shows the infection level on a particular day each year. For each value of $\beta_1$, we start at the fixed point of the unforced system and iterate the equations forwards, ignoring the first few hundred years of transients. We note that strictly multiple attractors can coexist – we shall consider this in more detail later.

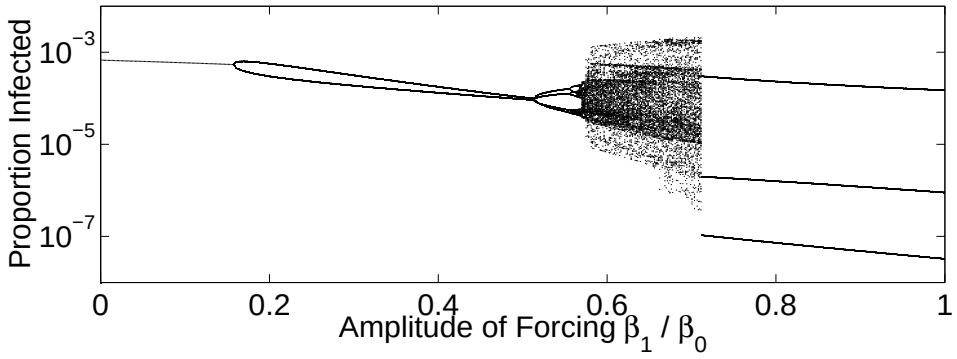
More recently a far more mechanistic approach has been adopted for childhood infections. It was realised (Fine and Clarkson 1982) that the major forcing component of such diseases was the greater mixing of children during school terms compared to school holidays. Therefore the contact (transmission) parameter was modified:

$$\beta(t) = \beta_0 + \beta_1 \times \text{Term(t)}$$




A bifurcation model of measles with sinusoidal, showing the various transitions at the forcing increases.

Where Term(t) is 1 during school terms and $-1$ during school holidays. As this mixing term is a series of step functions there is appears to be little or no change of doing any detailed mathematical analysis, instead we must rely on computer simulation. The figure below shows the corresponding bifurcation diagram for the term-time forcing model. Clearly, the region of stable biennial dynamics (similar to the observed patten) is much bigger, with the chaotic regime reduced significantly.




A bifurcation model of measles with term-time forcing, showing the various transitions at the forcing increases.

However, recent work has discovered a simplified means of examining the complex dynamics that arises. Consider a long school term, during that period $\beta$ is constant so orbits will undergo damped oscillations to the fixed point. We now switch suddenly to school holidays when the mixing is less. During the holidays $\beta$ is again constant, so we have damped oscillations to a different fixed point. The dynamics of these childhood diseases (with term-time forcing) can be considered as flipping between two attracting fixed points – the eigenvalues at these points and the rate of flipping will determine the deterministic attractor (Keeling, Rohani and Grenfell 2001).

Two more recent advancements have been made which can improve on the general picture for measles. Firstly Earn *et al* (2000) considered a data-driven time-varying $\beta$ value. Age-structured models (see next set of lectures) show that the seasonal forcing predominantly comes from the mixing between susceptible and infectious primary school children. As such the aggregate level of seasonality is determined by the distribution of infection (and susceptibility) across ages. It

is therefore feasible to use the expected results of an age-structured model, or the aggregate observed dynamics to generate a forcing function that accounts for both school terms and variations in age-profiles. We note however that such models are parameterised for the deterministic attractor and are therefore not necessarily reliable for transient behaviour.

A second more statistical approach has been developed in recent years by Grenfell and coworkers. This is a discrete-time model (and therefore related to the Reed-Frost model), in which the states of the system are updated every two weeks – which fits in neatly with the available data and the infection period of the disease.

$$\begin{aligned} S_{t+1} &= B_{t+1} - \beta_t S_t I_t^\alpha \\ I_{t+1} &= \beta_t S_t I_t^\alpha \end{aligned}$$

where the parameter $\alpha$ allows us to mimic non-random mixing. In practice, data is available on the yearly birth rate $B_t$, and the number of weekly reported cases – which is assumed to be proportional to $I_t$. We note that

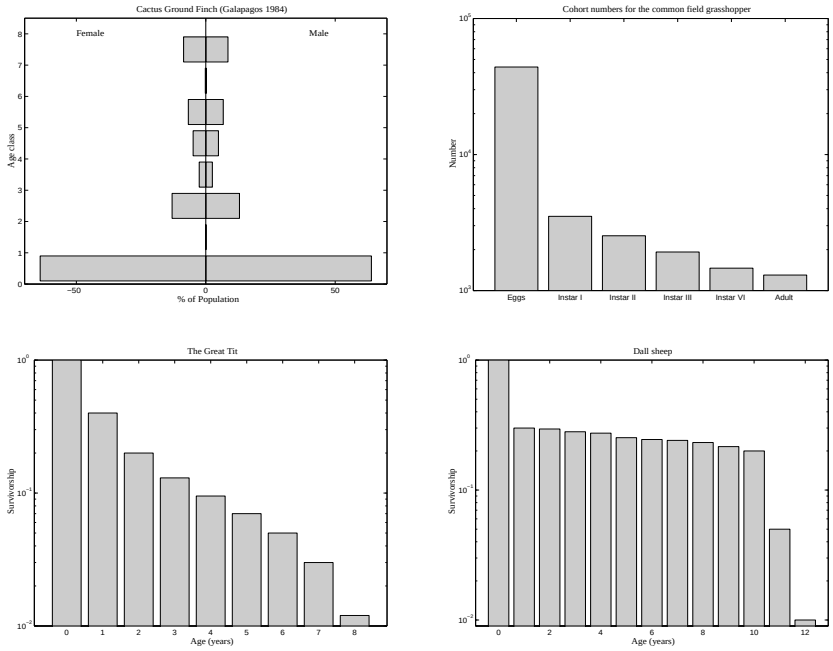
$$S_t = S_0 + \sum_{s=1}^{t} (B_s - I_s)$$

By assuming that the level of susceptibility doesn't vary dramatically, and that everyone is eventually infected, we can find the proportionally constant between reported cases and $I_t$. Secondly, if we log the infection equation:

$$\log(I_{t+1}) = \log(\beta_t) + \log(S_t) + \alpha log(I_t) = \log(\beta_t) + \log(S_0) + \sum_{s=1}^{t} (B_s - I_s)/S_0 + \alpha log(I_t)$$

It is therefore a relatively simple linear problem to estimate all the required elements $S_0$ and the seasonally varying $\beta_t$. This approach leads to a very accurate predictive model for measles. However, it is not a truly mechanistic model, many of the parameters need to be estimated from detailed data – it cannot therefore be applied to new situations (such as vaccination or changes in school-terms) for which no data exists.

So far we have thought of all the organisms of a given species to be identical. One step toward incorporating the individuality of populations is to acknowledge the effects of age on an individual. For many organisms, including humans, the chances of reproducing or dying are very much age dependent. For other species, such as many insects and amphibians, there are several different stages that are passed through on the route to maturity (eg, caterpillars and butterflies or tadpoles and frogs); with such organisms the different stages have radically different behaviour. Finally, in some species (especially fish), large adults may feed upon juvenilles which introduces interesting feedbacks and delays into the dynamics.




Examples of age profiles for a variety of species. Top left, the Cactus Ground Finch shows a high degree of stochasticity in the number of individuals of a given age, but remarkable symmetry between males and females. Top right, the Common Field Grasshopper displaying the number of individual from a single cohort to make it to each life-stage. Bottom left, the Great Tit shows almost a fixed survival probability each year irrespective of age. Bottom right, in contrast the Dall sheep shows high mortality in the first year and after 10 years.

Due to its extra complexity, we shall generally only consider age structure for single species growth models, but of course the principle can be extended to include intra-specific competition or more than one species. Let us consider a population, where the density of creatures of age $a$ is given by $x(a)$ - the dynamics of this species will be given by a P.D.E. If an organism of age $a$ has a reproductive rate $r(a)$ and a death rate $d(a)$, then the population dynamics are given by,

$$\frac{\partial x(a)}{\partial t} \;=\; \delta(a)\int_0^\infty r(b)x(b)db \;-\; d(a)x(a) \;-\; \frac{\partial x(a)}{\partial a}$$

where the Dirac-delta function ensures that young are born aged zero, and the partial derivative with respect to $a$ accounts for the individual becoming older. Notice that even for such a simple

model as this, there is no clear analytical solution.

We can alternatively treat the births as giving a boundary condition at $a = 0$, however this boundary condition will depend on an integral over all $a$.

$$\frac{\partial x(a)}{\partial t} \;=\; -\,d(a)x(a) \;-\; \frac{\partial x(a)}{\partial a} \qquad\qquad \text{and} \quad x(0) \;=\; \int_0^\infty r(a)x(a)da$$

We can now ask two simple questions,

a) *is there an age profile which is a fixed point of the system ?*

$$\frac{dx(a)}{dt} \;=\; 0 \qquad\qquad \Rightarrow \qquad \frac{dx}{da} \;=\; -d(a)x \qquad \text{with} \quad x(0) \;=\; \int_0^\infty r(a)x(a)da$$

From the differential equation we get that,

$$x(a) = X_0 \exp\left(-\int_0^a d(b)db\right)$$

hence to satisfy the boundary condition at $a = 0$ we require

$$\int_0^\infty r(a) \exp\left(-\int_0^a d(b)db\right) da = 1$$

Consider a definite example. We have a population of organisms which reach sexual maturity at two years old, after which they have a constant reproductive rate. The death rate of the organisms increases linearly with age. Therefore the underlying PDE is,

$$\frac{\partial x(a)}{\partial t} \;=\; r\delta(a) \int_2^\infty x(b)db \;-\; Dax(a) \;-\; \frac{\partial x(a)}{\partial a}$$

Looking for a fixed point of the system,

$$x(a) = X_0 \exp\left(-\frac{Da^2}{2}\right)$$

and we only have a solution when,

$$r \int_2^\infty \exp\left(-\frac{Da^2}{2}\right) da = 1$$

and for each $r$, this only has a solution for a unique value of D - it is therefore reasonable to assume that such a system will not exist in the real world.

b) *is there an age profile which grows in time ?*

$$\frac{\partial x(a)}{\partial t} \;=\; \lambda x(a) \;\; = -d(a)x(a) \;-\; \frac{\partial x(a)}{\partial a} \qquad \text{with} \qquad x(0) \;=\; \int_0^\infty r(a)x(a)da$$

From this differential equation we get that,

$$x(a) = X_0 \exp\left(-\int_0^a d(b)db \;-\; \lambda a\right)$$

and so the solution of the boundary condition gives us the eigenvalue of the age profile,

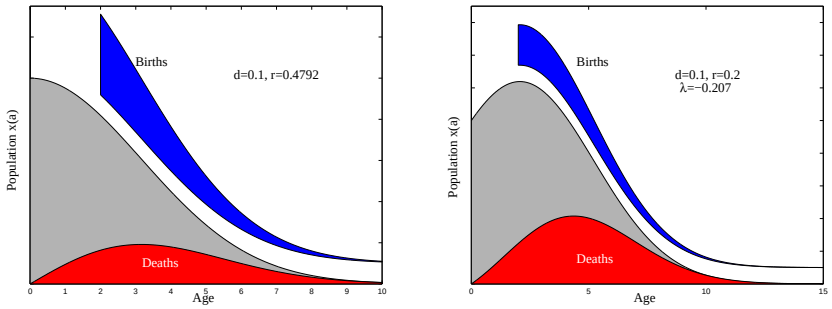
$$\int_0^\infty r(a) \exp\left( - \int_0^a d(b)db \, - \, \lambda a \right) da = 1$$

from which we can obtain $\lambda$, the population growth rate.

Returning to the above example, with $D = 0.1$, $r = 0.2$ we obtain

$$\int_2^\infty \exp\left( -\frac{a^2}{20} \, - \, \lambda a \right) da = 5$$

Numerical calculations show that $\lambda \approx -0.207$ and so the population will decay to extinction.




Theoretical age profiles showing the number that survive as well as the death rates and reproduction rates for the example above. The left-hand graph is for the particular relationship between $r$ and $D$ which gives a fixed population distribution. The right-hand figure shows a population distribution which is an eigenfunction of the system.

Of course such age structured models for a population can also be combined with other factors described earlier. The intra-specific logistic model can be generalised to have arbitrary competition between age-groups

$$\frac{\partial x}{\partial t} \; = \; -d(a)x(a) \int_0^\infty c(a,b)x(b)db \, - \, \frac{\partial x}{\partial a} \qquad \text{with} \qquad x(0) \; = \; \int_0^\infty r(a)x(a)da$$

The interaction between a predator and an age structured victim can be captured by the equations,

$$\begin{aligned} \frac{dP}{dt} \; &= \; -D_P P \, + \, \int_0^\infty A(b)V(b)db \, P \\ \frac{\partial V(a)}{\partial t} \; &= \; -D_V(a)V(a) \, - \, \int_0^\infty C(b)V(b)db \, P - \frac{\partial V(a)}{\partial a} \\ &\text{where} \quad V(0) \; = \; \int_0^\infty B(b)V(b)db \end{aligned}$$

The solution of the one species growth model is simplified if we move from a continuous to a discrete time system. This generally corresponds to systems where there is a strong seasonal component and a particular breeding season – we sample this system on an annual basis. We need to make a further simplifying assumption of a maximum age class - for example age 5. We can now represent the age-structured population by a 5 dimensional vector $\underline{x}$, and the map

relating this year to next year by a matrix which operates on this vector. Such a matrix is called a **Leslie Matrix**, $L$

$$L = \begin{pmatrix} r_1 & r_2 & r_3 & r_4 & r_5 \\ 1-D_1 & 0 & 0 & 0 & 0 \\ 0 & 1-D_2 & 0 & 0 & 0 \\ 0 & 0 & 1-D_3 & 0 & 0 \\ 0 & 0 & 0 & 1-D_4 & 1-D_5 \end{pmatrix}$$

Here $r_a$ is the average number of offspring (which survive until the annual sampling) produced by an individual of age $a$ and $D_a$ is the probability of death during a year for individuals of age $a$. Notice that the term $1 - D_5$ is the proportion of individuals of the maximum age which survive and therefore remain within the maximum age class. The dominant eigenvalue of $L$ tells us whether the population will increase or decrease in the long term, and the associated eigenvector is the final distribution of ages that will be achieved. Notice that such a system is alway linear.

We can now return to the idea of harvesting. Suppose we have a Leslie matrix,

$$L = \begin{pmatrix} 0 & 0 & r \\ S & 0 & 0 \\ 0 & S & T \end{pmatrix}$$

where the three age classes represent juveniles, immatures and adults. Now suppose that after the reproduction phase we can harvest a fixed proportion of each class $H_1$, $H_2$ and $H_3$. It is reasonable to ask what are the optimal proportions such that the harvest is maximised, and the population does not suffer extinction. This translates to finding values of $\underline{H}$, such that the maximum eigenvalue of

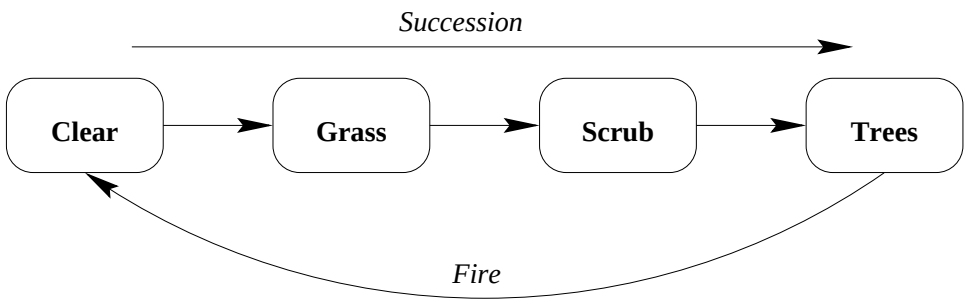
$$L = \begin{pmatrix} 0 & 0 & r \\ S(1-H_1) & 0 & 0 \\ 0 & S(1-H_2) & T(1-H_3) \end{pmatrix}$$

is one and that the amount harvested is maximised.

**Succession Models**

An alternative area where such matrix models can also be of use is in the study of successional species. In this type of system, after an initial disturbance, the habitat is colonised by a succession of species, with the later species only able to invade once the earlier species have improved conditions. Commonly studied examples of this are the colonisation of the inter-tidal habitat by mussels (Paine and Levin 1981), fungal competition (Halley *et al* 1994), and plant communities (Robertson *et al* 1988, Hoffmann 1999). Although the dynamics of these systems have a very strong spatial component, the temporal behaviour can be represented by a transition matrix.

Consider the following example. From bare earth, a habitat can be rapidly colonised by grasses. These grasses bond the soil together, trap moisture and allow dense scrub to form. The scrub offers sufficient protection for trees to grow. Finally the trees are susceptible to forest fires, which clears the habitat back to bare earth.

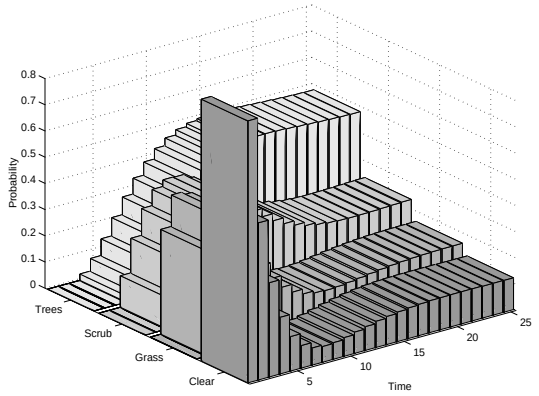



Idealised view of succession between three 'species', with fire acting as the major disturbance.

If we can capture the probability of succession by a transition matrix,

$$M = \begin{pmatrix} 0.6 & 0 & 0 & 0.1 \\ 0.4 & 0.7 & 0 & 0 \\ 0 & 0.3 & 0.8 & 0 \\ 0 & 0 & 0.2 & 0.9 \end{pmatrix}$$

Then over time the probability of finding each type of species is given by the eigenvector of $M$ with the largest eigenvalue (because the amount of habitat does not change, the largest eigenvalue is 1).




Successive iterations of the transition matrix $M$, shows convergence to the dominant eigenvector.

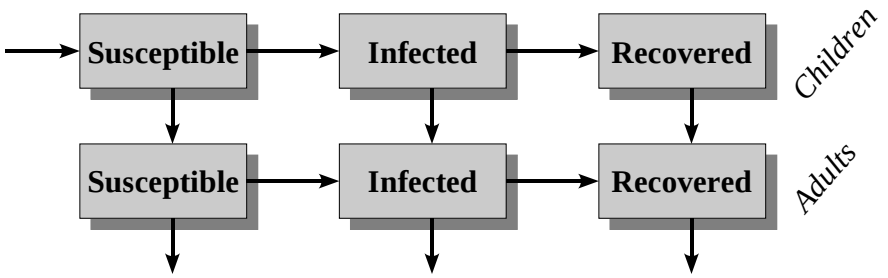
Note: To model this system realistically requires a spatial context - with the fires acting to synchronise large areas of habitat, and proximity increasing the risk of colonisation.

---

Flipse, E. and Veling, E.J.M. 1984 An Application of the Leslie Matrix Model to the Population-Dynamics of the Hooded Seal, *Cystophora-Cristata Erxleben. Ecological Modelling* **24** 43-59
Halley, J.M., Comins, H.N., Lawton, J.H. and Hassell, M.P. 1994 Competition, succession and pattern in fungal communities - towards a cellular automaton model. *Oikos* **70** 435-442
Hoffmann, W.A. 1999 Fire and Population Dynamics of Woody Plants in a Neotropical Savanna: Matrix Model Projections *Ecology* **80** 1354-1369
Leslie, P.H. 1945 On the uses of matrices in certain population mathematics *Biometrika* **33** 182-212
Leslie, P.H. 1948 Some further notes on the use of matrices in population mathematics*Biometrika* **35** 213-245

Paine, R.T. and Levin, S.A. 1981 Intertidal Landscapes : Distribution and the Dynamics of Pattern. *Ecological Monographs* **51** 145-178
Robertson, G.P., Huston, M.A., Evans, F.C. and Tiedje, J.M. 1988 Spatial variability in a successional plant community: patterns of nitrogen availability. *Ecology* **69** 1517-1524
Smith, G.C. and Trout, R.C. 1994 Using Leslie-Matrices to Determine Wild Rabbit-Population Growth and the Potential for Control *Journal of Applied Ecology* **31** 223-230

## RISK-STRUCTURED DISEASE MODELS




A two-class age-structured model, showing the routes between the compartments.

There are many examples of diseases which tend to be age-specific. There are a host of *childhood infections*, such as measles, whooping cough, rubella, chickenpox and mumps, which before vaccination were usually first caught in early childhood. A simple model can explain what common factor meant that these diseases were contracted in childhood. These diseases can be studied using the SIR model. Consider a group of $C$ children that are all born susceptible, ignoring deaths we have:

$$\frac{dC}{dt} = -\beta I^* C \qquad \Rightarrow \qquad C(t) = C(0)\exp(-\beta I^* t) \approx C(0)\exp(-d[R_0 - 1]t)$$

So the number of susceptible children left decays exponentially as they get older, and the decay is faster the larger $R_0$ is. So what defines a childhood disease is simply a high $R_0$, so that you catch the disease before you become an adult. However, the most important thing about childhood diseases, is that they affect children - and that children have very different mixing patterns to adults.

First we consider the slightly simpler case of risk-structured models for sexually transmitted diseases. This is easier as the SIS model is lower dimensional and demography (births and deaths) can be ignored. We separate the population into two distinct classes: high risk individuals who have many sexual partners and low risk individuals who have few partners. Intuitively we would expect the high-risk individuals to have a greater chance of contracting any given STD, but to comprise a smaller proportion of the population. Labelling the two groups by subscripts we maintain the familiar form of the equations,

$$\begin{aligned} \frac{dI_H}{dt} &= S_H\lambda_H - gI_H \\ \frac{dI_L}{dt} &= S_L\lambda_L - gI_L. \end{aligned}$$

where $\lambda$ is the force of infection experienced by each group. Obviously the force of infection to the high-risk group is determined by the number of infectious individuals:

$$\lambda_H = \beta_{HH}I_H + \beta_{HL}I_L$$

where $beta_{HH}$ refers to transmission from the high risk group to itself and should dominate this expression. I will assume that $I_H$ refers to the proportion of the total population that are infected and in the high-risk group ($S_H + I_H + S_L + I_L = 1$). The alternative assumption, that $I_H$ refers to the proportion of the high risk group that are infected ($I_H + S_H = I_L + S_L = 1$) has both advantages and disadvantages.

The full equations for an $n$ class SIS model are therefore:

$$\frac{dI_m}{dt} = S_m \sum_p \beta_{mp} I_p - g I_m$$

The scalar transmission parameter is now replaced by a matrix, known as the *Who Acquires Infection From Whom* matrix. Clearly from this form of the equation the invasion criterion is that one of the eigenvalues at the disease-free equilibrium must have a real part greater than one. How does this compare to the calculation of $R_0$?

If you remember $R_0$ is defined as the average number of secondary cases produced by an average infectious individual in a totally susceptible population. We can easily see what $R_0$ is for each class

$$R_0^m = \sum_p \frac{\beta_{pm} S_p}{g}$$

However, what is the "average infectious individual" – well it is the distribution of infection in the early stages of the epidemic, which is the eigenvector associated with the dominant eigenvalue. From this insight we see that

$$R_0 = \frac{\lambda}{g} + 1$$

where $\lambda$ is the dominant eigenvalue of $\widehat{\beta_{pm}} = \beta_{pm} S_p$. Hence the actual value of $R_0$ is bounded by the individual values of the $R_0^m$.

We now wish to use this formulation to develop an age-structured model. There are other bits of accountancy due to individuals getting older and moving between classes. In general, we can also write down equations for a model with $n$ age classes:

$$\begin{aligned}
\frac{dS_n}{dt} &= B_n + L_{n-1} S_{n-1} - \sum_m \beta_{nm} I_m S_n - L_n S_n - d_n S_n \\
\frac{dI_n}{dt} &= L_{n-1} I_{n-1} + \sum_m \beta_{nm} I_m S_n - L_n I_n - d_n I_n
\end{aligned}$$

So $B_0$ is the birth rate (and all other $B_n$ are zero), $L_n$ is the rate at which individuals mature and leave the $n$ class ($1/L_n$ equals the age-span in class $n$), $d_n$ is the natural death rate in class $n$ and $\beta_{nm}$ is the rate of transmission from class $m$ to class $n$.

There are several complications with studying this model compared to the standard SIR and SIS models.

**1) $\boldsymbol{\beta}$ is now a matrix of values.**
Previously, for a one class model we generally knew the average infectious period, the level of seroprevalence in the population ($1 - S^*$), and the life expectancy. This allowed us to calculate the single parameter $\beta$. In simple terms the information about one age class allows us to estimate one $\beta$ parameter.

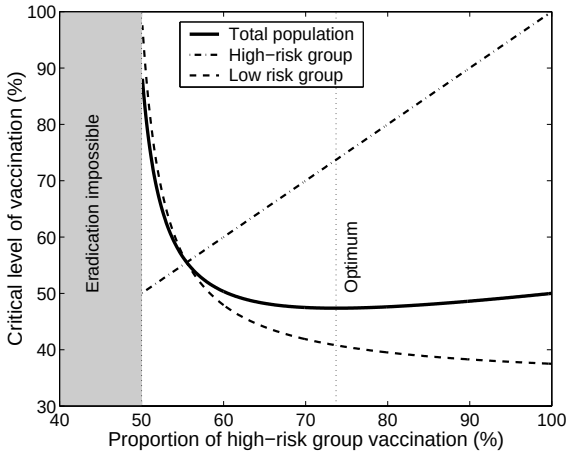
However, for age structured models, we have $n$ age-classes (and hopefully $n$ levels of seroprevalence) but $n^2$ terms in the $\boldsymbol{\beta}$ matrix and hence $n^2$ parameters to estimate – this is impossible. We therefore seek a simpler from of the matrix (with just $n$ parameters) so that we can successfully estimate all the terms. Many different forms of the WAIFW matrix have been tried, in general we need an $n \times n$ matrix which only has $n$ distinct parameters. The most common example is:

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_1 & \beta_1 & \beta_3 & \beta_4 \\ \beta_1 & \beta_2 & \beta_3 & \beta_4 \\ \beta_3 & \beta_3 & \beta_3 & \beta_4 \\ \beta_4 & \beta_4 & \beta_4 & \beta_4 \end{pmatrix}$$

which mimics the increased mixing between school children (the groups are pre-school, primary school, secondary school and adults.
For sexually transmitted diseases the situation is somewhat more straightforward. We can make three simplifying assumption. The first is that we only stratify the population in terms of number of sexual partners; second, that the risk of catching and transmitting infection is proportional to the number of partners; and finally that partnerships form at random. This leads to a much simplified version of the mixing matrix,

$$\boldsymbol{\beta}_{ij} = \beta \times i \times j$$

This again means that we only have one parameter to find, but we have ignored many of the heterogeneities that exist with such systems. In general we observe assortative mixing such that high-risk (many partner) individuals are much more likely to partner other high-risk people.




The critical level of vaccination needed to eradicate a STD, as a percentage of the entire population, as the coverage in the high-risk group increases. In this example, the transmission matrix is $\boldsymbol{\beta} = \begin{pmatrix} 10 & 1 \\ 1 & 2 \end{pmatrix}$, with $N_H = 0.2$, $N_L = 0.8$ and $g = 1$ as before.

**2 Simple Relationships**

Once this additional structure is added to the equations all of the simple results relating $R_0$, the equilibrium levels and the final size of the epidemic. Many of these now become complex matrix equations, however the results for vaccination are worth considering.

Firstly, if we vaccinate at random then the classic threshold result still applies $V_C = 1 - 1/R_0$. Disappointingly, this level of vaccination is often much larger than would be predicted if we

(foolishly) ignored the risk-structure. We can however use the heterogeneities to target our vaccination campaign, focusing on the core-groups that are most responsible for the spread of infection. Often, we can do better than the standard (non-structured) models would predict.

It is still very much an open problem to know who and when to vaccinate. It is clear from the above graph that we do much better if we focus out vaccination at the high risk group. It is also clear that we're safer to focus too much on the high-risk group than not enough. For the parameters given in the graph:

the true value of $R_0 = 2.2472$.

the value of $R_0$ if we ignored structure is 2.

the level of random vaccination needed is 55.5%.

the level of targeted vaccination needed is 47.4%.

This echos the general form of results that occur for almost all structured models. However, in age structured models the situation is far more complex as individuals propagate through classes.

# Multi-Species, Multi-Host & Multi-Strain

The models considered so far have all involved just one or two species. The 2-D nature of the associated differential equations means that the dynamics are relatively straight-forward. The next obvious modification is to include more species, and in particular another trophic level.

The classic model with three trophic levels can either be considered as resource-prey-predator or prey-predator-top predator. Either way, the same model structure is used:

$$\begin{aligned}\frac{dR}{dt} &= BR - ERV \\ \frac{dV}{dt} &= FRV - CVP \\ \frac{dP}{dt} &= AVP - DP\end{aligned}$$

Such models can easily lead to much more complex dynamics including chaos. An interesting modification to this model has been studied by (Blasius *et al* 1999) to understand the Lynx-Hare cycle in the Canadian arctic. Here the variables are lynx ($L$), snowshoe hares ($H$) and vegetation ($V$):

$$\begin{aligned}\frac{dV}{dt} &= V - 0.1LV - 1.5 \\ \frac{dH}{dt} &= 0.1LV - 0.6HL - H \\ \frac{dL}{dt} &= 0.6HL - 10L + 0.1\end{aligned}$$

This slight modification is capable of producing regular 'outbreaks' of lynx, but with chaotic maxima.

Once we have three species, a host of alternative arrangements of feeding-structure exist. For example, we could have two predator species with a shared prey:

$$\begin{aligned}\frac{dV}{dt} &= BV - C_1VP_1 - C_2VP_2 \\ \frac{dP_1}{dt} &= A_1VP_1 - D_1P_1 \\ \frac{dP_2}{dt} &= A_2VP_2 - D_2P_2\end{aligned}$$

Here the two predators compete for a limited resource, and the most fit predator (the one that can survive on the lowest level of prey) will force the other species to extinction. Considering the proportion of $P_1$ to the total predator population, $p_1$:

$$\frac{d}{dt}\frac{P_1}{P_1+P_2} = \frac{dp_1}{dt} = p_1p_2\left[(A_1V - D_1) - (A_2V - D_2)\right]$$

Although the two predator species do not directly interact, they suffer competition by driving down the prey of the other species.

Or we could have two prey species with a common predator:

$$\begin{aligned}\frac{dV_1}{dt} &= B_1V_1 - C_1V_1P \\ \frac{dV_2}{dt} &= B_2V_2 - C_2V_2P \\ \frac{dP}{dt} &= (A_1V_1 + A_2V_2)P - DP\end{aligned}$$

Again competition works in the same way:

$$\frac{d}{dt}\frac{V_1}{V_1+V_2} = \frac{dv_1}{dt} = v_1v_2\left[(B_1 - C_1P) - (B_2 - C_2P)\right]$$

the competition is even more indirect as the two species interact by increasing the level of predation on the other species. This model however, takes a very simplistic view of predation with the predator consuming both prey; if non-linear switching functions are used, so that the predator preferentially feeds on the more abundant prey then co-existence is possible.

---

---

## Generalised Lotka-Volterra Type Models

Let $\underline{x} \in I\!R^n_+$ be a vector of population densities, then a generalised Lotka-Volterra type model can be written as,

$$\begin{aligned}\frac{dx_i}{dt} &= \left[r_i + \sum_j a_{ij}x_j\right]x_i = \phi_i(\underline{x})x_i \\ \Rightarrow \qquad \frac{d}{dt}log(x_i) &= \phi_i(\underline{x})\end{aligned}$$

We note that internal equilibrium points are given by,

$$\underline{\phi}(\underline{x}) = 0 \qquad \Rightarrow \qquad A\underline{x} = \underline{r} \qquad A = a_{ij}$$

This whole family of equations can incorporate logistic growth and mass-action type interaction between species. It therefore provides a fairly robust method of looking at multi-species food webs.

For such a system we find that,
a) No interior equilibrium $\Longleftrightarrow$ every orbit converges to the boundary.

($\Leftarrow$) trivial
($\Rightarrow$) Set $K = \phi(int\, \mathbb{R}^n_+)$. Due to the linear nature of $\phi$, $K$ is convex and $\exists$ hyperplane $H \ni \underline{0}$ disjoint from $K$.
$\exists\ c \in \mathbb{R}^n_+$ such that $c \perp H$ and $c.y > 0\ \forall y \in K$
Now set $V(x) = \sum_i c_i log(x_i)$ therefore

$$\dot{V} = \sum_i c_i \frac{\dot{x_i}}{x_i} = c.\phi > 0 \qquad \text{as } \phi \in K$$

Suppose there exists an orbit $x(t)$ which does not converge to the boundary. Then we can find a sequence $T_i$ such that,

$$f^{T_i}(x) = x(T_i) \to y$$

Therefore,

$$V(f^{T_i}(x)) \leq V(f^{T_i+t}(x)) \leq V(f^{T_j}(x)) \qquad \forall T_i \leq t_i + t \leq T_j$$
$$\Rightarrow \qquad V(y) = V(f^t(y)) \qquad \forall t$$

Therefore $y$ must be an equilibrium point of the system.

b) If there exists an attractor $\Lambda$ which is bounded away from infinity, and $B(\Lambda)$ is its basin of attraction, then for $\underline{x}(0) \in B(\Lambda)$

i)

$$\overline{\underline{\phi}} \;=\; \lim_{T\to\infty} \frac{1}{T} \int_0^T \underline{\phi}(\underline{x}(t))dt \;=\; 0$$

$$\log(x_i(T)) - \log(x_i(0)) \;=\; \int_0^T \frac{d}{dt} \log(x_i)dt \;=\; \int_0^T \phi_i(x)dt$$

and as the left-hand side is bounded, so $\overline{\underline{\phi}} = 0$.

ii)

$$\overline{\underline{x}} \;=\; \lim_{T\to\infty} \frac{1}{T} \int_0^T \underline{x}(t)dt \text{ is a fixed point}$$

Consider,

$$\phi\left(\lim \frac{1}{T} \int_0^T x(t)dt\right) \;=\; \lim \frac{1}{T} \int_0^T \phi(x(t))dt \qquad \text{as } \phi \text{ linear}$$
$$\phi(\overline{\underline{x}}) \;=\; 0$$

Hence $\overline{\underline{x}}$ is a fixed point of the system.

These generalised Lotka-Volterra models, include competition models, complex multi-trophic food webs, disease models *etc*; therefore the simple properties of all such systems have wide ranging applications.

**Multi-Host Diseases**

If a disease can infect multiple hosts, then the models that can be applied are the same as those for age-structure. We can qualify this with a simple example of the spread of foot-and-mouth between buffalo ($B$) and domestic cattle ($C$).

$$\begin{aligned}
\frac{dS_B}{dt} &= B_B - (\beta_{BB}I_B + \beta_{BC}I_C)\, S_B - d_B S_B \\
\frac{dI_B}{dt} &= (\beta_{BB}I_B + \beta_{BC}I_C)\, S_B - g_B I_B - d_B I_B \\
\frac{dS_C}{dt} &= B_C - (\beta_{CB}I_B + \beta_{CC}I_C)\, S_C - d_C S_C \\
\frac{dI_C}{dt} &= (\beta_{CB}I_B + \beta_{CC}I_C)\, S_C - g_C I_C - d_C I_C
\end{aligned}$$

For this example, we would expect the matrix $\beta$ to be dominated by the diagonal terms, and from field observations $g_B < g_C$.

A more complex set of dynamics can occur when there are obligate hosts involved in the transmission process. Such a scenario exists for the spread of malaria (where the parasite is transmitted from person to person via mosquitoes), or bubonic plague (where the bacterium are transmitted from rodent to rodent via infected fleas and occasionally to humans). For malaria, the following would be a plausible model of mosquitoes ($M$) and humans ($H$).

$$\begin{aligned}
\frac{dS_H}{dt} &= B_H - \beta_{HM} I_M S_H - d_H S_H \\
\frac{dI_H}{dt} &= \beta_{HM} I_M S_H - m_H I_H - d_H I_H \\
\frac{dS_M}{dt} &= B_M - \beta_{MH} I_H S_M - d_M S_M \\
\frac{dI_M}{dt} &= \beta_{MH} I_H S_M - d_M I_M
\end{aligned}$$

For this model it might be more sensible to let $S_M$ and $I_M$ refer to the density of mosquitoes (rather than proportion) as we would expect bite-rate and hence transmission to increase with density.

**Multi-Strain Diseases**

We wish to consider two strains of disease that are interacting. The simplest model is to assume that the two strains are completely cross-reactive, such that immunity to one strain automatically confers immunity to the other. Following the basic SIR model, we now have two $I$ classes dependent on the infecting strain.

$$\begin{aligned}
\frac{dS}{dt} &= B - \beta_1 I_1 S - \beta_2 I_2 S - dS \\
\frac{dI_1}{dt} &= \beta_1 I_1 S - g_1 I_1 - dI_1 \\
\frac{dI_2}{dt} &= \beta_2 I_2 S - g_2 I_2 - dI_2 \\
\frac{dR}{dt} &= g_1 I_1 + g_2 I_2 - dR
\end{aligned}$$

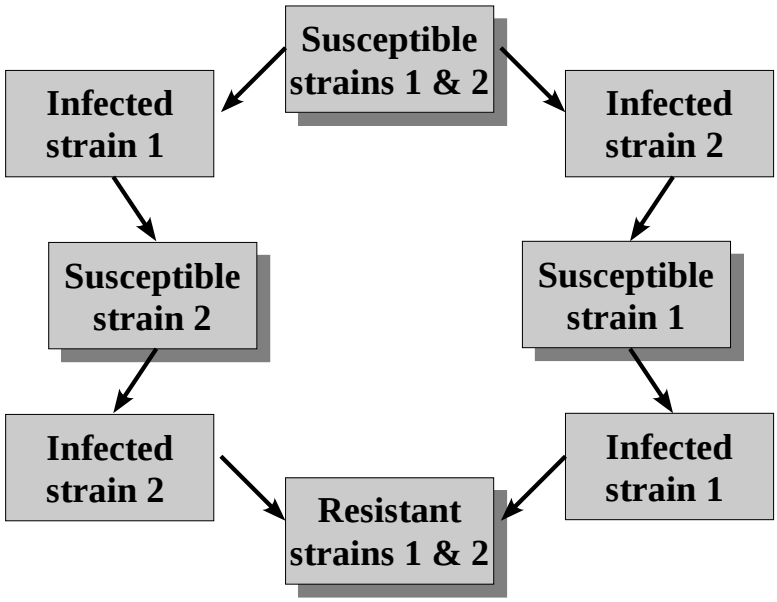
From this simple model we see that one strain ALWAYS dominates. This is clear if we look at the ratios of the two strains,

$$\frac{d}{dt}\frac{I_1}{I_1 + I_2} = \frac{I_1 I_2}{(I_1 + I_2)^2}\left[(\beta_1 S - g_1) - (\beta_2 S - g_2)\right]$$

From which we find that the strain with the largest $R_0$ always dominates.

Much more complex dynamics are obtainable when two strains only confer partial immunity.

This model has eight compartments, and a variety of different transmission parameters (which measure the extent of cross-immunity). The most difficult part of working with such models is developing an intuitive notation. A common example, which can be extended to model any number of strains is to have a string of subscripts; thus $N_{SI}$ is the number of individuals that




A two-strain model, showing the routes to complete immunity.

are susceptible to strain 1 and infected with strain 2. The dynamics of each subscript can then be written down as before: eg.

$$\begin{aligned}
\frac{dN_{SS}}{dt} &= B_{SS} - \beta_1^{SS} \sum_X N_{IX} N_{SS} - \beta_2^{SS} \sum_X N_{XI} N_{SS} - dN_{SS} \\
\frac{dN_{SI}}{dt} &= \beta_2^{SS} \sum_X N_{XI} N_{SS} - \beta_1^{SI} \sum_X N_{IX} N_{SI} - g_2 N_{SI} - dN_{SI} \\
\frac{dN_{SR}}{dt} &= g_2 N_{SI} - \beta_1^{SR} \sum_X N_{IX} N_{SR} - dN_{SR}
\end{aligned}$$

**Multiple Malaria Strains, No cross immunity**

An example of such models with no cross-immunity comes from the epidemiology of *Plasmodium falciparum*, the malaria-causing parasite. The basic reproductive ratio, $R_0$ of malaria is generally estimated from the increase with age ($a$) of the proportion seropositive, that is the proportion that are no longer susceptible, $1 - S(a)$:

$$S(a) \;=\; \exp\left(-\frac{R_0 - 1}{L}\, a\right) \tag{1}$$

where $L$ is the average life expectancy. From such calculations $R_0$ has been estimated in the range 50 to 100, which makes its eradication extremely difficult. However, these calculations assume that malaria is a single disease causing organism and therefore following infection there is complete immunity. However, recent evidence suggests that a diverse range of antigenically distinct strains may cause the disease that is labelled malaria.

Following the work of Gupta *et al.* (1994), we consider several strains of *Plasmodium falciparum* and suppose that the strains offer no cross immunity. A relaxation of this assumption, so that strains confer partial immunity does not radically change the general conclusions. Thus, for strain $i$:

$$S^i(a) \;=\; \exp\left(-\frac{R_0^i - 1}{L}\, a\right)$$

However, as there is no cross immunity the strains act independently; therefore the proportion of individuals that are totally susceptible and have no malaria antibodies again any strain can be calculated:

$$S^{\text{total}}(a) \;=\; \prod_i S^i(a) \;=\; \exp\left(-\frac{\sum_i R_0^i - 1}{L}a\right) \tag{2}$$

Thus, comparing equations (1) and (2) we see that the $R_0$ that is derived under the assumption of a single malaria strain is actually the sum of the separate $R_0$'s when multiple strains co-exist. Hence the true value of $R_0$ for each strain is likely to be greatly reduced compared to standard estimates; Gupta *et al.* (1994) calculate that $R_0$ may be as low as 6 or 7, so that instead of having one disease that is very hard to eradicate this analysis suggests that we have multiple strains each of which may be far easier to eliminate.

**Enhanced Susceptibility**

We now focus on the situation where co-infection with two or more strains is more likely than pure chance would dictate. The classic example here is sexually transmitted infections, where the presence of one infection can increase the susceptibility of the host to others. This enhanced susceptibility can lead to some surprising results as discussed below. Sexually transmitted infections usually conform to the SIS paradigm, where after treatment infectious individuals are once again susceptible. There are four differential equations corresponding to the two possible states (S and I) and the two diseases:

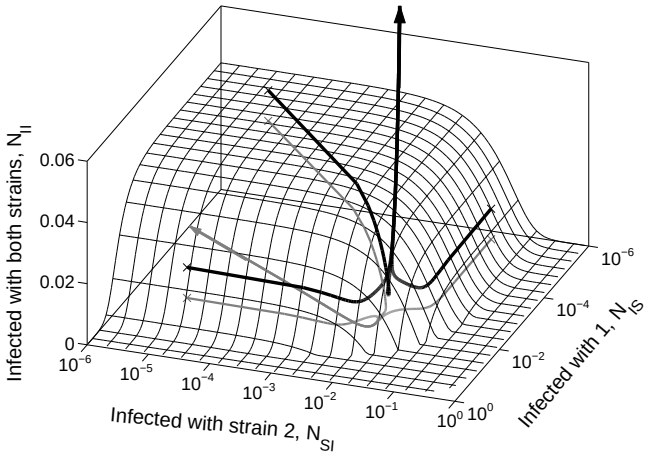
$$\begin{aligned}
\frac{dN_{SS}}{dt} &= -\beta_1 N_{SS} I_1 - \beta_2 N_{SS} I_2 + g_1 N_{IS} + g_2 N_{SI} + g_3 N_{II} \\
\frac{dN_{IS}}{dt} &= \beta_1 N_{SS} I_1 - g_1 N_{IS} - \widehat{\beta}_2 N_{IS} I_2 \\
\frac{dN_{SI}}{dt} &= \beta_2 N_{SS} I_2 - g_2 N_{SI} - \widehat{\beta}_1 N_{SI} I_1 \\
\frac{dN_{II}}{dt} &= \widehat{\beta}_1 N_{SI} I_1 + \widehat{\beta}_2 N_{IS} I_2 - g_3 N_{II} \\
& I_1 = N_{IS} + N_{II} \qquad I_2 = N_{SI} + N_{II}
\end{aligned} \tag{3}$$

These equations assume that those individuals with both infections would be treated for both simultaneously (at rate $g_3$); we also make the simplifying assumption that infections are passed on independently, such that those with both infections do not necessarily pass both on to each individual they infect. We are now particularly interested in the case where being infected with one disease increases the susceptibility to the other, which translates to $\widehat{\beta}_1 > \beta_1$ and $\widehat{\beta}_2 > \beta_2$.

One interesting feature of such enhanced susceptibility is its effect on the invasion and persistence of the two infections. For either disease to invade we require, as usual, that $R_0$ is greater than one; in particular for both diseases to be able to invade we need:

$$\beta_1 > g_1 \qquad \text{and} \qquad \beta_2 > g_2$$

At invasion, because the density of infection is low and hence co-infection every rarer, the terms $\widehat{\beta}_1$, $\widehat{\beta}_2$ and $g_3$ do not enter into the invasion criterion. However, once the diseases are established and co-infection common these terms can play a pivotal role in maintenance of these infections. In particular, if $\beta_1$ and $\beta_2$ are small but $\widehat{\beta}_1$ and $\widehat{\beta}_2$ are large, we can experience an Allee effect, whereby the diseases cannot invade but may persist once they become established. The figure shows an example of this; with orbits that start above a critical prevalence tending to an endemic equilibrium whereas orbits that start below tend to zero.




Example of orbits from the enhanced susceptibility model clearly demonstrating the Allee effect. The surface separating persistence from extinction is also shown. Grey orbits start just below the surface and lead to extinction, whereas black orbits start just above the surface and tend to the same fixed point. ($g_1 = g_2 = g_3 = 1$, $\beta_1 = 0.9$, $\beta_2 = 0.85$, $\widehat{\beta_1} = 8$, $\widehat{\beta_2} = 7$.)

**Evolutionary Implications: a simplification**

Substantial simplifications can be made given two reasonable assumptions (Gog & Grenfell 2003). The first is that immunity acts to reduce transmission, not susceptibility, this means that all individuals are equally at risk to any strain irrespective of their epidemic history. Physiologically this assumption implies that even when an individual has immunity to a particular strain, when challenged they mount a full immune response but clear the pathogen before transmission can occur. The second assumption is that partial immunity acts by making a proportion of all infected individual totally immune , rather than giving all individuals a reduced transmission rate. These two assumptions mean that we only need to consider the number of individuals susceptible to and infectious with each strain. Hence for a model with $n$ strains there are only $2n$ equations.

$$\begin{array}{rcl} \frac{dI_i}{dt} & = & \beta_i S_i I_i \, - \, g_i I_i \, - \, dI_i \\ \frac{dS_i}{dt} & = & B - \sum_j \beta_j \sigma_{ij} S_i I_j \, - \, dS_i \end{array} \tag{4}$$

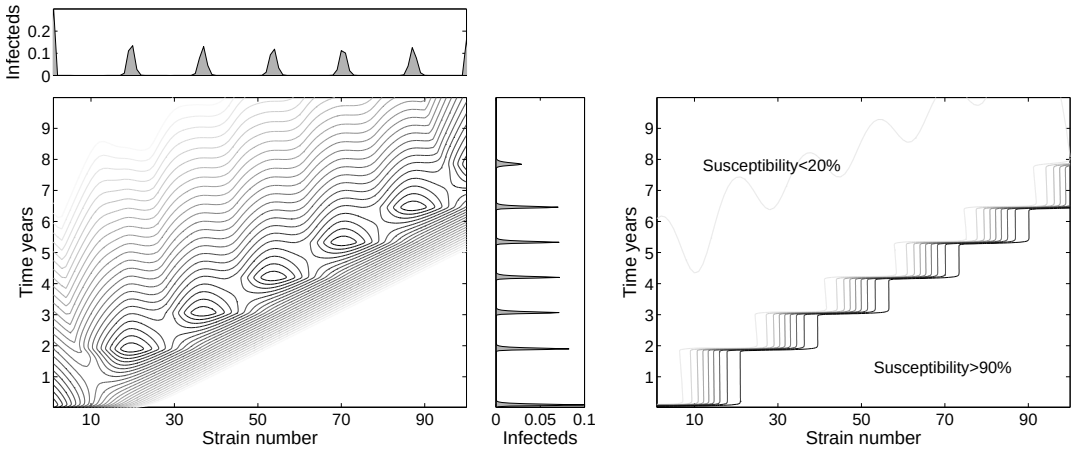
where $0 \leq \sigma_{ij} \leq 1$ informs about the gain of complete immunity to strain $i$ due to infection (or challenge) with strain $j$. We would normally insist that $\sigma_{ii} = 1$ so that the single strain dynamics are SIR. It should be noticed that the susceptibility classes are not mutually exclusion, an individual susceptible to strains 1 and 2 will belong to both $S_1$ and $S_2$, therefore the population components no longer sum to the population size.

This simple, but powerful model, can be used to investigate antigenic drift which is evolution driven by the immunity of the population with no discernible change in disease characteristics (so $\beta_i = \beta$ and $g_i = g$). We position strains in a one-dimensional line and assume the immunity is conferred most strongly to nearby strains:

$$\sigma_{ij} \; = \; \exp(-A[i-j]^2)$$

It is also presumed that random mutation, at a rate $\varepsilon$, can lead to the spontaneous creation of adjacent strains. This modifies the basic equation:

$$\frac{dI_i}{dt} \; = \; \beta S_i I_i \, - \, g I_i \, - \, d I_i \, - \, \varepsilon I_i \, + \, \frac{1}{2}\varepsilon I_{i+1} \, + \, \frac{1}{2}\varepsilon I_{i-1}$$




Dynamics of a strain-structured model of influenza (eqn 4), where partial immunity is conferred to nearby strains and is assumed to act on the transmission. The left-hand graphs show the level of infection with the population; the main graph giving the strain specific level over time $I_i(t)$ (contours plotted on logarithmic scale), the top graph shows the total number infected with each strain $\int I_i(t)dt$, the right graph shows the total infected over time $\sum_i I_i(t)$. The right-hand graph gives the level of susceptibility to each strain $S_i(t)$, here the contours are linearly spaced. ($B = d = 5 \times 10^{-5}$ per day, $g = 1/7$ per day, $R_0 = 3$, $A = 0.01$, $\varepsilon = 0.01$ per year.)

The figure shows an example of the type of drift-dynamics predicted by this model. With 100 strains, a tradition approach which tracks all infection histories would require over $6 \times 10^{31}$ equations which is computationally infeasible. Two clear features emerge from this model which reflect the observed behaviour of influenza. Firstly, there is clear oscillations in the over-all level of infection driven by the strain-structure and rate of mutation; in practice any oscillations in the prevalence of influenza is reinforced by seasonal effects. Secondly, due to the cross immunity invoked in nearby strains the next epidemic strain must be sufficiently distinct from previous strains.

Stochasticity can be a major driving force in the behaviour of epidemics, and yet it is one of the most difficult concepts to deal with mathematically. Most of the results shown therefore come from computer simulations of stochastic populations. There are three main forms of stochasticity, environmental, additive and demographic.

**Environmental**
Environmental stochasticity, as its name suggests, comes from sources external to the disease population. This is generally modelled as random noise added to the parameters; most commonly we set $\beta = \beta_0 + \beta_1\xi = \beta_0(1 + F\xi)$, where $\xi$ is some type of noise term. Such a form of stochasticity can be easily studied mathematically, but neglects the individual nature of the population.

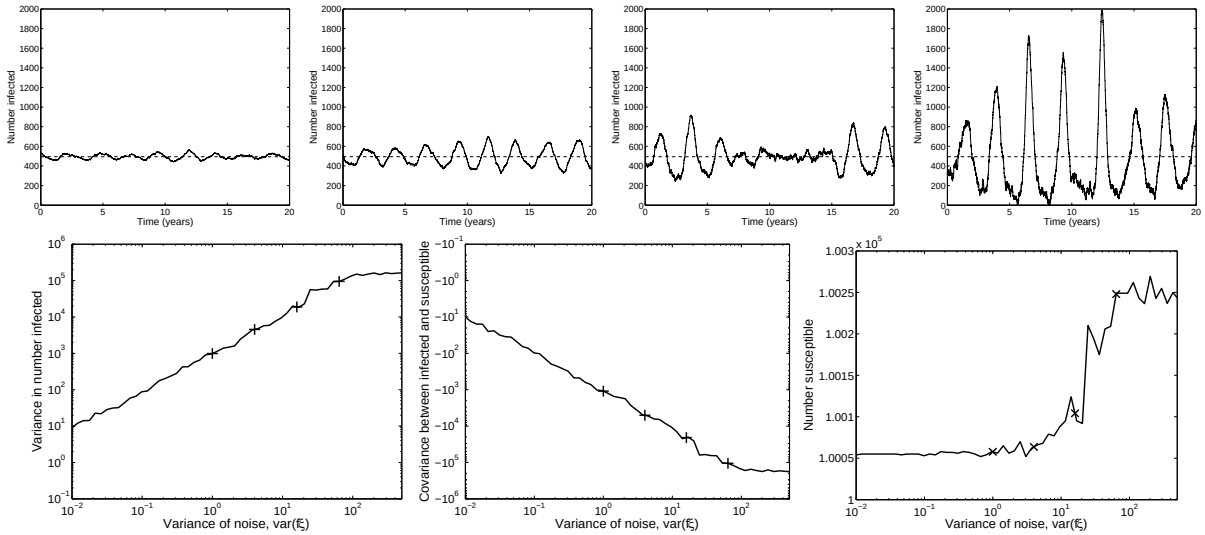
**Additive**
Additive stochasticity, or noise, is random perturbations applied to the basic variables (usually the number of infecteds). The better models assume that the noise scales in some non-linear way with the level of infection. Thus our equation becomes:

$$\frac{dI}{dt} = \beta IS - gI + f(S, I)\xi$$

where $\xi$ is again some appropriate noise term. Although this type of model can be very effective, it again ignores the individual nature of the population.

When $f$ is constant, the long-term dynamics as the value of $f$ and hence the amount of noise is varied. As expected, when more noise is added to the equations, the dynamics deviate further from the deterministic equilibrium. However, rather than these deviations being completely random, there is a distinct oscillatory nature and frequency to the behaviour. This is caused by the interaction between deterministic and stochastic forces – a common feature in all stochastic models (Renshaw 1991; Rohani, Keeling and Grenfell 2002). The noise terms allow the system to wander away from the deterministic (fixed-point) equilibrium, but the underlying deterministic model forces it back towards the equilibrium point. This return movement closely matches the deterministic prediction of decaying oscillations with a natural period (see Chapter 1). This is a classical physical phenomena, that random noise causes systems to oscillate at (or very near to) their natural frequency. We can therefore see noise (and most forms of stochasticity) as over-coming the deterministic behaviour when it is weak close to the equilibrium point, but dominated by the deterministic behaviour far from the equilibrium point. This battle between deterministic and stochastic forces, means that the transient dynamics on the approach to the deterministic equilibrium often dominate the observed stochastic behaviour. However, if the noise term is very large (greater than shown) it can completely swamp any deterministic component and a pure random-walk is observed where negative infection levels are possible.

The figure also shows how the aggregate properties of the dynamics vary with the noise level. The variance in the number of infectious individuals increases almost linearly with the variance in the noise. For the very highest levels of noise however the variance does not increase linearly; this is due to the action of strong non-linear forces that operate when the amplitude of the resonant epidemics becomes large and the number infectious is occasionally forced close to zero. Perhaps more surprisingly, there is a corresponding decrease in the covariance between susceptible and infected individuals. This is because if a particular value of the noise is 'good' for




The top row gives examples of the dynamics of the SIR model with births and deaths ($B = d = 5.5 \times 10^{-5}$, $R_0 = 10$, $g = 1/10$, $N = 10^6$). The amount of noise added to the transmission terms increases from left to right. The bottom row shows how the variance in the number infected, the covariance between susceptible and infecteds and the average proportion susceptible change with the variance in the noise. From biological considerations, if noise ever forced the number of infected individuals below zero, the number was reset to zero – this only occurs for the largest levels of noise.

the infected population it is 'bad' for the susceptibles; hence noise generates a strong negative covariance.

This negative covariance has implications for the average dynamics. The rate at which cases arise ($\beta SI$) is now reduced, as the product of $S$ and $I$ is smaller than expected due to the negative covariance. This reduction in transmission (and therefore a reduction in the effective reproductive ratio $R$) is reflected by an increase in the mean level of susceptibles. This is a very important concept; stochasticity does not just cause variations about the mean (deterministically predicted) values, but can actually cause significant changes to these means. It is the non-linear nature of the disease equations (in particular the product of $S$ and $I$ in the infection term) that allows stochasticity to modify the mean values. From the graph it is clear that stochasticity has little effect on the mean values when the oscillations are small and therefore the localised dynamics are almost linear.

**Inclusion of Noise in Differential Equations**

Consider the simplest stochastic differential equation model:

$$\frac{dx}{dt} = \text{Noise}$$

The most simple means of solving such equations is using the Euler method of integration, breaking time into small components, $\delta t$.

$$x_{t+\delta t} = x_t + \delta t \frac{dx}{dt} = x_t + \delta t \text{Noise} = x_0 + \delta t \sum_{1}^{\frac{t}{\delta t}} \text{Noise}$$

Thus the equations progress as the summation of many small noise terms. However, if the noise terms are all independent, then the variance of $x$ at any time decays as the step size is made smaller. Thus, in the limit $\delta t \to 0$, when the updating method is exact all the noise

terms effectively cancel. This is a reflection of the well-known problem that there is no simple mathematical method of expressing the noise term.
The simplest solution to this problem is to scale the noise term with respect to the integration step. Throughout this section we shall assume that:

$$\xi = \frac{RANDN}{\sqrt{\delta t}}$$

such that as the time step of integration decreases the amplitude of the noise increases. If this scaling is used with higher-order integration methods (such a Runge-Kutta) then the noise should be calculated *before* each integration step.
This new definition of $\xi$ has the properties that we require. Such that if

$$\frac{dx}{dt} = f\xi$$

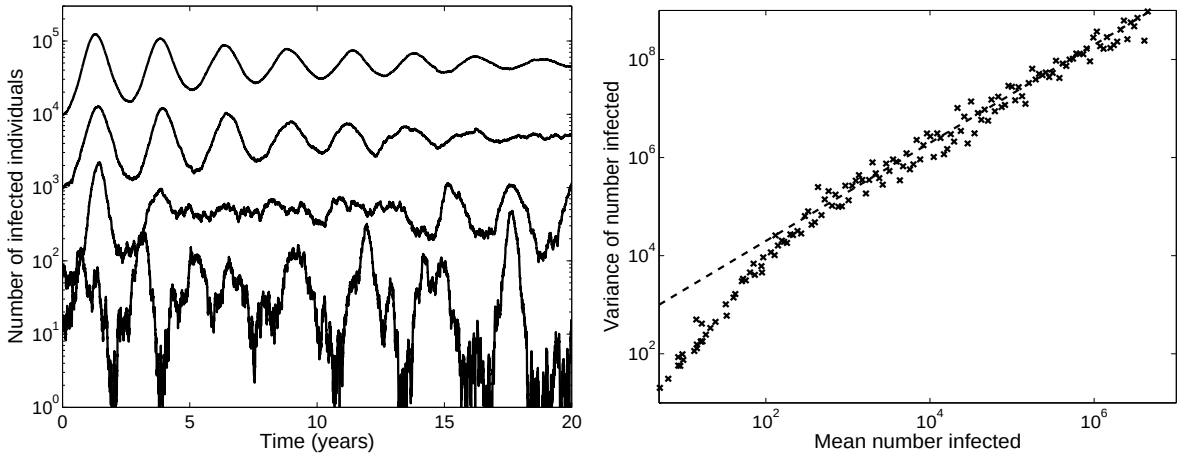
then, averaged over multiple simulations, the mean of $x$ is zero while the standard deviation grows like $ft$, and therefore the dynamics correspond to a random walk.

**Scaled Additive Noise**
While the simple method of including noise outlined above provides variability, the amplitude of such random terms has so-far been assumed to remain constant. In practice this is rarely the case, generally the absolute magnitude of the variability increases with increasing population size, although the relative magnitude usually decreases (Keeling & Grenfell 1999; Bjørnstad, Finkenstädt & Grenfell 2002). If events, such as infection, occur at random then in any short time interval the number of events will be Poisson distributed and hence the variance will be equal to the mean. This provides a direct means of determining the magnitude of the noise term, denoted $f$ in the original equation, such that $f = \sqrt{rate}$. This concept can be extended still further, such that each event is given an associated noise term:

$$\begin{aligned}
\frac{dS}{dt} &= [BN + \sqrt{BN}\,\xi_1] - [\beta SI/N + \sqrt{\beta SI/N}\,\xi_2] - [dS + \sqrt{dS}\,\xi_3] \\
\frac{dI}{dt} &= [\beta SI/N + \sqrt{\beta SI/N}\,\xi_2] - [gI + \sqrt{gI}\,\xi_4 - [dI + \sqrt{dI}\,\xi_5] \\
\frac{dR}{dt} &= [gI + \sqrt{gI}\,\xi_4] - [dR + \sqrt{dR}\,\xi_6]
\end{aligned} \tag{5}$$

This set of equations now contains six distinct noise terms ($\xi_1, \ldots \xi_6$) one for each event-type; the same noise is used for each event even though it may appear in multiple equations. Note that although at equilibrium the event rates cancel (such that $\frac{d}{dt} = 0$), the noise terms add together. Thus, the amount of noise is determined by the absolute magnitude of opposing rates (such as births, which replenishes susceptible and death and infection, which decreases susceptibles) rather than by the rates of change themselves.

The dynamics with this more complete form of scaled noise. Large populations (10 million or 100 million) are effected little by such noise terms, their dynamics is close to that predicted by the deterministic model. In contrast, smaller population sizes (100 thousand) experience proportionally more noise and their behaviour lies further from the deterministic ideal (Keeling & Grenfell 1999; Rohani, Keeling & Grenfell 2002). Once again, in the smaller population size, the presence of noise can be seen to induce cycles (years 15-20) close to the natural epidemic period. A log-log plot of variance against mean number infected provides a concise method of summarising the variability across a range of population sizes. For scaled noise, and large population sizes (with more that 200 infected individuals on average), the variance scales linearly




The dynamics of SIR epidemics ($B = d = 5.5 \times 10^{-5}$, $R_0 = 10$, $g = 1/10$) with scaled noise. The left-hand graph shows how the oscillatory behaviour smaller populations becomes disrupted by noise, whereas large populations conform close to the deterministic ideal ($N = 10^5, 10^6, 10^7, 10^8$). The right-hand graph show the mean-variance relationship for a range of population sizes from 10 thousand to 100 million. The dashed line represents $var = 10 \times mean$, clearly showing how the scaling operates at large population sizes.

with the mean and therefore with the population size also. This linear behaviour is expected for a wide range of stochastic epidemiological (as well as ecological) models, irrespective of the detailed dynamics or the way that stochasticity is introduced (Keeling & Grenfell 1999; Keeling 2000a). Thus while large populations behave more like the deterministic equations, the absolute amount of variation is greater than in small populations. For very small populations, this linear relationship between the mean and variance is destroyed due to the strong nonlinearities that operate when relatively large amplitude epidemics are triggered.

**Comparison of Noise Terms**

Let us contrast the differences between these various models. We do this by noting that $f$ in the additive formulation is equal to $\beta FSI/N$ in the environmental model. Hence, we have four distinct noise scenarios:

1) Plain Additive Noise. $f$ is constant.
2) Scaled Additive Noise. $f = \sqrt{\beta SI/N}$.
3) External Parameter Noise. $F$ is constant, so $f \propto \beta SI/N$
4) Heterogeneous Parameter Noise. $F \propto I^{-\frac{1}{2}}$, so $f \propto \beta S\sqrt{I}/N$ .

Where 4 reflects the fact that different individuals may have different underlying parameters. From this comparison, consider how these noise terms scale as the population size, $N$, changes. Scenario 1 is independent of the population size, scenarios 2 and 4 scale with the square root of the population size, and scenario 3 is proportional to the population size. Thus when the population size is large, the dominant error term is due to externally-driven fluctuations in the parameters, scenario 3. Also, when there is considerable variability in individual's response to infection, the parameter noise due to this factor (scenario 4) can easily exceed the scaled additive noise (scenario 2) due to dynamic variability. However, despite the potential importance of parameter noise, it has received relatively little attention in the epidemiology literature.

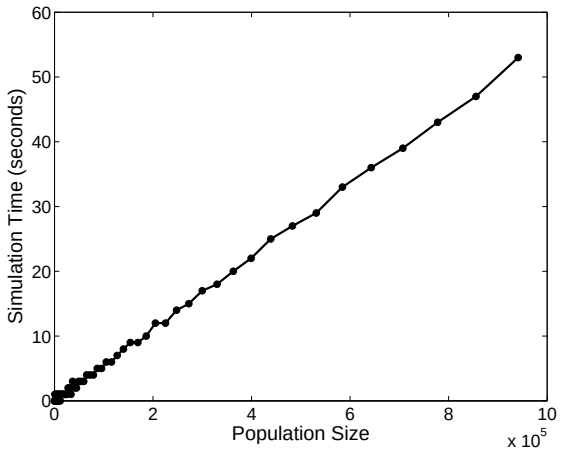
**Demographic**

This type of stochasticity arises because the population is composed of whole individuals, so events occurs at random and make whole number changes.
The following "program" performs demographic iterations.

1) Make a list of all possible events, $E$, and the rate at which they occur, $R_E$.
2) The rate at which any event occurs is $R_{total} = \sum_E R_E$.
3) Pick a random number, $RAND_1$; the time until the next event is $\delta T = -1/R_{total} \log(RAND_1)$
4) Pick another random number, $RAND_2$; find event $i$ such that

$$\sum_{E<i} R_E < RAND_2 < \sum_{E \leq i} R_E$$

5) Perform event $i$, increase the time by $\delta T$, and return to 1.




The time (in seconds) to simulate 1200 years of SIR epidemics ($B = d = 5.5 \times 10^{-5}$, $R_0 = 10$, $g = 1/10$). As the population size increases so does the simulation time, for very large populations this increase could become prohibitive.

This process is then repeated time and time again, as the computer program moves forward one small increment after another. Note that as the population size increases, so the rate at which events occur decreases and the average step size gets smaller – thus (unlike the deterministic models) stochastic models get slower as the population considered gets larger.

The following are a set of common observations for all models with demographic stochasticity.
**Individuals.** The population is composed of whole individuals, so changes must occur in integer multiples. We must therefore think about probabilistic rates and not differential rates. We are best to deal with numbers and not proportions.

**Population Size.** The most obvious factor is that larger populations suffer less from stochasticity than smaller ones. Basically this is because one individual in a small population is worth more than one individual in a large population. Thus as populations become very large (and it's not clear how large this should be) their behaviour approximates the deterministic model.

The effects of population size can be captured by considering the relationship between the mean $M$ and variance $V$ of the number of cases. Almost all populations conform to Taylor's Power Law ($M \propto V^\alpha$) where the exponent $\alpha$ generally lies between 1 and 2. First, it is clear that the mean should increase proportional to the population size. However, the standard deviation relative to the mean decreases as the population size gets bigger, and thus there is relatively less variation about the equilibrium value. (Note, when the population size becomes very large, we

often find that $\alpha \to 1$.)

**Transients.** The behaviour of stochastic models is a trade-off between two competing forces, the deterministic part pulls the orbits back towards the equilibrium (or any other attractor) while the stochastic part pushes trajectories away. Thus how the transient dynamics of trajectories return to the equilibrium dominates the dynamics. Stochasticity (of any form) is known to excite oscillations at the natural frequency of the epidemic.

Stochastic dynamics often cause resonance at (or close to) the natural frequency. This is because while deviations away from the fixed point are random, the return to the fixed point is dominated by the deterministic behaviour. We are seeing a biased random walk back to the equilibrium. (If we think about the dynamics in $r - \theta$ space about the equilibrium, then $r$ is subject to a constant amount of stochasticity, while *theta* becomes more deterministic as $r$ increases. Thus large deviations return in a deterministic manner.)

**Extinctions.** In general we are interested in situations where $R_0 > 1$. However, even under these conditions chance events can drive a population to extinction. This leads to the idea of a *critical community size* as the smallest population that does not suffer from frequent stochastic extinctions. This is obviously very important when considering eradication by vaccination.

We note that the *critical community size* (CCS) must depend on the disease dynamics, population demography and level of imports of infection from outside the population. Despite this, numerous studies of case reports of measles have shown that the CCS is consistently around 300,000 – this figure has been derived independently for communities in the UK, communities in the USA and isolated islands. Standard stochastic models generally over-estimate the CCS, predicting many more extinctions than are actually observed.

**Failure to Invade.** The same reasoning also holds for invading diseases; there is a chance that an infected individual will recover before passing on the disease to any secondary cases. We can make this argument more mathematically explicit, using what is known as a branching process. Suppose that one infectious individual arrives in a totally susceptible population and let $P_{ext}$ be the probability that the infection goes extinct before it causes a major epidemic. Initially one of two events can happen, either the infectious individual recovers (rate $g$) or it causes a secondary case (rate $\beta S/N = \beta$). If the event is recovery, then extinction is guaranteed, otherwise we need to consider the probability of extinction given that two individuals are now infectious. This probability is $P_{ext}^2$, as it requires the lineages from both infections to go extinct independently. Thus:

$$\begin{aligned} P_{ext} &= \frac{g}{\beta + g} \times 1 + \frac{\beta}{\beta + g} \times P_{ext}^2 \\ \Rightarrow P_{ext} &= \frac{g}{\beta} = \frac{1}{R_0} \end{aligned} \tag{6}$$

Thus, not only does a high $R_0$ mean that an invading disease is difficult to control, it also means that there is a low probability of chance extinction. Although this form of calculation works well for invading diseases when the growth is exponential, the only way to assess the risk of extinction for diseases that are currently endemic is through repeated simulation of the stochastic model. Similar reasoning and calculations show that if the outbreak starts with $n$ infected individuals in a population with some immunity then:

$$P_{ext} = \frac{1}{R^n}$$

where $R = R_0 S/N$ is the effective reproductive ratio. Finally if the infectious period is a constant interval (rather than the standard assumption that individuals recover at random), then a similar argument reveals that in a totally susceptible environment the risk of extinction is:

$$P_{ext} \;=\; 1 - R_\infty$$

where $R_\infty$ is the final size of epidemic as predicted by the deterministic equations.

**Deterministic Approximations**

One approach which can provide an analytical solution to stochastic problems is moment closure. Here it is illustrated for the simplest example, the SIS model. We consider the expected value $\langle . \rangle$ for the number of cases.

$$\begin{aligned}
\frac{d}{dt}\langle I \rangle &= \langle \beta SI - gI \rangle \\
\frac{d}{dt}\overline{I} &= \langle \beta(N-I)I - gI \rangle \\
&= \left\langle \beta NI - \beta I^2 - gI \right\rangle \\
&= \beta N\overline{I} - \beta(\overline{I}^2 + var(I)) - g\overline{I} \\
&= \beta(N-I)I - gI - \beta var(I).
\end{aligned}$$

Hence, the expected level of infection is reduced to the stochastic variability of the system. We could approximate the variance in terms of the mean, but its more elegant to refine the approximation by trying to calculate the variance. We do this by looking at changes to $\left\langle I^2 \right\rangle$; this is done by considering the effects of both events and the rates at which they occur.

$$\begin{aligned}
\frac{d}{dt}\left\langle I^2 \right\rangle &= \langle \beta SI(2I+1) + gI(-2I+1) \rangle \\
\frac{d}{dt}(\overline{I}^2 + V) &= \left\langle \beta(2NI^2 + NI - 2I^3 - I^2) - g(2I^2 - I) \right\rangle \\
2\overline{I}\frac{d\overline{I}}{dt} + \frac{dV}{dt} &= 2\beta N(\overline{I}^2 + V) + \beta N\overline{I} - 2\beta(\overline{I}^3 + 3\overline{I}V + T) - 2g(\overline{I}^2 + V) + g\overline{I} \\
\frac{dV}{dt} &= 2\beta NV + \beta N\overline{I} - 4\beta\overline{I}V - 2gV + g\overline{I} + (T).
\end{aligned}$$

We can now make the approximation that the third order terms are zero ($T = 0$), and solve for $\overline{I}$ and $V$.

A second approach that works extremely well for the SIS model and when the population size is small is to use Master equations. Master equations (also known as Ensemble or Kolmogorov-forward equations) are the integer-valued and event-driven equivalent of the Fokker-Plank equations described above. Essentially they require the formulation of a separate differential equation for the probability of finding the population in every possible state. For example $P_{S,I}(t)$ is the probability of having $S$ susceptibles and $I$ infected, where $S$ and $I$ are both integers, at time $t$. Events, such as birth, infection or recovery, move populations between various states and this is reflected in the equations for the probabilities.

Again, because of its simplicity and the fact that only two events can occur, we initially focus on the SIS equation. We let $P_I(t)$ be the probability that $I$ individuals are infectious, noting that there is no need to explicitly keep track of the susceptibles. It is often conceptually easier to think of a large (infinite) number of simulations, and to consider $P_I$ to be the proportion

of simulations that have $I$ infecteds. Four processes can occur that modify the proportion of simulations in state $I$.
1) A simulation in state $I$ can have an infected individual recover at rate $gI$, such that there are now only $I-1$ infected.
2) A simulation in state $I$ can have an susceptible become infected at rate $\beta SI/N = \beta(N-I)I/N$, such that there are now $I+1$ infected.
3) A simulation in state $I+1$ can have an infected individual recover at rate $g(I+1)$, such that $I$ infected remain.
4) A simulation in state $I-1$ can have an susceptible become infected at rate $\beta(S+1)(I-1)/N = \beta(N-I+1)(I-1)/N$, such that there are now $I$ infected.

Processes 1 and 2 cause the loss of a simulation in state $I$, whereas processes 3 and 4 are associated with the gain of a state $I$. Formulating an explicit equation for these processes we obtain:

$$\frac{dP_I}{dt} = -P_I[gI] - P_I[\beta(N-I)I/N] + P_{I+1}[g(I+1)] + P_{I-1}[\beta(N-I+1)(I-1)/N] \quad (7)$$

This generates $N+1$ differential equations ($I = 0 \ldots N$), each of which is coupled to the two nearest probabilities. However, **these equations are always linear.**

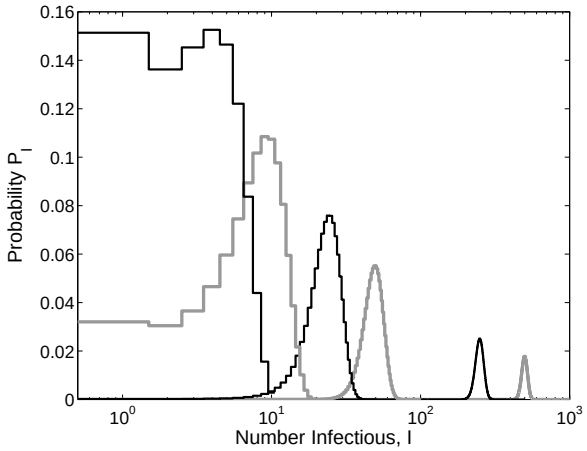
As well as being able to iterate the probabilities forward from any particular condition, one of the great benefits of these models is in understanding how stochasticity affects the final distribution of disease prevalence and the expected variation. However, a significant difficulty exists with this model, which also occurs for many models that use event-driven stochasticity – the fact that extinction events are permanent. Let us consider the dynamics of $P_0$, noting that $P_{-1}$ is zero by definition:

$$\frac{dP_0}{dt} = -P_0[g0] - P_0[\beta(N)0/N] + P_1[g(1)] = gP_1$$

Hence, there is no escape from extinction and eventually every simulation falls into this absorbing state although this may take thousands if not millions of years. So, the modeller is often trapped between the mathematical certainty of extinction and the epidemiological observation that extinctions are rare. Two solutions exist, either to modify the equations such that a low level of infectious imports arrive in the population or to ignore those simulations that have gone extinct and find the distribution of the disease conditional on the fact that it extant (not extinct). This latter method is often preferable as it retains the simplicity of the standard equations.

If we are interested in the final distribution of cases then, rather than having to iterate all the equations forward, it is often far simpler to solve for when $\frac{dP_I}{dt} = 0$. For this one-dimensional SIS model, the equilibrium solution can be found by matching the proportion of simulations moving from $I$ to $I+1$ with the proportion moving from $I+1$ to $I$. If the model is at equilibrium then these two "movements" must be equal or else some of the $P_I$ would be changing. This observation leads to iterative equations for the equilibrium distribution:

$$\begin{aligned} P^*_{I+1}[g(I+1)] &= P^*_I[\beta(N-I)I/N] \qquad \text{such that } \sum_{I=1}^{N} P^*_I = 1 \\ \Rightarrow \qquad P^*_{I+1} &= P^*_1 \prod_{J=1}^{I} \frac{\beta(N-J)J}{Ng(J+1)} \\ P^*_I &= P^*_1 \frac{(N-1)!}{(N-I)!I} \left(\frac{\beta}{gN}\right)^{I-1} \end{aligned} \quad (8)$$




Final equilibrium probability $P_I^*$, from the master equations for the SIS model. $R_0 = 2$, $N = 10, 20, 50, 100, 500$ & $1000$. Note that the number of infection individuals is plotted on a log scale. The integrals beneath all the curves are equal to one, even though the log scale makes the large epidemics appear to occupy a smaller area.

Hence, from either the initial iterative equation or the final explicit form, the quasi-equilibrium distribution of infection (conditional on non-extinction) can be calculated. Such conditional distributions are shown in the figure from which several key factors emerge. For large population sizes ($N$ greater than 500), the conditional equilibrium distribution is approximately Gaussian. This is because over the range of likely values the deterministic dynamics are approximately linear and the integer-valued nature of the population is largely irrelevant. (The smaller values of $P_I^*$ for these population sizes is because the distribution is spread over a wider range of $I$ values). For smaller population sizes the distributions are much more skewed as the non-linear behaviour further from the fixed point plays a more dominant role. For the very smallest population sizes, the distribution becomes bimodal with a peak at $I = 1$ as well as near the expected mean.

This approach can also be used to calculate the extinction rate conditional of the disease still being present. Extinctions of infection can only occur for one particular situation, that is when there is only one infected individual and it recovers. Thus the rate of extinction is simply $gP_1$. Using the fact that all the terms in the distribution sum to one, equation (8) gives us:

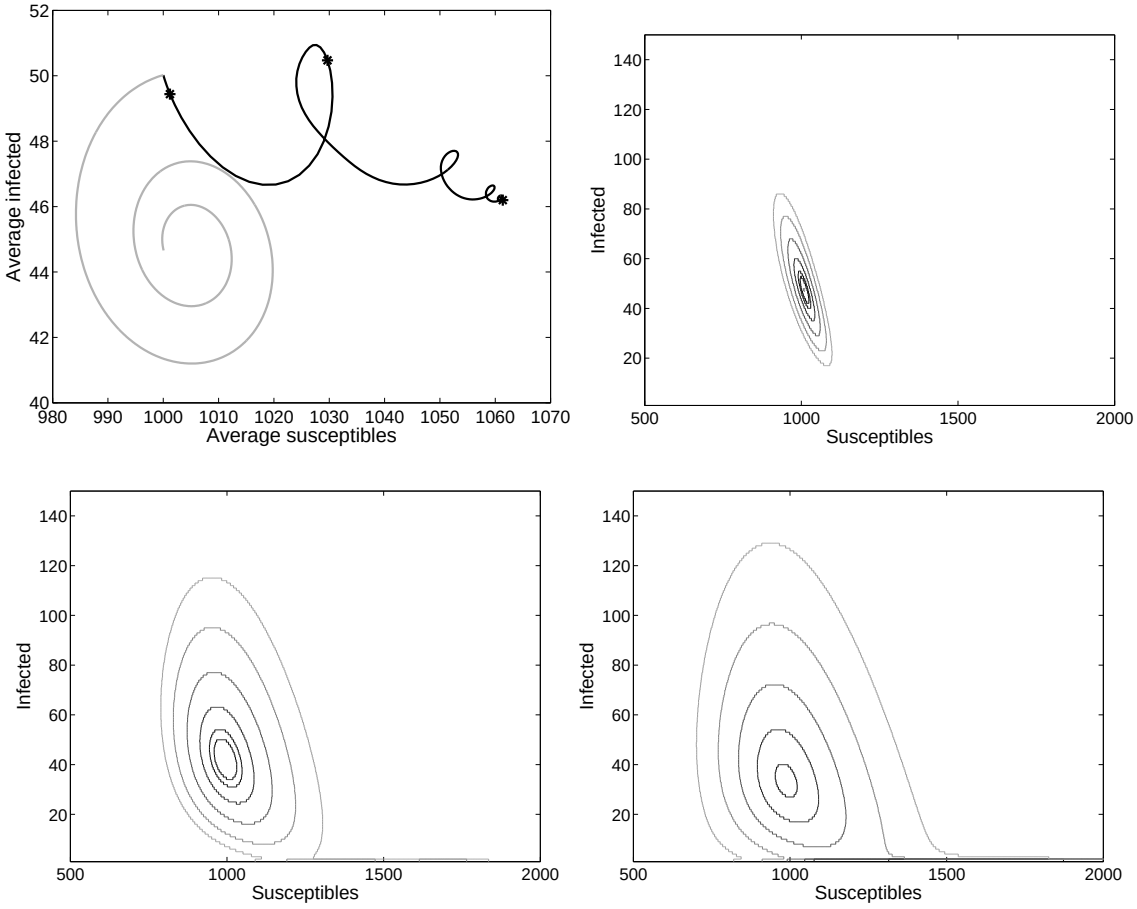
$$\text{Extinction rate} \;=\; g\left[\sum_{I=1}^{N}\frac{(N-1)!}{(N-I)!I}R_0^{I-1}\right]^{-1}$$

So, at least in principle, the long-term asymptotic extinction rate (condition on the disease currently being present) can be calculated analytically.

This mechanism can be extended to examine the stochastic dynamics of the SIR equation, this is conceptually straightforward, but algebraically awkward. The probability is now two-dimensional, depending on both the number of infected and susceptible individuals. In a population of $N$ individuals, whereas the master-equations for the SIS model have $N+1$ equations, for the SIR model there are $\frac{1}{2}(N+1)(N+2)$ equations and so iterating the equations is far more computationally intensive. The equations are also slightly more complex to formulate as there are now six possible events (infection, recovery, birth, death of susceptible, death of infected and death of recovered).

$$\begin{aligned}\frac{dP_{S,I}}{dt} \;=\; & -[\beta SI/N + gI + BN - dS - dI]P_{S,I} \;+\; [\beta(S+1)(I-1)/N]P_{S+1,I-1} \\ & +[gI]P_{S,I+1} \;+\; [BN]P_{S-1,I} \;+\; [d(S+1)]P_{S+1,I} \;+\; [d(I+1)]P_{S,I+1}\end{aligned}$$

Results from this equation with a population size of $N = 10000$ are shown in figure. There are clearly many differences between the average dynamics of the master equation and the behaviour of the standard deterministic equations. The master equation has a higher level of susceptibles at equilibrium as expected due to the negative correlations that develop between the numbers of susceptibles and infecteds. Interesting, due to the higher-dimensionality of the master equations, the trajectories of the average quantities can cross – this is not observed for the standard model. Looking at the distributions in more detail, the first snap-shot (graph B) shows the initial development of negative correlations such that high level of infection are associated with low levels of susceptibles. At later times (graphs C and D), there is a much wider distribution of values and the effects of disease extinction can be observed leading to high levels of susceptibles with few or no infecteds.




Dynamics of the SIR master-equation ($R_0 = 10$, $g = 0.1$, $B = d = 5 \times 10^{-4}$, $N = 10^4$, imports of infection occur at the rate of two per year, $\varepsilon \approx 0.055$). Note that the values of $B$ and $d$ lead to a much faster demographic turn-over than we have previously assumed, the average life expectancy is around five and a half years. This has been done to increase the level of infection and hence improve disease persistence. Graph A shows the dynamics of the average number susceptible and infected predicted by the master equations (black) and predicted by standard differential equations (grey). Graphs B to D show the predicted distribution of infected and susceptibles at three times (indicated with stars on graph A). The contours correspond to 95, 90, 75, 50, 25 and 10% confidence intervals. The master equation was initialised with $P_{1000,50} = 1$ and all other terms zero. The equations were simulated within the rectangle $500 \leq S \leq 3000$, $0 \leq I \leq 500$, and hence contain one and a quarter million individual equations.

Such master equations are clearly a powerful but computationally intensive tool for understanding the dynamics of stochastic disease models. In many respects they should act as the template against which other modelling strategies are judged. With the ever-increasing power of computers, the simulation of such large sets of differential equations will become increasingly feasible.

# Spatial Heterogeneity

Spatial heterogeneity is the new boom area in mathematical-biology. The issues involved are simple to explain: how do we cope with populations that are not randomly mixing and have some form of spatial structure. This spatial structure can come in multiple forms, from community level, to local house-hold structure. Here we will review some basic types of models and consider there uses.

**Metapopulations**
Metapopulation are frequently used in ecology and epidemiology to stabilise deterministic systems or to prevent the extinction of stochastic ones. The idea being that sub-populations driven to extinction would have time to recover before being re-colonised from other sub-populations. Two ecological systems have seen the majority of applied research using metapopulation techniques,

• Infectious diseases. If we assume that there is a large amount of mixing within each town or city, but only limited movement between them, then diseases conform to the metapopulation ideal - with each community being represented by a sub-population.

• Butterflies. In general many species of butterfly occupy isolated, specialised habitats with only limited movement between distant sites - again conforming to the metapopulation ideal. This coupled with the fact that many butterfly populations are under decline has lead to much research into understanding and predicting their stochastic behaviour.

◦ Observation of the Checkerspot butterfly on Jasper Ridge in California show that 97% of individuals remain in their natal habitat.

◦ Habitats with butterfly populations number in the hundreds, generally persist for around 3 to 30 years. Smaller population sizes are rarely observed, whereas populations of over a thousand usually persist for much longer than any field studies.

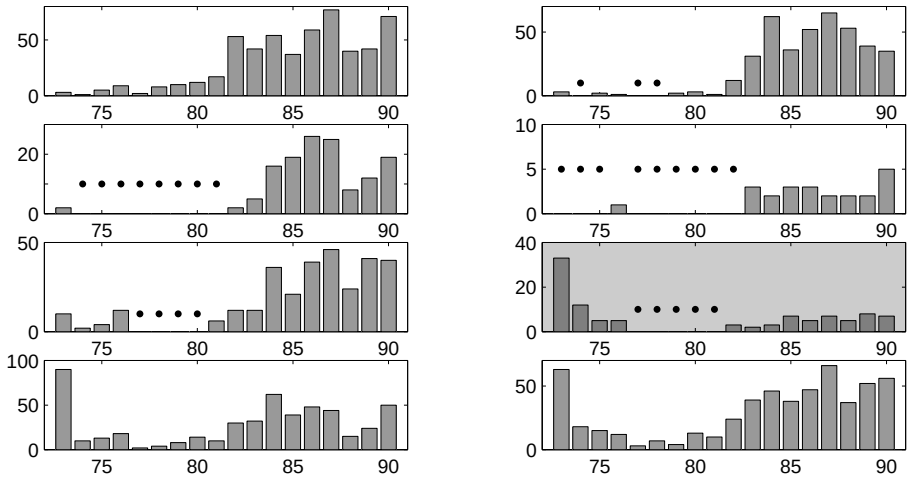
◦ The stochastic effects which lead to the extinction of small subpopulations can be enhanced by the gregarious nature of the caterpillars of many species, which tend to aggregate on just a few host plants. Hence we either see a thriving population or extinction, which corresponds well to the Levins metapopulation model.

◦ As well as extinctions due to the intrinsic stochastic nature of the system, extinctions may also be attributable to external effects such as vegetative dynamics or environmental effects. Effects such as climate may act to synchronise populations over very large areas; Hanski and Woiwood (1993) observed synchronized fluctuations over areas as large as $10^5$ km$^2$. It is due to climatic fluctuations that we frequently observe the greatest persistence for species which occupy a diverse range of habitats. These synchronised extinctions go against the Levins ideal.

◦ Colonisation of vacant habitats has been observed as a linear wave, moving through the habitats with a constant speed. This would indicate that dispersal is a very localised phenomenon.

Two extremes of model are frequently used. Both models describe each habitat as either having a species present or absent, with no consideration given to the internal dynamics of the population.

**Levins Metapopulations:** implicit discrete space, presence-absence models.
*Levins Metapopulation.* All subpopulations are equal.
One of the simplest models of metapopulation extinction and recolonisation comes from Levins (1969). Populations randomly become extinct with probability $e$, and extinct habitats are

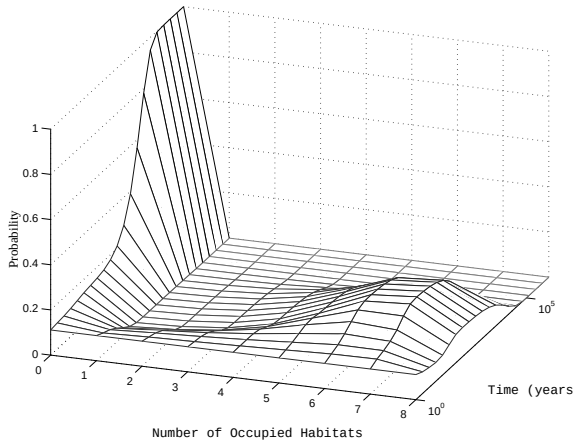



The annual number of Ringlet Butterflies *Aphantopus hyperantus* in 8 subpopulation, from 1973 to 1980. Dots indicate an absence of butterflies from that patch. The data is from Sutcliffe *et al* 1997.

recolonised at a rate proportional to the number of occupied habitats.

$$\begin{aligned} \frac{dE}{dt} &= e(1-E) + rE(1-E) \\ \Rightarrow \qquad E &\rightarrow \frac{e}{r} = \frac{\text{Extinction rate}}{\text{Recolonisation rate}} \end{aligned}$$

Although in many cases this gives a reasonable approximation to the proportion of extinct subpopulations $E$ (Glanville fritillary - Hanski *et al* 1995), the parameters $e$ and $r$ cannot be determined from the population dynamics, but only from long term observations. From the above extinction data, we would estimate $e \approx 0.043$ and $r \approx 0.216$. Using these parameters, we can estimate the long-term behaviour of the entire metapopulation.




The probabilistic distribution of occupied sites. Notice that the time axis is on a logarithmic scale. Hence, for this particular system there is a very long transience when 5 or more habitats are occupied, before the inevitable stochastic extinction. The model predicts that, in general, the entire metapopulation will be extinct after a hundred thousand years - which is a very long time scale for conservation issues.

In a similar manner to the calculation of gene extinction due to random genetic drift, Nisbet and Gurney (1982) considered a stochastic version of Levins' model to predict the time to global

extinction $T_G$ from the average time for a local extinction $T_L \approx \frac{1}{e}$,

$$T_G \;=\; T_L \exp\left(\frac{NO^{*2}}{2(1-O^*)}\right)$$

where $N$ is the total number of habitats and $O^*$ is the proportion of habitats occupied at the steady state of the deterministic equations. This predicts that time to extinction increases exponentially as the number of habitats increases.

One potential problem with this method is separating empty (but possible) habitats from unsuitable ones. This means that the researcher needs to have a very good understanding of an organisms requirements and preferences.

*Mainland-Island Metapopulation* Mainland persists and colonises islands.
The probability of an island, area $A$ and a distance $d$ from the mainland being occupied is given by,

$$\begin{aligned}\frac{dO(A,d)}{dt} &= rd^{-k}(1-O) \;-\; eA^{-K}O \\ O &\rightarrow \frac{rd^{-k}}{rd^{-k}+eA^{-K}}\end{aligned}$$

A good example of a mainland-island model is the work of Harrison *et al* (1988) on the Edith Checkerspot butterfly.

Obviously most biological population will fall somewhere between these two extremes. Hanski and Thomas (1994) produced a spatially explicit meta-population model for *Hesperia Comma* the Skipper butterfly, where much of the theory of metapopulation systems was able to predict the qualitative trends observed.

$$\frac{dO_i}{dt} \;=\; r(1-O_i)\sum_{j\neq i} {d_{ij}}^{-k} O_j A_j \;-\; e{A_i}^{-K} O_i$$

Clearly such a complex and specific model can only be solved numerically.

**Patch Models or Continuous Metapopulations:** implicit discrete space, continuous or stochastic populations.
Metapopulation models are by far the simplest modification to the standard SIR equations (or equivalent). In essence they spilt the population into various subpopulations (or patches) and consider the dynamics on each one. Interaction between the patches comes from between-patch coupling, which mimics the effects of movement. The most general form of the metapopulation is as follows:

$$\begin{aligned}\frac{dS_i}{dt} &= B_i - \beta_i S_i\left[\sum_j \sigma_{ji} I_j\right] - d_i S_i \\ \frac{dI_i}{dt} &= \beta_i S_i\left[\sum_j \sigma_{ji} I_j\right] - d_i I_i - g I_i\end{aligned}$$

This highly complex set of equations takes into account different movement levels between the patches, different mixing rates with the patches, as well as different demography. However,

there are far too many parameters to fully explore this type of model. Current work is using census data and movement-to-work information to try to parameterise such a model for the UK, where patches are either counties, districts, wards or even postcodes. We note that for such a model the rate of disease increase depends on the starting patch, hence (as seen in age or risk structured models) we need a matrix approach to calculate $R_0$.

Note that there is a rigorous theory underpinning the choice of coupling. If we assume that individuals move been patches relatively quickly (compared to the infectious period of the disease) then for two patches

$$\sigma_{ii} = \frac{\gamma^2 + (1-\gamma)^2}{N} \qquad \sigma_{ij} = 2\gamma(1-\gamma)$$

where $\gamma$ is the expected proportion of the time that an individual spends away from their home patch. The factor of two in $\sigma_{ij}$ originates because coupling can come from either the movement of susceptibles or the movement of infecteds. The quadratic terms occur due to two individuals with the same home patch meeting in the away patch.Using this sort of formulation the value of $R_0$ is independent of the coupling.

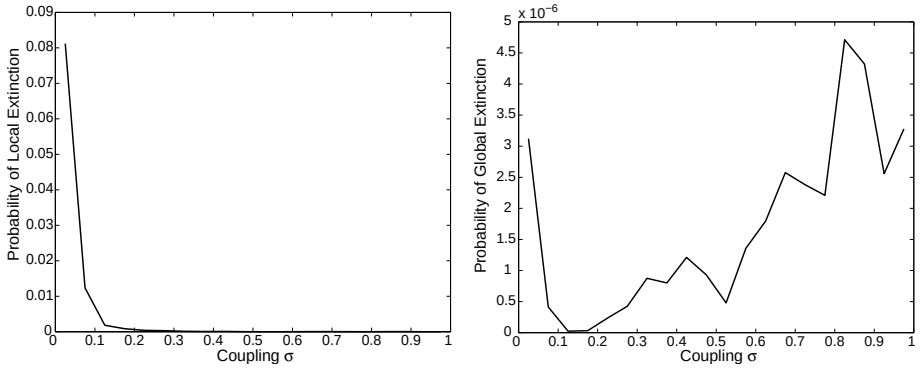
In most cases we look at the simpler metapopulation model where all patches are identical ($B_i = B$, $\beta_i = \beta$ and $d_i = d$), and the coupling is global ($\sigma_{ij} = \mu$, $\sigma_{ii} = 1-(N-1)\mu$). In general, for even weak level of coupling, this global model asymptotes to a uniform solution – which isn't very interesting. To achieve more meaningful dynamics we need to include stochasticity. The random nature of infection and other events acts to push populations apart, while the deterministic attractor and the coupling act to bring populations together. For most populations, the mean number of cases ($M$) increases linearly with the population size ($N$), whereas the variance ($V$) obeys a power-law $V \propto N^{\alpha}$; where $1 < alpha < 2$. Let us now compare the dynamics of one large population, of size $nN$, with the aggregate of $n$ weakly coupled subpopulations of size $N$. The mean number of cases will be equal in each, but the variances will be very different:

$$V_{large} \;=\; kn^{\alpha}N^{\alpha} \;>\; V_{metapop} \;=\; nkN^{\alpha}$$

So a weakly coupled metapopulation will show less variation than a single large patch (or a highly coupled metapopulation).

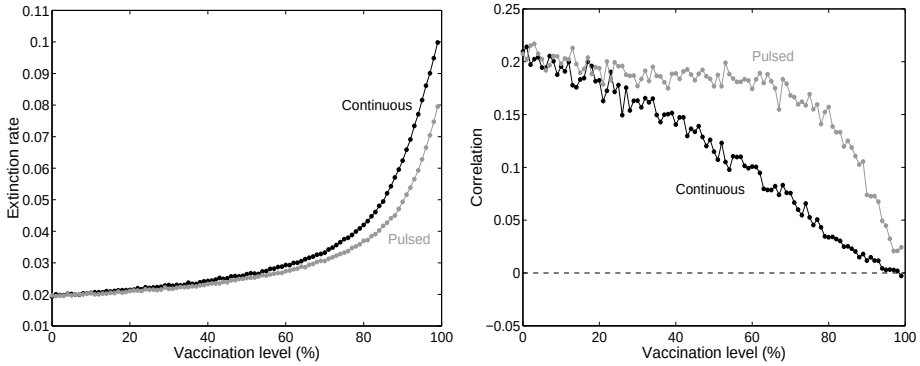
When comparing the extinction rate of a disease across several small populations to the extinction in one large population there are conflicting elements; although small population go extinct more readily we would require them all to be extinct at the same time to globally eradicate the infection, whereas the large population only needs to go extinct once. The effects of coupling between subpopulations on the overall risk of global disease extinction has received much attention due to the interesting trade-offs that arise. If the coupling is very small, then the subpopulations act independently and there is little or no chance of the disease being reintroduced from another subpopulation; there is no rescue effect. If the coupling is very large, then the subpopulations act like one large homogeneously mixed population and thus stochastic effects may force the entire metapopulation extinct.

An extremely applied use of metapopulation modelling for diseases is in understanding the effects of pulse vaccination and synchrony on the persistence of infection. The global persistence of a disease is determined by two main factors, the local extinction rate and the rate of recolonisation – which in turn is related to the heterogeneity of the metapopulation. The figure below shows how these two facets change as the level of vaccination increases; we first consider the solid black line which corresponds to continuous random vaccination. Below the




Probability of local (subpopulation-level) and global (metapopulation-level) extinction for a stochastic SIR model with 20 coupled subpopulations. (Infectious period = 13 days, subpopulation size = 30 thousand, basic reproductive ratio $R_0 = 14$, birth rate = $5.5 \times 10^{-5}$ per person per day).

critical vaccination level of 90% the local extinction rate shows only a moderate increase with the level of vaccination, so that the expected length of an epidemic decreases slowly. In contrast, the correlation between two coupled subpopulations starts to decrease from the onset of vaccination.




The effects of vaccination on the characteristics of unforced SIR epidemics. Black symbols refer to constant random vaccination at birth, whereas grey symbols correspond to pulse-vaccinating randomly at regular four year intervals – similar results are achieved for more frequent yearly pulses. (A) The change in the extinction rate (per day) of an isolated population. (B) The change in the correlation between two subpopulations that are coupled at a level $\sigma = 0.01$. (Population size = ten thousand, $R_0 = 10$, $g = 10$ days, import rate = 5 per year.)

The balance between vaccination increasing the stochastic extinction rate but reducing the synchrony between populations depends on the demographic and epidemiological parameters. Thus while moderate levels of vaccination will always act to reduce the total number of cases they may surprisingly increase the global persistence of the disease if the loss of synchrony is dramatic enough. However, as the level of vaccination approaches the critical eradication threshold the rapid rise in the rate of local extinctions will overwhelm any rescue effects and global extinction will inevitably follow.

Obviously, vaccination would be a much more effective tool if as well as reducing the number of cases it could also decrease the global persistence of the disease. Pulse vaccination offers this possibility, shown in grey on the figure. The first observation is that pulse vaccination is associated with a slightly lower local extinction rate, and also more cases of the disease; this is because in the gaps between the vaccination pulses children that would have been immunised under continuous vaccination have a chance of catching the infection. However, this could be compensated for by only vaccinating susceptible children and therefore not wasting vaccine. In contrast, for the correlation between populations the difference between pulsed and continuous vaccination is more dramatic. The significant perturbation caused by a periodic vaccination campaign acts to synchronise the dynamics of the two populations – thus for pulsed vaccination the correlation remains approximately constant for vaccination levels below 60%.

**Partial Differential Equations:** continuous space, continuous populations
A more mathematically tractable way of examining spatial heterogeneity is through the use of PDE models. A standard SIR type equation can be made spatial by adding diffusion to the system:

$$\begin{aligned}\frac{\partial S(x,t)}{\partial t} &= B - \beta S(x,t)I(x,t) - dS(x,t) + D\frac{\partial^2}{\partial x^2}S(x,t)\\ \frac{\partial I(x,t)}{\partial t} &= \beta S(x,t)I(x,t) - dI(x,t) - gI(x,t) + D\frac{\partial^2}{\partial x^2}I(x,t)\end{aligned}$$

This type of diffusion model can lead to some interesting dynamics, and can be used to predict wave-speeds of infection. This was used to great effect in understanding the historical dynamics of bubonic plague, and accurately predicted the wave-speed from individual-level assumptions about movement and mixing, and the spread of rabies among foxes.

Travelling wave solutions are obtained by looking for solutions of the form, $I(x,t) = I(x-ct)$ and $S(x,t) = S(x-ct)$. Ignoring births and deaths, the wave speed turns out to be

$$V = 2\sqrt{\beta S_0 D - gD}$$

This tells us that the wave speed increases with both the diffusive nature of the host and the basic reproductive ratio $R_0$ of the disease.

A second form of continuous-space continuous-population density model is used when individuals are sedentary and spread of infection is airborne (eg many plant diseases). Here the equations are intro-differential in nature:

$$\begin{aligned}\frac{\partial S(x,t)}{\partial t} &= B - \beta S(x,t)\int K(y)I(x-y,t)dy - dS(x,t)\\ \frac{\partial I(x,t)}{\partial t} &= \beta S(x,t)\int K(y)I(x-y,t)dy - dI(x,t) - gI(x,t)\end{aligned}$$

Here the infection kernel $K$ plays a similar role to the coupling parameters in the metapopulation models. This again leads to travelling wave solutions; when the kernel is thin-tailed (has a finite variance) the wave moves with uniform speed, however if the kernel is thick-tails (has an infinite variance) the wave accelerates. For this sort of model it is easy to calculate the initial growth rate and $R_0$:

$$R_0 = \frac{\beta S_0}{g}\int K(y)dy.$$

This assumes that growth is exponential, although eventually if there is an asymptotic wave speed the growth must be linear.

These continuous-space continuous-population density models have a major problem associated with them. Extremely low levels of infection occur at vast distances from the source, and these can trigger out-breaks. Therefore while in the real world wave-speed is primarily determined by whether or not an infection reaches a location, in these models infection reaches *all* locations, which then start the slow growth of an epidemic. Another side-effect of this formalism is that what is predicted as an accelerating wave, is often seen in practice as a front that starts new epidemics a long distance away.

**Coupled Map Lattices:** discrete lattice, continuous populations
Consider a simple map, which takes the population level this year to the population level next year, eg.

$$x_{t+1} \;=\; f(x_t).$$

One of the most natural means of including true Euclidean space (as opposed to the more abstract space of the metapopulation) is the coupled map lattice. First we **discretise** space into a grid or lattice of cells, usually this lattice is a regular 2-D square or a single line. In general the dynamics of each cell are governed by homogeneous, mean-field dynamics within the cell, together with coupling to neighbouring cells,

$$x'_{i,j} \;=\; \left[1 - \sum_{k,l} \sigma(i,j\,;\,k,l)\right] f(x_{i,j}) + \sum_{k,l} \sigma(k,l\,;\,i,j) f(x_{k,l}).$$

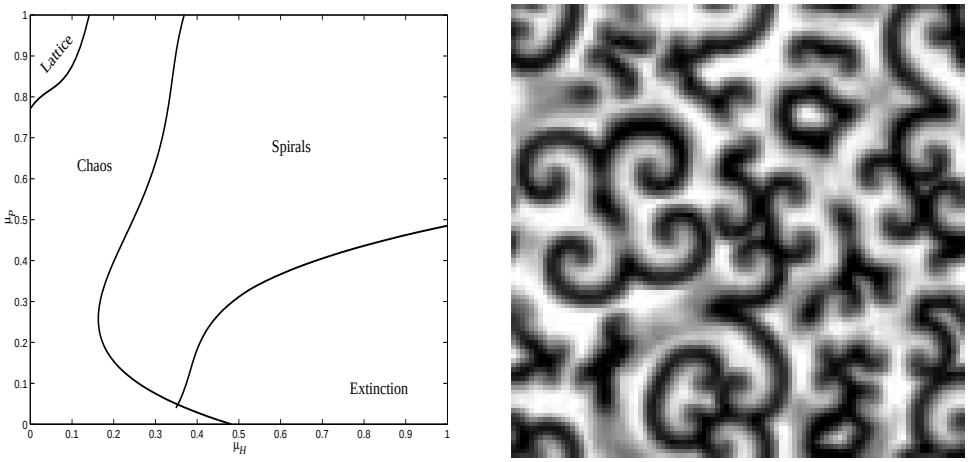
We usually expect the coupling to act in both directions, $\sigma(i,j\,;\,k,l) = \sigma(k,l\,;\,i,j)$, and to be homogeneous, $\sigma(i,j\,;\,i+x,j+y) = \sigma(k,l\,;\,k+x,l+y)$. In this formulation, $\sigma \in [0,1]$ gives the proportion of individuals moving from one site to another, and hence defines the neighbourhood. For a given site its **neighbourhood** is the set of all sites that interact or couple to it,

$$Nbd_{i,j} \;=\; \{k,l \,|\, \sigma(i,j\,;\,k,l) + \sigma(k,l\,;\,i,j) > 0\}$$

Commonly, the neighbourhood of a site is the adjacent 4 cells (Von Neumann neighbourhood) or the adjacent 8 cells (Moore neighbourhood) - this means that all interactions are local in space. This coupling $\sigma$ can often be thought of as the random movement (or diffusion) of individuals. An intuitive way to visualise a CML is to consider a population of beetles living on fir trees in a plantation. Each tree has its own population dynamics with occasional migration to, and immigration from, neighbouring trees.

The Nicholson Bailey CML
One of the most famous and well-cited examples of coupled map lattices being used in ecology is the work of Hassell, Comins and May (1991). As discussed earlier, the standard Nicholson-Bailey model for the interaction of hosts and parasitoids is unstable, with rapidly expanding oscillations in both populations. By discretising space into a square lattice, and by including coupling between neighbouring cells, Hassell *et al* were able to stabilise the dynamics. As the coupling is varied many different forms of spatial pattern are observed including spiral waves, crystal lattice and seemingly chaotic behaviour.

Advantages and Problems
The main advantage of CML is that they are a relatively straight forward modification to the




Results from the Nicholson-Bailey coupled map lattice (Hassell *et al* 1991). The left-hand figure shows the type of persistence spatial patterns observed for a range of couplings, $\mu_H$ and $\mu_P$. The right-hand figure is a snap-shot of the spatial lattice giving the density of parasitoids at each site.

deterministic homogeneous dynamics; they also give us many insights into the role of spatial heterogeneity and allow some degree of analytical rigour (see work on discrete breathers). They are fast to compute and are the simplest way to implement a spatial model.
On the down-side there are some theoretical flaws in their discretisation of space. The CML assumes that there is complete global mixing within a lattice cell but only limited interaction with neighbouring sites. While this assumption is not totally invalid for beetle populations in a coniferous plantation it is not a true model of Euclidean space and great care must be taken over choosing a suitable scale. It is also possible for the regular nature of the grid to impose symmetry breaking (preferred directions) on the spatial dynamics. A further complication is that each site generally obeys a set of deterministic rules, a more natural assumption would be to have stochastic dynamics within each cell, which would give rise to localised extinctions and rescue effects.

**Cellular Automata:** discrete lattice, discrete stochastic populations
As with the coupled map lattice, the Cellular Automata discretise space into a regular lattice of cells. However for this type of model there are only a finite number of states each cell can be in, eg. empty or occupied. Hence in the majority of applications each cell corresponds to a single individual. There is a large literature on the behaviour of deterministic cellular automata (Wolfram 1986), but these are not very applicable to biological problems. The major of these models therefore usually have a set of probabilistic rules which govern the transitions of each lattice cell, these rules usually just depend on the current state of the cell and the states of surrounding cells. For example the probability of a susceptible organism becoming infectious will depend on the number of infectious organisms in the surrounding neighbourhood.
For CA there are two different types of updating. Synchronous updating can be compared to discrete time maps where every cell is updated at the same time, using the previous values of the neighbours. The alternative, asynchronous updating is more akin to continuous time differential equations, here cells are updated randomly one at a time and although this may be more realistic it is far slower numerically.

The Contact Process
One of the most simple probabilistic cellular automata is the Contact Process. Consider a linear array of cells, which can be in one of two states, lets call them '*on*' and '*off*' - we can equate this model to the spatial spread of disease in which case the states become '*infectious*' and '*suscep-*

*tible*'. *On* sites can spontaneously decay back to the *off* state at a rate $g$; *off* sites switch to *on* at a rate proportional to the number of *on* neighbours (eg $n\tau$ where $n \in \{0, 1, 2\}$ is the number of *on* neighbours). As *on* sites tend to create *on* neighbours, there is a high degree of spatial aggregation in the contact model. These correlations lead to many interesting persistence results (Mollison 1977, Durrett 1980).

The Forest Fire Model

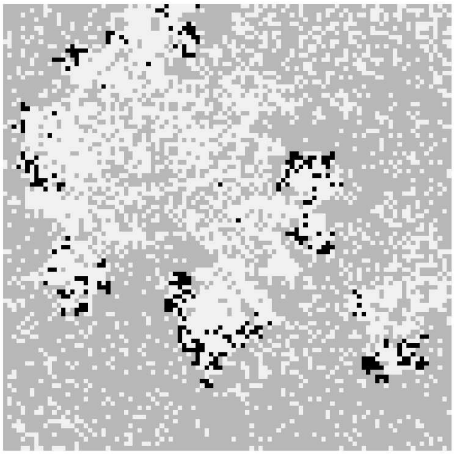
This second type of probabilistic cellular automata can also be regarded as a model of disease transmission. Forest fire models generally assume a 2-D regular lattice of sites. There are now three states, long lived trees (*cf* susceptible), short lived fires (*cf* infectious) and empty sites (*cf* recovered). Trees can catch on fire from burning neighbours, fires randomly decay to empty sites and empty sites are eventually re-colonised by trees. Once again there is a large literature on the persistence and fractal nature of this system, as well as whether it displays Self Organised Criticality (Bak *et al* 1990, Rand *et al* 1995).

A disease model

The simplest disease model would have the following rates of transition:

$$\begin{array}{lll} S \to I & \text{Rate} & = \beta \times \text{number of I in neighbourhood} \\ I \to R & \text{Rate} & = g \\ R \to S & \text{Rate} & = d \end{array}$$

This is clearly related to the forest-fire model. This models a locally spreading epidemic, with very severe competition for space (so that when a recovered dies it is immediately replaced by a new-born susceptible). We now face the tricky question of what is $R_0$ in this situations, and



An example of a cellular automaton model. Cells are colour coded, black for infectious, white for recovered, grey for susceptible.

hence can we predict if an infection will spread or fail. From the very initial growth rate, we expect:

$$R_0 = \frac{\beta}{g} \times \text{number of neighbours}$$

however after this first generation, the each individual must has at least one less susceptible neighbour (who infected them). Hence in these localised spatial models there is a rapid depletion of susceptibles in the immediate vicinity, although the global level has hardly changed. Whether the infection can spread is a question of percolation, and hence we need a high initial

growth rate to maintain the epidemic. The precise levels needed in a general situation is still an open problem in mathematics.

**Network Approaches:** implicit network, continuous populations
The one main draw-back with cellular automaton models is their very rigid structure, each individual has the same number of local neighbours. A network approach is far more flexible and allows individuals to have different numbers of contacts some of which may be distant. For airborne diseases, such as influenza or measles, deciding who is a contact is very difficult though still important. However, for sexually transmitted diseases, its far more clear how to define contacts even if the data is somewhat more sensitive. We look at how pair-wise models can help us understand the dynamics of STDs on networks.

Let us look at the number of infected individuals which we write as $[I]$:

$$\frac{d[I]}{dt} = \tau[SI] - g[I]$$

that is infection occurs at rate $\tau$ across a connected pair of $S$ and $I$. We could now approximate this as $[SI] = n[S][I]/N$, (where $n$ is the average number of partners) – this ignores all spatial structure. Obviously the next stage is to formulate an equation for [SI]:

$$\frac{d[SI]}{dt} = \tau[SSI] + g[II] - \tau[SI] - g[SI] - \tau[ISI]$$

This requires that we know $[II]$, however we know that $[IS] + [II] = n[I]$, but we also need the triples $[SSI]$ and $[ISI]$. In theory we could write down equations for this term, but that becomes really messy. We therefore assume that:

$$\begin{aligned}[SSI] &= \xi\frac{[SS][SI]}{[S]} \\ [ISI] &= \xi\frac{[SI]^2}{[S]}\end{aligned}$$

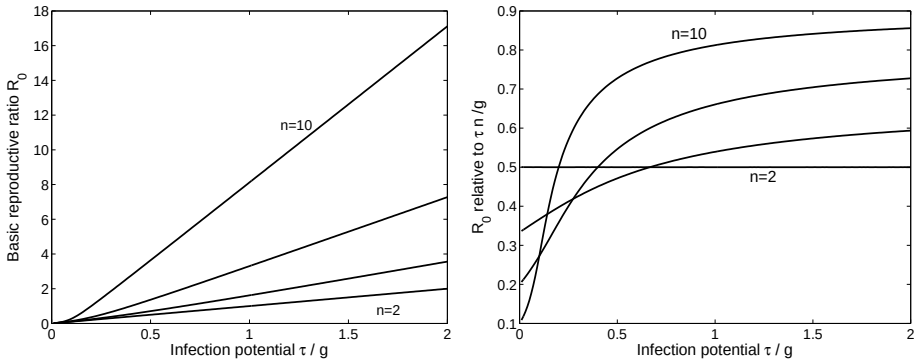
where $\xi$ is one if the number of partners is Poisson distributed, or $1-1/n$ if everyone has exactly $n$ partners. We can solve these equations to give:

$$\begin{aligned}[I] &= \left(1 - \frac{g}{\tau(n-1)}\right)N \\ [SI] &= \frac{g[I]}{\tau}\end{aligned}$$

This shows two things, first the level of infection is less in these models than expected in random mixing models and this is due to the negative correlation that builds up between susceptible and infectious individuals ($[SI] < n[S][I]/N$).

Finally we can use this approach to calculate $R_0$. Note that in the first generation we would expect to cause $\tau n/g$ secondary cases (or from an individual perspective $n\tau/(\tau+g)$). To find $R_0$ we can either use an eigenvalue approach on the governing equations, or we can look for a quasi-equilibrium solution to $[SI]$ which is proportional to $[I]$, when $[I]$ is small.

$$\frac{d}{dt}\frac{[SI]}{[I]} = \frac{1}{[I]}\left\{\tau\left(\frac{[SS][SI]}{[S]}\right) + g(n[I] - [SI]) - \tau[SI] - g[SI] - \tau\left(\frac{[SI]^2}{[S]}\right)\right\} - \frac{[SI]}{[I]^2}\frac{d[I]}{dt}$$




The left hand figure shows the calculated $R_0$, once [SI] is at quasi-equilibrium, for different average numbers of neighbours ($n = 2, 3, 5, 10$). Notice that this is slightly non-linear, especially when $R_0$ is small. The right hand figure shows how much $R_0$ is reduced relative to the first generation reproductive ratio $n\tau/g$.

which has the quasi-equilibrium solution:

$$[SI]^* \;=\; \frac{\tau(n-1) - g + \sqrt{[\tau(n-1) - g]^2 + 4\tau n g}}{2\tau}[I]$$

This allows us to calculate the actually asymptotic growth rate, and hence a sensible $R_0$:

$$R_0 \;=\; \frac{\tau[SI]^*}{g[I]}$$

From the above graph we clearly see the effects of depleting the level of susceptibles. The asymptotic level of $R_0$ is clearly much lower than in the first generation, and this effect is exacerbated if $\tau/g$ or $n$ is small.

**Full Stochastic Simulations:** continuous space, discrete populations
This type of model is by far the most advanced, most biologically realistic and therefore the least mathematically tractable of all the spatial models. For a disease system, the rate of infection of a susceptible individual ($i$) is a function of its susceptibility ($S_i$), the transmission rate from surrounding individuals ($T_j$) and the distance between them ($d_{ij}$):

$$Rate_i \;=\; S_i \sum_{j \in I} T_j K(d_{ij})$$

where $K$ is the transmission kernel and measures how infectivity decreases with distance. This is the basis for the models that were used during the 2001 Foot-and-Mouth epidemic. For this type of transmission, and assuming that identical individuals are uniformly dense in 2D, we can approximate the basic reproductive ratio $R_0$

$$R_0 \;=\; ST \int_{\mathbb{R}^2} \rho K(|\underline{d}|)\, d\underline{d} \;=\; 2\pi ST\rho \int_{\varepsilon}^{\infty} rK(r)\, dr$$

where $\rho$ is the density. The $\varepsilon$ term is there to take account of spacing of individuals; a susceptible cannot be on top of an infected. In general, $K$ is either assumed to be *normal*, *exponential* or most often *power-law*. For power-laws, $K(d) \propto d^{-\alpha}$, it is important that $\alpha$ is greater than 2 if the reproductive ratio is to be bounded.

Further work by Mollison (1977) has shown that a second quantity:

$$ST \int_{\mathbb{R}^2} \rho |\underline{d}| K(|\underline{d}|)\, d\underline{d} \;=\; 2\pi ST\rho \int_{\varepsilon}^{\infty} r^2 K(r)\, dr$$

has an important role to play in the spatial dynamics. If this value is finite then the disease asymptotes to a travelling wave of fixed velocity; whereas if the value is infinite then the infection keeps accelerating in a series of long jumps forward followed by back-filling. Hence in the power-law case, we get a constant wave speed if $\alpha > 3$.

# Control

One of the main practical achievements of mathematical epidemiology has been in the understanding and prediction of the optimal and appropriate methods of controlling a variety of diseases. With epidemic control time is usually the critical factor, early interventions are key to rapid and easy eradication. This means that models play an important role as it is too time consuming, as well as costly (in both public-health and economic terms) to experiment with the real system.

We will start with a basic review of mass-vaccination, before discussing the many techniques used to target control more effectively.

**Mass Vaccination**

Mass prophylactic vaccination before the arrival of an epidemic, or continual vaccination again endemic diseases is by far the most common mechanism of disease control. The critical level of vaccination needed is given by :

$$V_C \;=\; 1 - \frac{1}{R_0}$$

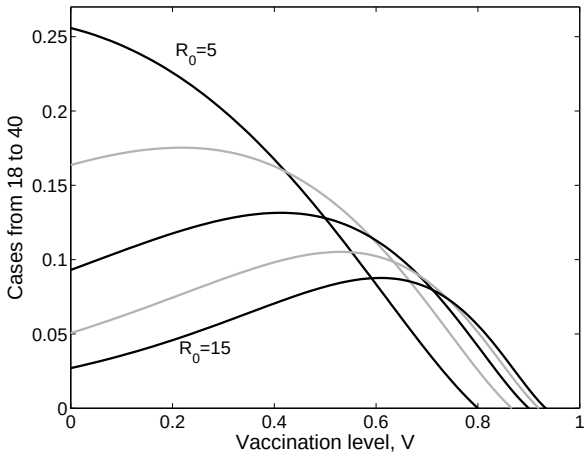
thus, not only due diseases like measles spread rapidly and invade easily, they are also difficult to control.

An interesting use of modelling with vaccination is the control of Rubella. Generally Rubella ($R_0 \approx 10$) is a harmless infection with SIR type dynamics, however if it is caught by pregnant women it can have severe consequences for the child. We therefore want a vaccination scheme that reduces the number of Rubella cases in the 18-40 year age group. Hence, using the standard age structured formalism and assuming a fraction $V$ of the population are vaccinated at birth:

$$S(a) \;=\; (1-V)\exp(-d[(1-V)R_0 - 1]a)$$

If we ignore deaths, the number of cases between 18 and 40 is equal to the drop in the level of susceptibles over this age range:

$$\text{Total Cases} \;=\; S(18) - S(40)$$




The number of cases in the 18-40 age group, as the level of vaccination is varied. Although high levels of vaccination are clearly optimal as they eradicate the infection, intermediate levels may lead to an increase in infection by increasing the average age of infection.

The results of this calculation are shown in the figure. When the value of $R_0$ is moderately large, intermediate levels of vaccination may have an adverse effect and increase the number of cases in the at-risk age groups. This occurs because although there are less susceptibles in the population, there are also less infectious individuals; this increases the average age of infection. Without vaccination, the high level of infectious cases meant that most women got infected before they were 18. Interesting, this shows that just vaccinating girls is not an optimal policy ($V = 0.5$) as (i) this does not eradicate the epidemic and (ii) any women that weren't protected may be put at greater risk.

Another practical issue that we have so far neglected is the fact the vaccines are rarely 100% effective. In practice, not all individuals that are vaccinated will go on to develop immunity. If we suppose that only a proportion $p$ develop immunity then in general we can compensate for this by vaccinating a higher fraction of the population:

$$V_C \;=\; \frac{1}{p}\left[1 - \frac{1}{R_0}\right]$$

However, we clearly run into difficulties if the proportion that are protected is so low that the vaccination threshold is not achievable.

In some circumstances the immunity generated by vaccination is not life-long (eg. vaccination against foot-and-mouth only protects cattle for about 6-12 months). In such scenarios regular repeat vaccinations will be required, this significantly increases the logistical effort associated with vaccination.

Mass vaccination *during* an epidemic can be seen as a race between the exponentially increasing disease and the constant logistical effort; there is a critical amount of vaccination effort needed to control the disease, for lower efforts the penalties are severe. In the early stages of an epidemic, constant levels of mass-vaccination will lead to a linear decrease in the proportion of susceptibles. Thus:

$$\begin{aligned} \frac{dI}{dt} &\approx \beta(S_0 - Vt)I \;-\; gI \\ I &\sim \exp\left([R_0 - 1]gt - \frac{1}{2}Vt^2\right) \end{aligned}$$

This race between vaccination and the epidemic is made worse by the delay between vaccination and protection, this can often be as much as two weeks which means that control always lags well behind the infection.

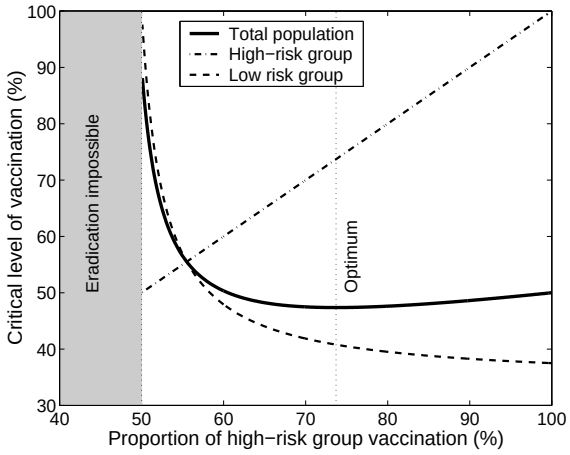
Finally, for some diseases the vaccination itself may have severe side-effects. For example, the vaccine against small-pox causes adverse reactions in individuals with asthma, eczema and those that are immuno-compromised (eg with AIDS) and can even be fatal. In such cases there is a trade-off between under-vaccinating (letting the disease progress) and over-vaccinating (causing many vaccine related problems). Clearly there is an optimal level of vaccination that minimises the public-health consequences:

$$\text{Consequences} \;=\; C_V V \;+\; C_D R_\infty \qquad \text{where} R_\infty \;=\; (1 - V)\left[1 - \exp(-R_0 R_\infty)\right]$$

where $C_V$ and $C_D$ are the consequences of the vaccine and disease respectively. Generally, $C_V \ll C_D$ and so vaccination at (or even above) the critical level is optimal.

**Targeted Vaccination**

One successful way to increase the efficiency of vaccination is to carefully target it towards specific risk group. It is frequently the case that focusing vaccination on the highest risk-groups allows the disease to be eradicated with the minimum of effort. For example, if we want to control a sexually transmitted disease there is little point targeting monogamous individuals.




As seen before, the critical level of vaccination needed to eradicate a STD, as a percentage of the entire population, as the coverage in the high-risk group increases. In this example, the transmission matrix is $\boldsymbol{\beta} = \begin{pmatrix} 10 & 1 \\ 1 & 2 \end{pmatrix}$, with $N_H = 0.2$, $N_L = 0.8$ and $g = 1$ as before.

It is still a very open problem to decide the precise distribution of vaccination effort that should be applied to different at-risk groups, but in general it is better to over-target the high-risk groups rather than under-target.

Another example of targeted vaccination is with age-structure. If the vaccine offers lifelong immunity, then it is alway better to vaccinate early in life. Consider vaccinating against measles, and the MMR vaccine. With the triple vaccine children get protected much early, say six months of age as opposed to twelve. For a given level of vaccination, V, the number of extra cases during the six month delay is approximately:

$$(1-V)[\exp(-[17(1-V)-1]/70) - \exp(-[17(1-V)-1]/140)]$$

So as the vaccination campaign becomes more successful, there is a decreased importance on early vaccination. From an individual point of view it is better for everyone else to vaccinate their children, but if everyone has this selfish attitude the vaccination programme will fail.

**Quarantining**

Quarantine, removing individuals from the active population, can be a very effective control strategy. It effectively works by removing infectious individuals at a faster rate than they recover. Ideally it should be used in conjunction with contact-tracing (see later), such that individuals are quarantined before they become infectious. For a given disease, we can calculate the rate of quarantining $q$ necessary to eradicate the disease:

$$R_0 = \frac{\beta}{g} \qquad R = \frac{\beta}{q} = 1 \;\Rightarrow\; q = R_0 g$$

Thus, $R_0$ determines how much faster we need to remove infection by quarantine compared to natural recovery. In practice, we may also need to take account of the fact that with many

diseases individuals are infectious *before* they are symptomatic.

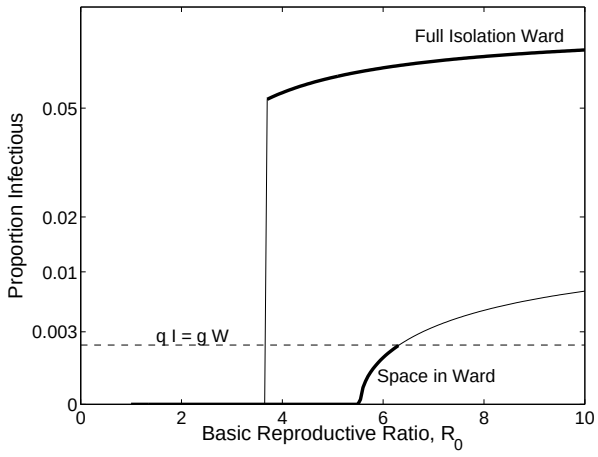
Frequently, public-health authorities have a limited amount of quarantining sites available. Let us suppose that infected individuals can be quarantined at a rate $q$, but that only $W$ places are available on the isolation ward. Two scenarios exist, (a) either the isolation ward is full $Q = W$ and there are excess cases:

$$\begin{aligned} \frac{dS}{dt} &= B - \beta SI - dS \\ \frac{dI}{dt} &= \beta SI - gI - dI - gW \\ \frac{dQ}{dt} &= gW - gQ \end{aligned}$$

or (b) the ward is not full and cases can be immediately quarantined:

$$\begin{aligned} \frac{dS}{dt} &= B - \beta SI - dS \\ \frac{dI}{dt} &= \beta SI - gI - dI - qI \\ \frac{dQ}{dt} &= qI - gQ \end{aligned}$$

The switch between the two occurs when $qI = gW$.




The dynamics of a disease with quarantining. $B = d = 0.01$, $W = 0.01$, $g = 0.1$, $q = 0.5$; this high level of births and deaths mimics movement into and out of hospitals. Two solutions exist corresponding to when the isolation ward is full or not; the thick black lines show where the solution is self-consistent.

We notice from the figure that, for intermediate values of $R_0$, two possible solutions exist. Thus in this region the behaviour can flip from a low level of cases that can be controlled by quarantining to a high volume that cannot, this switch could be triggered by small stochastic fluctuations in the incidence of infection. A reverse switch, from high to low levels of infection, is unlikely to occur due to stochastic effects and would require concerted public-health efforts.

**Contact-Tracing**

Contact tracing is by far the most efficient form of targeted control. In an ideal setting, infected individuals are diagnosed early and all possible contacts (to whom they could have spread the disease) are traced and, isolated or treated. In this case the disease would be eradicated after

the second generation. In practice, tracing is not 100/someone on the train, are completely random. Thus, while contact-tracing is the ultimate targeted control strategy it is also very labour intensive, hence it can only be an effective strategy when there are few cases. For this reason contact-tracing is frequently used for sexually transmitted infections where contacts are obvious, prevalence is low and the disease dynamics are slow. If the number of cases is large, then the tracing capacity becomes swamped and the epidemic goes out of control (cf finite-sized isolation wards).

One method of modelling the effects of contact tracing is to subdivide the infected population into those that can and cannot be traced:

$$\begin{aligned}
\frac{dS}{dt} &= -\beta(I_T + I_U)S \\
\frac{dI_T}{dt} &= \beta p(I_T + I_U)S - gI_T - TI_T \\
\frac{dI_U}{dt} &= \beta(1-p)(I_T + I_U)S - gI_U \\
\frac{dR}{dt} &= gI_U + gI_T + TI_T
\end{aligned}$$

where $p$ is the proportion of infections that are traceable, $T$ is the contact tracing rate and any traced or diagnosed individuals are removed from the general population. For this situation, we can estimate $R_0$:

$$R_0 = p\frac{\beta}{g+T} + (1-p)\frac{\beta}{g}$$

Hence, contact tracing can only stop the epidemic if $(1-p)\beta/g < 1$, that is the untraceables are unable to sustain the epidemic.

While the above model gives us a good insight into the role of tracing, many of the subtleties are lost being amalgamated into one of the estimated parameters. A true model of contact-tracing would need to take into account the network structure of contacts (see Muller *et al.* 2000). When the transmission network is random, so that there is little or no clustering of contacts, then (and probably not too surprisingly) a proportion $1/R_0$ of all contacts must be traced if the disease is to be controlled. However, when there is clustering within the network the critical level of contact-tracing is reduced.

From these results we see that contact-tracing is easier in clustered networks (such as those associated with airborne diseases) as the clustering provides multiple routes of tracing the same person.

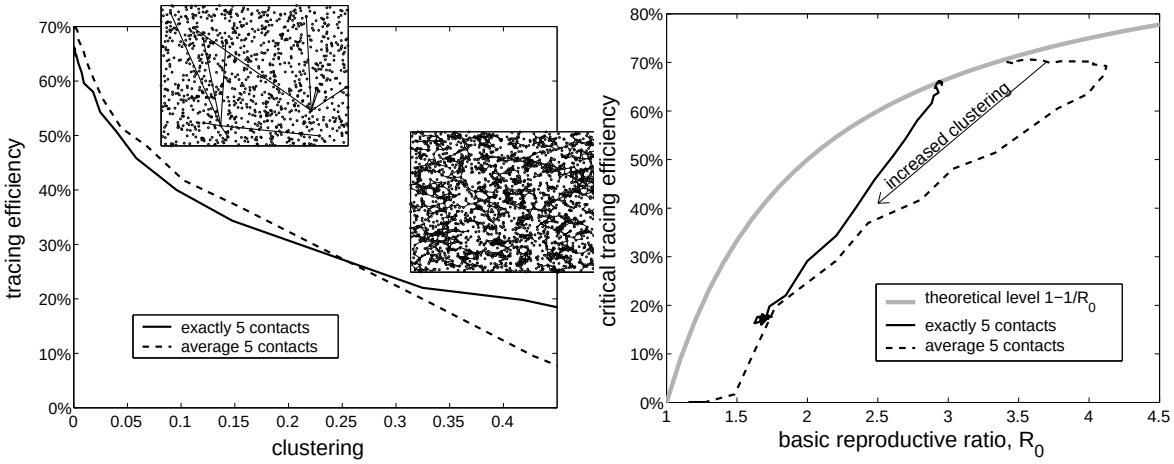
**Ring-Cull or Ring-Vaccination**

Ring vaccination or ring culling are two other forms of highly targeted control measures. Here, spatial proximity to a source of infection is the key risk factor. If we utilise the formulation developed in the previous spatial-modelling lectures, then in the presence of ring control out to a radius $r_c$ the reproductive ratio is reduced to:

$$R = 2\pi STP\rho \int_{r_c}^{\infty} rK(r)\,dr$$

hence if the kernel, $K$, has a power law distribution ($K(r) = Pr^{-\alpha}$), then

$$R = \frac{2\pi STP\rho r_c^{2-\alpha}}{\alpha - 2} \qquad \Rightarrow \quad r_c > \left(\frac{\alpha-2}{2\pi STP\rho}\right)^{\frac{1}{2-\alpha}} \text{ for eradication}$$




The effects of clustering on the level of contact tracing. The left-hand graph shows two example networks in addition to the critical level of contact tracing needed. The right-hand graph shows how clustering causes the critical contact tracing level to drop below the predicted value.

So, whenever the power-law decays faster than $\alpha = 2$, there is a locally targeted control measure which can eliminate the disease. If $\alpha$ is close to 2, then the control radius needed may be huge.

This has interesting analogies to the control of foot-and-mouth during the 2001 epidemic, where CP (contiguous premise) culling was used to slow the spread of the disease. Vaccination was not a reliable option due to the large number of animals involved and the delay between vaccination and protection. Culling is obviously much more extreme than vaccination; from the point of view of a farmer, (s)he is always better to take a chance that they might not get infected than to loose their animals to culling. However, from a local perspective it is optimal to cull a large number of at risk farms. Finally from a national perspective it is optimal to cull all at risk farms to prevent the disease spreading into new regions. It is this change in optimal behaviour at different scales that causes all the arguments and difficulties.

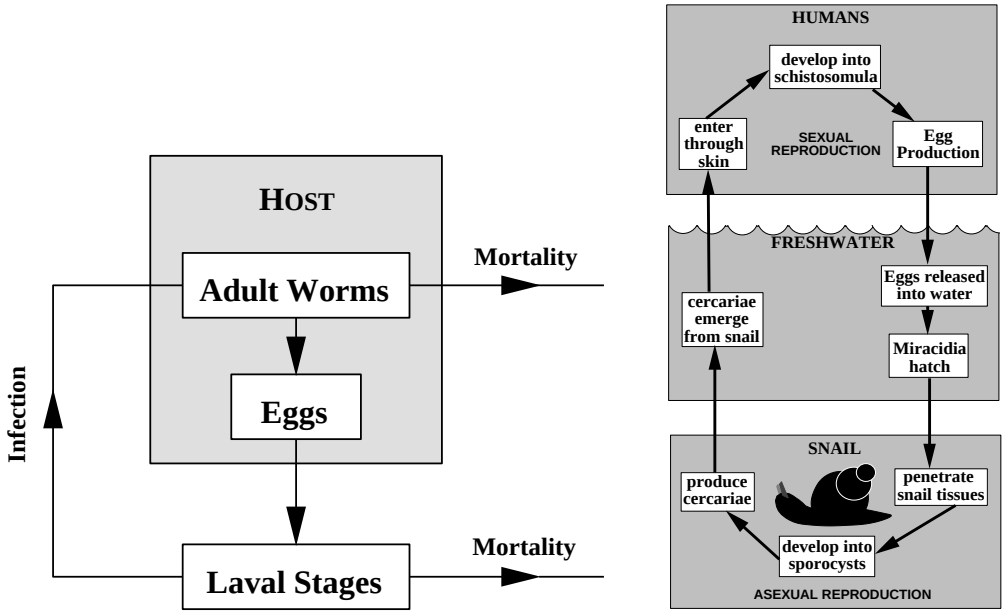
Whereas the study of micro-parasites concentrates on bacterial and viral diseases, the study of macro-parasites considers infection by larger organisms, such as **protozoa**, **helminths** and **arthropods**.

**Protozoa** are single celled organisms including **Amoebae**, **Flagellates** (diseases include Sleeping Sickness), **Cilliates** and **Sporozoa** (diseases include Malaria). As these simple creatures reproduce very quickly within the host, the mathematical models which describe these infections are comparable with standard micro-parasite disease models.

**Arthropods** Most arthropods which attack humans are blood feeders, examples include mosquitoes, ticks and fleas which can act as **vectors** for other infectious agents, transporting the disease from one host to the next.

**Helminths** are parasitic worms, including tapeworms, flukes and roundworms. In this lecture we shall concentrate on the study of this type of macro-parasite.

What differentiates the study of helminth infections from standard disease models is the variability in **worm burden** and the complexity of the parasite life cycle. Although for most micro-parasitic diseases hosts can either be classified as susceptible, infected or recovered, for helminth infections we must also track the number of worms (or burden) within each host. Many helminths spend some part of their reproductive cycle out side the host, often inside a host of a different species.




The left-hand figure shows a generalisation of the life-cycle of macro-parasitic worms. The right-hand figure is the life-cycle of **schistosomes** or blood flukes, which involves reproduction within humans and snails as well as free-swimming stages. Schistosomiasis causes extensive damage to the blood vessels surrounding the lungs and liver.

For macro-parasite systems we need to modify the definition of the basic reproductive ratio $R_0$,

$$R_0 = \begin{array}{l}\text{average number of offspring produced by a mature parasite that themselves}\\ \text{survive to reproductive maturity in the absence of density-dependent effect.}\end{array}$$

That is, $R_0$ is the number of offspring expected to complete a full reproductive cycle, when there is no intraspecific competition.

Some general features of macro-parasite systems which influence the construction of models are,

• The time the parasite spends within the human host is often much longer than the rest of its life cycle. Therefore, researchers often just consider the dynamics within the human population and assume that the rest rapidly achieve equilibrium.

• In general, it is found that a mass-action type formulation is the most accurate, where,

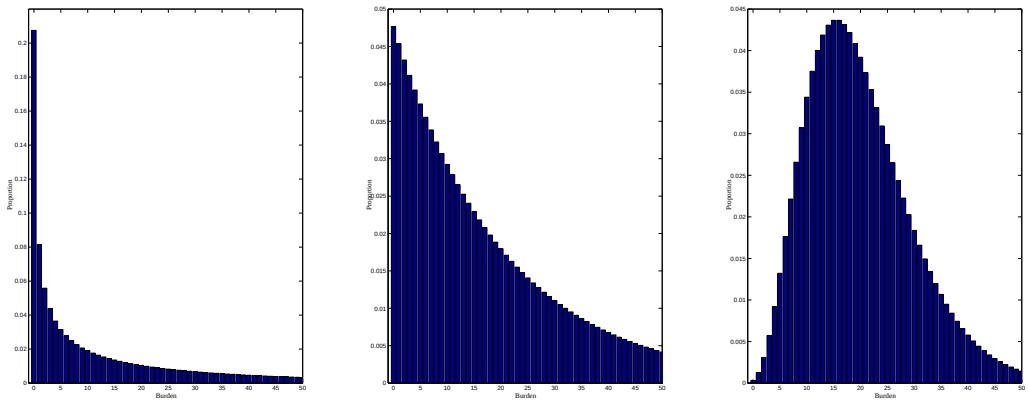
$$\text{The rate of infection } \propto \text{ density of host } \times \text{ density of infective stages}$$

• There is frequently a delay been the infection of a human host and the onset of reproductive maturity - this leads to the formulation of delayed differential equations. (Time-delayed differential equations frequently arise in population biology where they most often have a destabilising effect)

• Macro-parasite distributions are highly aggregated. The distribution is often termed **over-dispersed** which states that the variance is greater than the mean. The best fit to the distribution of worm burdens is a **negative binomial**,

$$I\!P(n|M) \;=\; \frac{(k+n-1)!}{n!(k-1)!}\left(\frac{M}{k}\right)^n\left(1+\frac{M}{k}\right)^{-k-n}$$

where $M$ is the mean worm burden and $k$ measures the degree of aggregation.



Figures showing how the distribution of the negative binomial changes with $k$. From left to right the level of aggregation is $k = 0.4$, $k = 1$ and $k = 5$; throughout $M = 20$.

As $k \to \infty$ the distribution tends to Poisson, and the variance is minimised. In general, it is found that $k$ lies between 0.1 and 1.0 (Anderson and May 1985). Notice that from the negative binomial distribution, we can find the **prevalence** of the macro-parasite. The prevalence is defined as the proportion of hosts that are infected, *ie* the proportion with a non-zero worm burden.

$$\text{prevalence} \;=\; 1 - I\!P(0) \;=\; 1 - \left(\frac{k}{M+k}\right)^{-k}$$

Prevalence levels for schistosomes in humans range from between 0.3% to 70%, depending upon location and species.

It has often been observed that many humans are predisposed to having a high or low worm

burden, although this may be partially attributed to life-style or nutrition, it is believed that much of this heterogeneity is due to differing levels of genetic resistance to infection.

Although many authors have devised complex, biologically realistic, mechanisms which produce a negative binomial distribution of worms, this distribution can also be formed from a much simpler basis. Assuming random infection at rate $\lambda$, a within host death rate of $d$, and a self-infection (or within host reproduction) rate of $r(< d)$, we obtain

$$\frac{d\mathbb{P}(n)}{dt} \;=\; [\lambda + r(n-1)]\mathbb{P}(n-1) \;+\; [d(n+1)]\mathbb{P}(n+1) \;-\; [\lambda + (d+r)n]\mathbb{P}(n)$$

The equilibrium distribution of this system is a negative binomial with,

$$M \;=\; \frac{\lambda}{d-r} \qquad k \;=\; \frac{\lambda}{r}$$

Hence, a simple birth-death model can also generate the observed levels of worm burden.

• Species of worm such as Schistosoma can produce between one hundred and three thousand eggs per female per day. However, density-dependence or crowding effects usually act to limit parasite survival and fecundity in those hosts with a high worm burden. In fact it has been observed that the number of cercariae emerging from a snail is independent of worm burden. Therefore snails can be considered as either susceptible or infected, and their precise burden ignored.

• Although most worm species reproduce asexually within the intermediate host (eg schistosomes reproduce asexually within the snail) the majority of parasites reproduce sexually within humans. This introduces the concept of a **mating function** $\phi$. Where $\phi(n)$ is the probability that a female worm is mated given that the worm-burden with the host is $n$. Note that in general $\phi(1) = 0$. Let us consider two cases where males and females are randomly distributed.

Polygamous

$$\begin{aligned}\phi(n) &= \sum_{f=0}^{n-1} \frac{f}{\frac{1}{2}n} \frac{n!}{(n-f)!f!} \left(\frac{1}{2}\right)^f \left(\frac{1}{2}\right)^{n-f} \\ &= 1 \;-\; \frac{1}{2^{n-1}}\end{aligned}$$

Monogamous

$$\begin{aligned}\phi(n) &= \sum_{f=0}^{n} \frac{min(f, n-f)}{\frac{1}{2}n} \frac{n!}{(n-f)!f!} \left(\frac{1}{2}\right)^f \left(\frac{1}{2}\right)^{n-f} \\ \phi(2n) \;=\; \phi(2n+1) &= 1 - \frac{1}{2^{2n}} \frac{(2n)!}{n!n!}\end{aligned}$$

[*Exercise: Check the above calculation from polygamous and monogamous worms*]. To calculate the rate of egg production, we need to sum the mating function and density dependent fecundity over the distribution of worms.

The differential equations for the mean number of adult worms per host and larvae is,

$$\begin{array}{rcccccc} \dfrac{dM}{dt} & = & \overset{\text{intake of larvae}}{\beta L(t-\tau_L)} & - & \overset{\text{worm death rate}}{\sum_n nd(n)\mathbb{P}(n|M)} & - & \overset{\text{host mortality}}{\sum_n nm(n)\mathbb{P}(n|M)} \\ \dfrac{dL}{dt} & = & \underset{\substack{\text{egg production}\\ \text{fertility \& mating fn}}}{N\sum_n E(n)\mathbb{P}(n|M(t-\tau_M))} & - & \underset{\text{larvae death}}{DL} & - & \underset{\text{uptake into hosts}}{N\beta L} \end{array}$$

where $N$ is the total number of potential hosts. This form of model has been used by many authors to study the dynamics of schistosome parasites and other helminths. Under the simplifying assumptions that $m$ and $d$ are independent of $n$, that $E$ is proportional to $n$ and that $\tau_L$ and $\tau_M$ are small, the equations reduce to,

$$\begin{aligned} \frac{dM}{dt} &= \beta L - dM - mM \\ \frac{dL}{dt} &= NEM - DL - N\beta L \end{aligned}$$

Hence, we can calculate $R_0$ for this system,

$$R_0 = \frac{NE}{d+m}\frac{\beta}{D+\beta N}$$

From this we observe that hosts must be above some critical density for the parasite to persist,

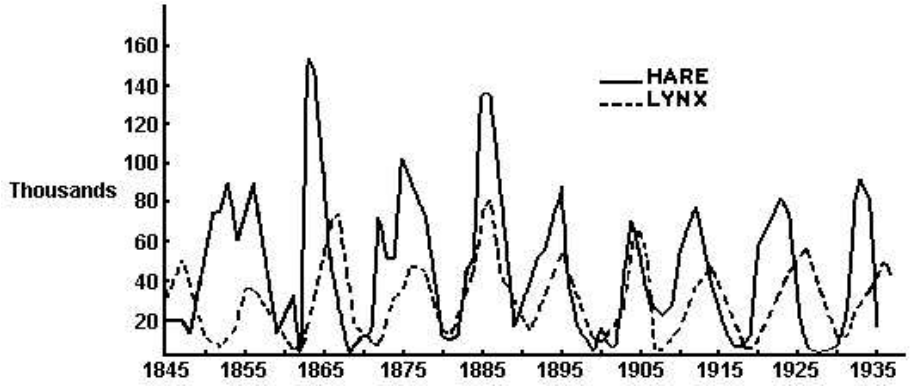
$$N > \frac{D(d+m)}{\beta E - \beta(d+m)}$$

If we allow the use of more general functions for $m$, $d$ and $E$, then usually the system is no-longer analytic and we have to resort to numerical calculations.

Mathematical ecology is very much a data-driven subject. One of the main sources of data is disease notification. Many diseases are termed notifiable, in which case all doctors have to notify a central authority each time they diagnose the disease. Obviously this provides a far larger data-set than could ever be gathered by a single researcher. In the next lecture we shall consider the data and models for measles more closely.

Being as collecting the ecological data is a long and painstaking process, there are a few major time-series which tend to receive the most attention. As with the diseases, the data in this time series is often collected by numerous people and later collated to provide a valuable resource. Below are three common examples, all demonstrate cyclic behaviour (no-one is interested in a stationary time series) and all come from northern latitudes where seasonality has a strong effect.

**Lynx in Canada**
Probably on of the best studied data sets is that of the lynx pelt harvests across Canada during this century.




Lynx-hare cycles in the Canadian arctic. Data is from the number of pelts sold to the trading companies in one year.

Early work on this time series attempted to relate the 8-10 year cycles to Sunspot activity, Lunar cycles (moonlight quality), and ozone or ultraviolet ray cycles. However, more modern research has looked for more biological explanations, including plant succession cycles and predator-prey dynamics.

By far the most popular explanation relies on the fact that the same 8-10 year cycles can also be observed in the population of snowshoe hares – a favourite prey of the lynx. A recent model for lynx ($L$), snowshoe hares ($H$) and vegetation ($V$) has produced some very interesting

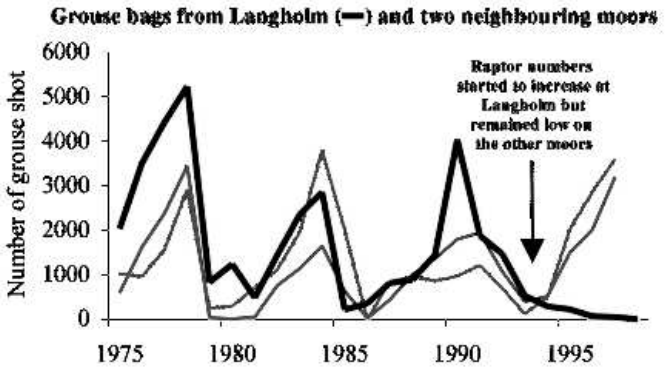
results (Blasius *et al* 1999)

$$\begin{aligned}\frac{dL}{dt} &= 0.6HL - 10L + 0.1 \\ \frac{dH}{dt} &= 0.1LV - 0.6HL - H \\ \frac{dV}{dt} &= V - 0.1LV - 1.5\end{aligned}$$

This is only a slight modification on the standard Lotka-Volterra tri-trophic model. It is capable of producing regular 'outbreaks' of lynx, but with chaotic maxima. It is still on going research to place models such as this in a spatial context and understand the observed spatial dynamics (Ranta *et al* 1997).

**Red Grouse in Scotland**

Another long-term data series, again from hunting records comes from the grouse moors of Scotland. The data displays a 4-5 year cycle, and although local moors usually cycle in phase this is not always the case. Data from other species of grouse in temperate climates also shows cycles, but the data is not as extensive.

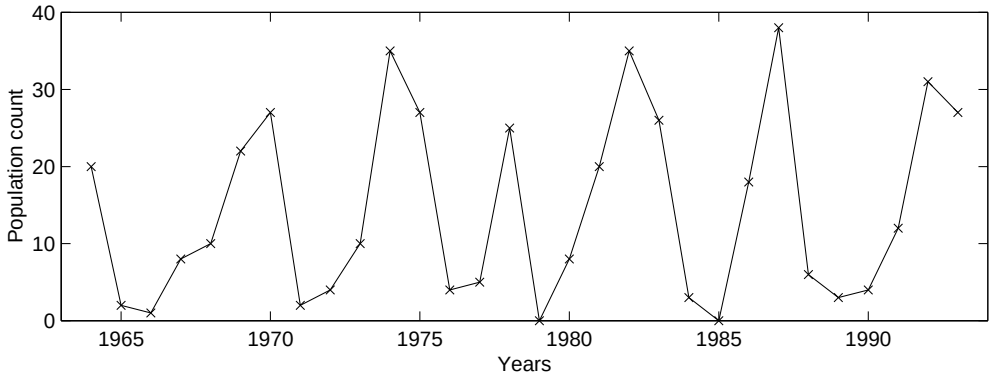



On the left a picture of a Red Grouse. On the right the number of grouse shot per year on three Scottish moors - data from ITE Banchory.

Although many explanations have arisen (including predation or competition and kin selection) by far the most popular is that the cycles are driven by macroparasitic worms (*Trichostrongylus tenuis*) and that the lack of synchronisation can be explained by the different environmental factors experienced by the moorlands (Hudson and Dobson 1997, Hudson *et al* 1998). It has been noted that within one area all the moors that fast westerly (and therefore receive more rain) cycle in phase, whereas all the other moors have a different phase; hence both intrinsic and extrinsic factors must be considered if we are to understand the dynamics.

**Voles in Scandinavia and Hokkaido**

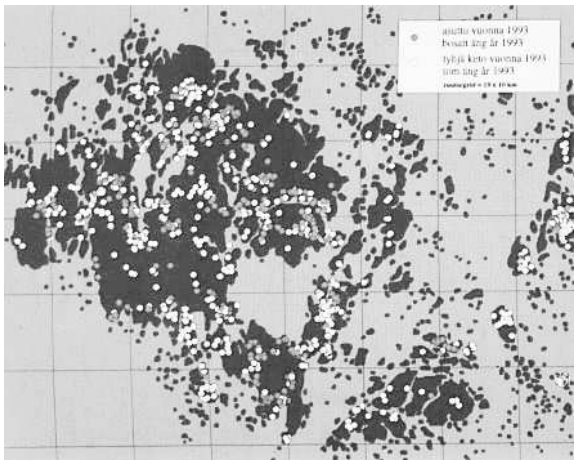
A new set of time series data being increasingly studied at the moment is the spatial distribution of voles. For many studies in both Britain and Scandinavia, researchers have observed slow linear travelling waves as well as 3-4 year cycles (Ranta and Kaitala 1997, May 1998). Such travelling waves can be explained by weak spatial coupling between local habitats, so that nearby points must be closely in phase.




Cycles of Vole populations in Lappi, Finland. Such cycles can be seen right across Scandinavia.

**Glanville Fritillary in Finland**

A group in Finland, lead by Ilkka Hanski (`www.helsinki.fi/science/metapop`), has been mapping the presence or absence of the Glanville Fritillary butterfly (*Melitaea cinxia*) in over 1500 suitable habitat patches in the Åland islands. In any one year about one third of the habitats are occupied, and in subsequent years there is a rapid turn-over of occupancy. This data which has been collected in detail for over 10 years is a valuable source of information on metapopulation dynamics.



Distribution of habitat occupied by the Glanville Fritillary butterfly in the Åland islands

**Disease Dynamics**

By far the richest source of high quality spatio-temporal data comes from disease outbreaks. The 2001 foot-and-mouth outbreak provided daily reports of farms diagnosed with infection, together with their location and the number and type of livestock. Similar data exist for the Swine-fever and Avian-influenza epidemics in the Netherlands. Finally, notifiable diseases offer an unparallelled set of detailed long-duration time series. The classic example is measles epidemics in England and Wales, where the number of reported cases per week in 1400 towns and cities is known from 1944 to the present. These data have proved invaluable in supporting and parameterising a growing number of modelling approaches. The data from 60 large cities and towns is freely available (`www.zoo.cam.ac.uk/zoostaff/grenfell/measles.htm`).