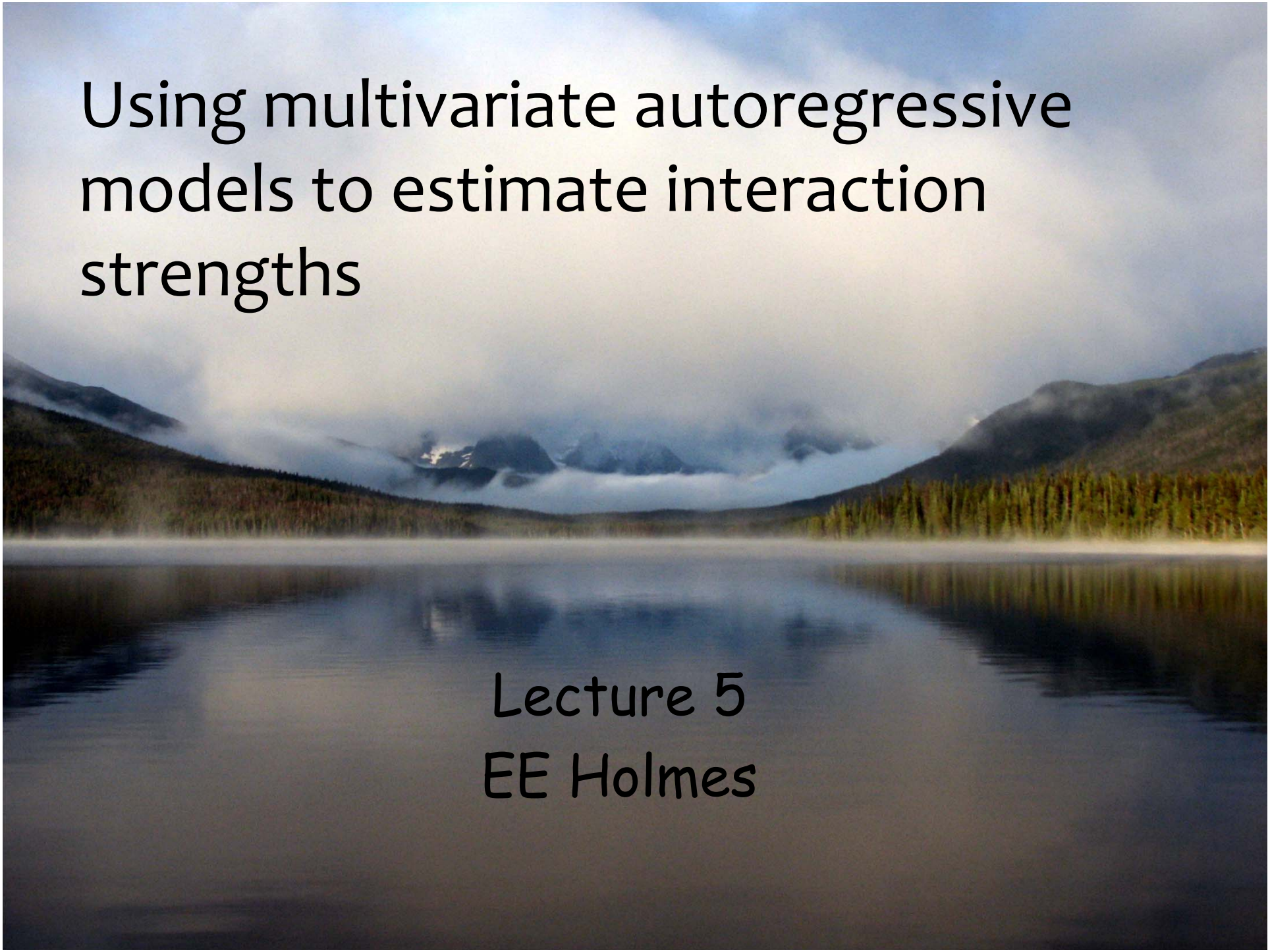
Using multivariate autoregressive models to estimate interaction strengths



Lecture 5

EE Holmes

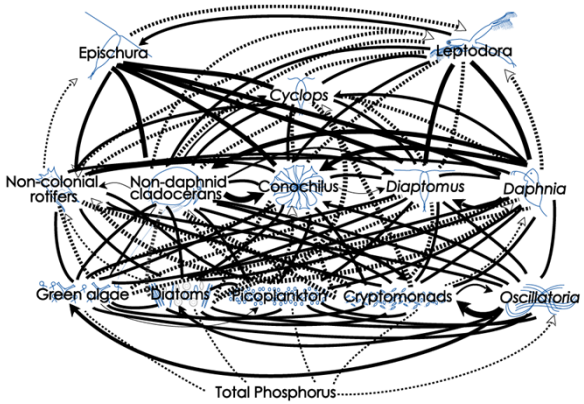
Univariate and multivariate Gompertz models

Morning: Univariate models and 2-spp models

- Estimating density dependence

Afternoon: Bigger multivariate models

- Estimating species interactions



The Gompertz model

univariate discrete time deterministic Gompertz model

$$n_t = n_{t-1} \cdot \exp\left[u + (b-1)\ln n_{t-1}\right]$$

- $|b| < 1$ "negative" density-dependence

  $b = 1$, no density-dependence

  $|b| > 1$, "positive" density-dependence (blows up)

- The closer b is to 0, the stronger the density-dependence (stronger the pull back to the mean). If b=1, there is no "pull" back to the "mean" (the mean is not in fact defined for this case).

Gompertz model written in log space

$$n_t = n_{t-1} \cdot \exp\left[u + (b-1)\ln n_{t-1}\right]$$

Taking the natural log of both sides

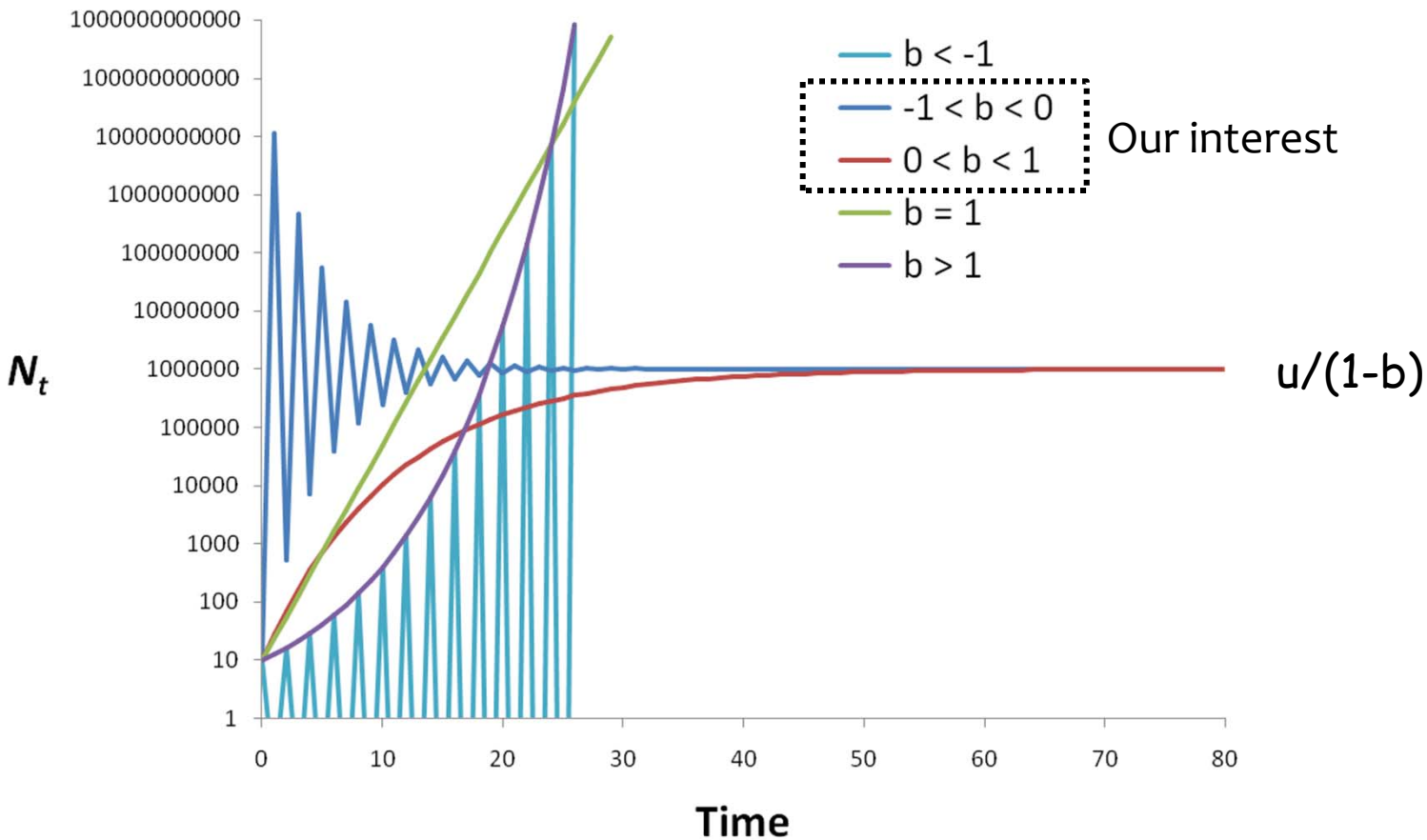
$$\begin{aligned}\ln n_t &= \ln n_{t-1} + u + (b-1)\ln n_{t-1} \\ &= \ln n_{t-1} + u + b\ln n_{t-1} - \ln n_{t-1} \\ &= u + b\ln n_{t-1}\end{aligned}$$

Substituting x_t for $\ln n_t$

$$\boxed{x_t = u + bx_{t-1}}$$

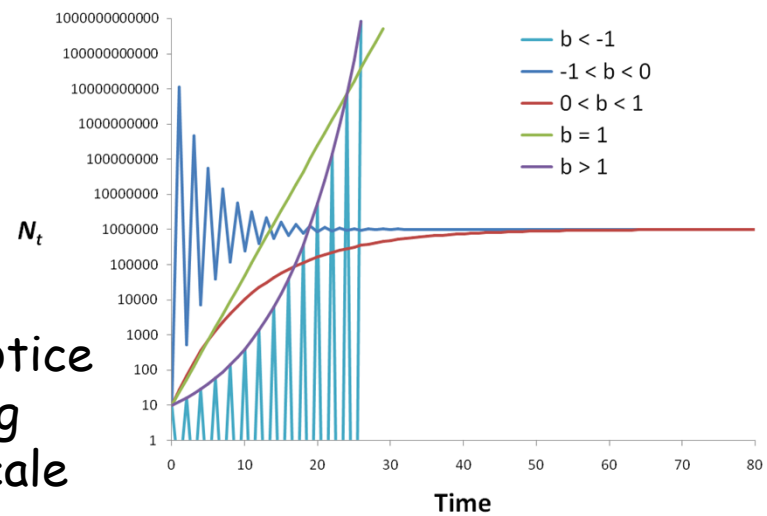
AR(1) minus the noise term

Examples of the Gompertz for different b



Notice that when b=1, we get a straight line

This line



notice log scale

The deterministic Gompertz eqn is just a linear (regression) model at b=1

$$x_t = u + x_{t-1}, x_0 = \pi$$

$$x_1 = u + \pi$$

$$x_2 = u + x_1 = u + u + \pi = u2 + \pi$$

$$x_3 = u + x_2 = u + u2 + \pi = u3 + \pi$$

This can be rewritten as

$$x_t = \pi + ut$$

π is the intercept
u is the slope

This emphasizes the relationship between the AR(1) model and a linear regression model, which is just **a hidden state model where b=1 and q=0 (no noise)**

Equilibrium for the deterministic Gompertz model

$$x_t = u + bx_{t-1}$$

the model reaches equilibrium at $t = \infty$, so we can write

$$x_\infty = u + bx_\infty$$

And via some algebra, we arrive at:

$$x_\infty = \frac{u}{1-b} \qquad \text{(provided } b \neq 1\text{)}$$

The equilibrium is a function of BOTH u and b.
This is rather unfortunate.

Add stochasticity (process error)

Adding stochasticity yields a univariate, lag-1 autoregressive or "AR(1)" process:

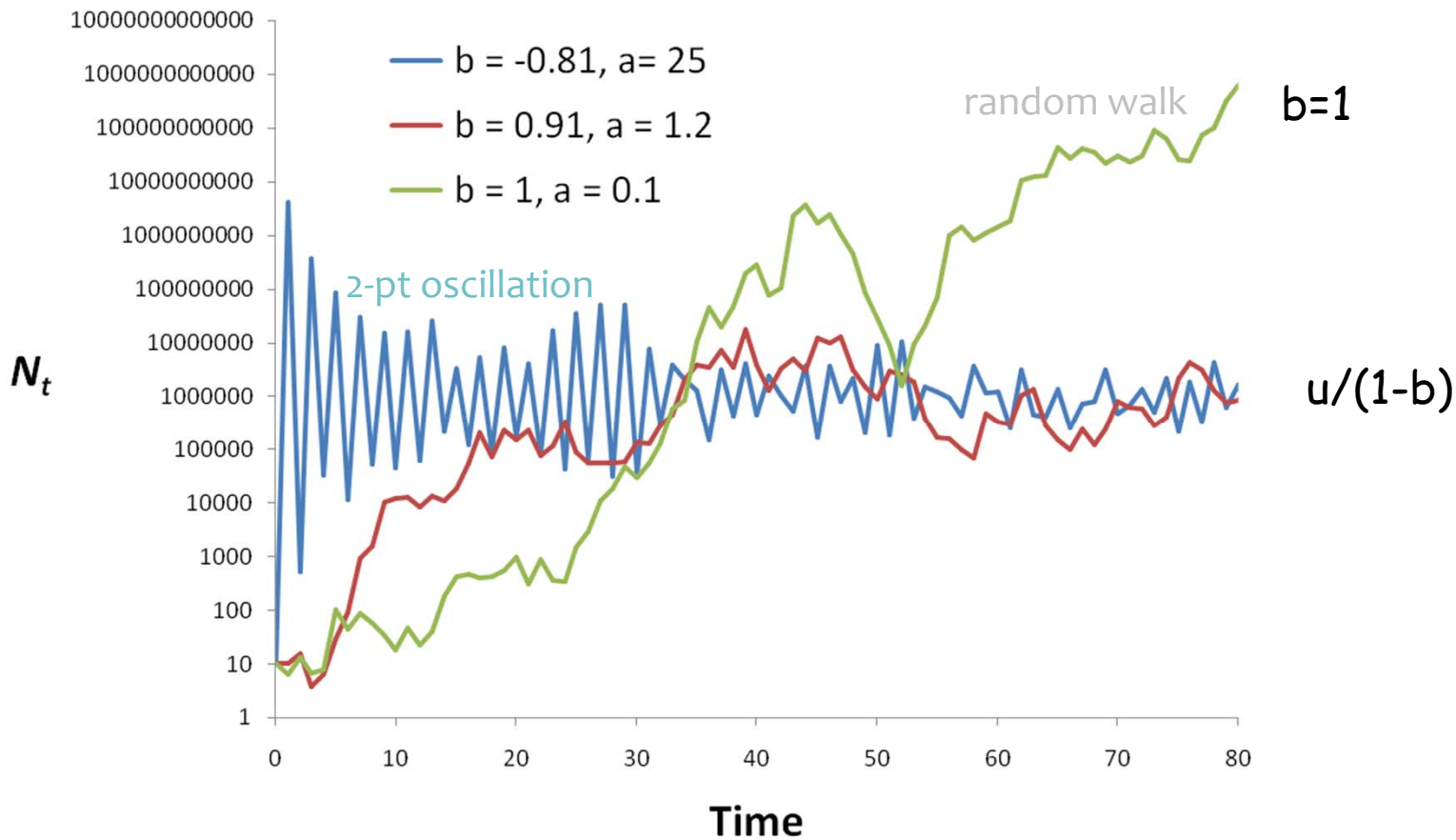
$$x_t = u + bx_{t-1} + w_t \qquad w_t \sim N\left(0, \sigma^2\right)$$

If $|b| < 1$, the process is "stationary"

If $b = 1$, the process is a "random walk" & "non-stationary"

Known as the discrete-time **Ornstein-Uhlenbeck** process in physics but as **Gompertz** or stochastic Gompertz model in population dynamics.

Example realizations



Equilibrium for the stochastic Gompertz process

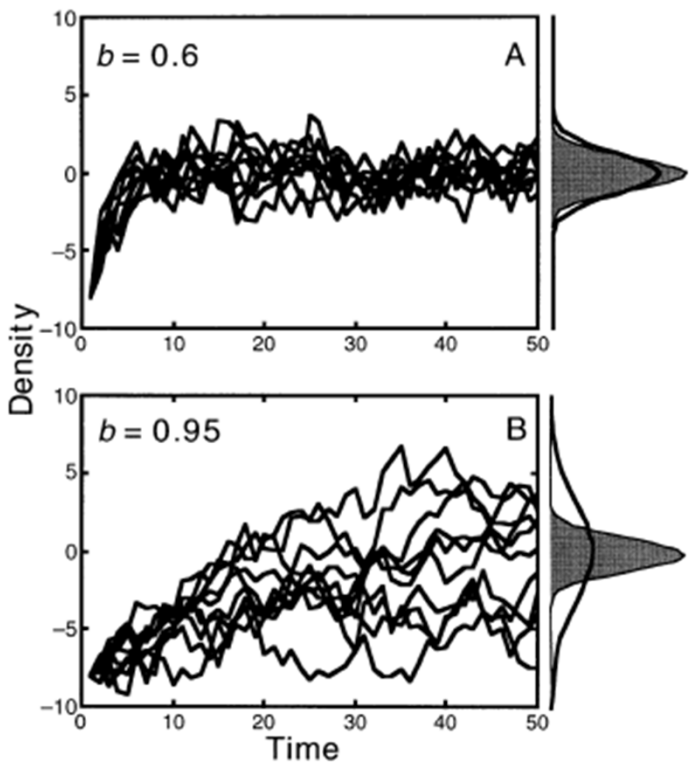


Fig. 1 - Ives et al. (2003)

<u>It has a stationary distribution</u>

probability distribution of X_t as $t \to \infty$ given $|b| < 1$

Normally distributed with mean μ_∞ and variance υ_∞

Properties of the stationary distribution

Assuming $|b| < 1$ (i.e. a stationary process)

mean $$\mu_\infty = \frac{u}{1-b} \qquad \text{(provided } b \neq 1\text{)}$$

variance $$v_\infty = q/(1-b^2)$$

Basic features of the Gompertz process

- Mean reverting, aka density-dependent

- Stationary, so it fluctuates around a mean

- Point equilibrium, unlike say Lotka-Volterra interaction models

- Can be seen as a **locally linear** approximation of other types of density-dependent interaction models

"locally linear" is jargon for "only holds for sure if density doesn't change too much".

Parameter estimation
Super easy in R

Open up R and follow after me

- lecture_5_example_0.R

So observation error is clearly a problem.
obs error = spurious density-dependence

ECOLOGY LETTERS

Ecology Letters, (2011) doi: 10.1111/j.1461-0248.2011.01702.x

LETTER

Are patterns of density dependence in the Global Population Dynamics Database driven by uncertainty about population abundance?

Jonas Knape* and Perry de Valpine
Department of Environmental Science, Policy and Management, 137 Mulford Hall #3114, University of California, Berkeley, Berkeley, CA 94720, USA
**Correspondence: E-mail: jknape@berkeley.edu*

Abstract
Density dependence in population growth rates is of immense importance to ecological theory and application, but is difficult to estimate. The Global Population Dynamics Database (GPDD), one of the largest collections of population time series available, has been extensively used to study cross-taxa patterns in density dependence. A major difficulty with assessing density dependence from time series is that uncertainty in population abundance estimates can cause strong bias in both tests and estimates of strength. We analyse 627 data sets in the GPDD using Gompertz population models and account for uncertainty via the Kalman filter. Results suggest that at least 45% of the time series display density dependence, but that it is weak and difficult to detect for a large fraction. When uncertainty is ignored, magnitude of and evidence for density dependence is strong, illustrating that uncertainty in abundance estimates qualitatively changes conclusions about density dependence drawn from the GPDD.

Keywords
Density dependence, GPDD, observation error, time series.

Ecology Letters (2011)

INTRODUCTION

Density dependence in population growth rates is a fundamental concept for ecological theory as well as for population management. Estimating density dependence in wild populations has, however, proved challenging. Ideally, density dependence in growth rates should be estimated directly from the effects of density acting on the traits contributing to population growth. Given current progress in statistical methods for jointly analysing data on both population size and demographic traits (Besbeas *et al.* 2005), and with long-term population studies involving demographic data becoming increasingly common, this approach holds a bright future. However, the number of such studies is currently limited and they only cover a rather narrow range of taxa. Long-term time series on population abundance are more common and can be used to estimate density dependence in population growth rates. Under this approach, density dependence is defined as a general tendency of per capita growth rates to decrease when population size is large and increase when it is small, and is identified as a statistical pattern not tied to any specific biological mechanism (Wolda & Dennis 1993).

It was noted early that estimates and tests of density dependence based on regressing log transformed current observed population size, y_t, on previous log transformed observed population size, y_{t-1}, are sensitive to uncertainty in the observations (St-Amant 1970; Kuno 1971; Ito 1972; Slade 1977). Similar concerns were aired about estimates from fisheries models of stock-recruitment data (Ludwig & Walters 1981; Walters & Ludwig 1981). Uncertainty inflates the Type I error rate of tests for density dependence (Shenk *et al.* 1998) and tends to bias estimates towards stronger density dependence if dynamics are under-compensatory and towards weaker density dependence if dynamics are over-compensatory (Benson 1973). Bulmer (1975) devised two tests for density dependence taking the time series nature of the data into account. One of those was designed to be robust against uncertainty about population size and has been shown to perform better than density dependence tests ignoring uncertainty in estimates of population abundance (Shenk *et al.* 1998). Simple procedures to correct for effects of uncertainty such as the SIMEX method have been suggested (Solow 1998; Freckleton *et al.* 2006) but typically require that the variance of the uncertainty about population size is known. A more direct approach to account for uncertainty is provided by state space models, first used for modelling population dynamics in the fisheries literature (e.g. Mendelssohn 1988; Sullivan 1992). State space models in these cases consist of a model of a population dynamical process combined with a model of the uncertainty in the abundance estimates, sometimes termed observation, measurement or sampling error, and may be used to estimate the variance of this uncertainty as well as to filter out its effects (de Valpine & Hastings 2002; Calder *et al.* 2003; Buckland *et al.* 2004; Dennis *et al.* 2006). Estimates derived from state space models tend to have smaller bias than estimates ignoring uncertainty about population abundance, but can also have large variances (Knape 2008), and the statistical properties of even simple state space model estimators are not fully understood (Dennis *et al.* 2006; Lebreton 2009).

The Global Population Dynamics Database (GPDD), containing over 5000 time series on population abundances obtained from various forms of population surveys, has provided an opportunity for ecologists to explore population dynamical patterns over a wide range of taxa (Inchausti & Halley 2001). Analyses using data in the GPDD have focused on, e.g., extinction risks (Fagan *et al.* 2001; Inchausti & Halley 2003; Brook *et al.* 2006), population cycles (Kendall *et al.* 1998; Murdoch *et al.* 2002) and effects of weather (Knape & de Valpine 2011) but, arguably, the studies stirring the most attention as well as debate have addressed population regulation and density dependence. These have explored patterns in the shape of density dependence (Sibly *et al.* 2005; Polansky *et al.* 2009) and in the strength of regulation and density dependence (Brook & Bradshaw 2006; Sibly *et al.* 2007;

© 2011 Blackwell Publishing Ltd/CNRS

2010

Ecology, 89(11), 2008, pp. 2994–3000
© 2008 by the Ecological Society of America

Reports

ESTIMABILITY OF DENSITY DEPENDENCE IN MODELS OF TIME SERIES DATA

Jonas Knape[1]
Department of Theoretical Ecology, Ecology Building, Lund University, SE-223 62, Lund, Sweden

Abstract. Estimation of density dependence from time series data on population abundance is hampered in the presence of observation or measurement errors. Fitting state-space models has been proposed as a solution that reduces the bias in estimates of density dependence caused by ignoring observation errors. While this is often true, I show that, for specific parameter values, there are identifiability issues in the linear state–space model when the strength of density dependence and the observation and process error variances are all unknown. Using simulation to explore properties of the estimators, I illustrate that, unless assumptions are imposed on the process or observation error variances, the variance of the estimator of density dependence varies critically with the strength of the density dependence. Under compensatory dynamics, the stronger the density dependence the more difficult it is to estimate in the presence of observation errors. The identifiability issues disappear when density dependence is estimated from the state-space model with the observation error variance known to the correct value. Direct estimates of observation variance in abundance censuses could therefore prove helpful in estimating density dependence but care needs to be taken to assess the uncertainty in variance estimates.

Key words: density dependence; state–space models; time series analysis.

Introduction

Density dependence can be loosely defined as a quantitative influence of population size on some life history or population trait of interest. The concept is of central importance to population ecology since it determines both the limiting and the short time behavior of the dynamics of populations. Empirical estimates of density dependence are therefore important from a scientific as well as from a management perspective. Assessment of density dependence in the dynamics of natural populations has however proved to be challenging (Dennis et al. 2006).

When relevant data are available, effects of density dependence can be directly linked to life history traits. For instance, density dependence in recruitment (e.g., Crespin et al. 2006) and survival (e.g., Festa-Bianchet et al. 2003) have been estimated by mark–recapture analyses and density dependence in fecundity has been inferred from data on reproduction (e.g., Solbreck and Ives 2007). Density dependence in life history traits influences density dependence in population growth rate (Lande et al. 2002). It can be argued that density dependence in population growth is the most important form of density dependence for determining long-term behavior of populations. However, since the link from demographic traits to population change is almost never known with good precision, density dependence in population growth rate is not easily inferred from life history data even if the effects of density dependence on several life history traits are well known. Time series analysis of population abundance data provides an alternative or complementary method that ideally could serve as a more direct way of estimating density dependence in population growth rate.

Estimates of density dependence must rely on measures of population density that are usually difficult to obtain with precision (Freckleton et al. 2006). This problem is particularly relevant to estimates of density dependence in growth rate derived from time series data on population size in that both the dependent and the independent variable are measured with uncertainty. Introducing observation error to dynamical data changes its dynamical structure (Dennis et al. 2006) and estimators relating to the dynamics of the data that do not account for observation errors are therefore often biased. Specifically, tests and estimators of density dependence based on time series data are known to be biased if observation errors are present but ignored for both direct (Kuno 1971, Walters and Ludwig 1981, Shenk et al. 1998, Freckleton et al. 2006) and delayed (Solow 2001) density dependence. An appealing method for overcoming this difficulty is provided by the state–space framework (Harvey 1990), a general term for statistical models of observations of hidden state variables that are dynamically linked through time. For time series data on population abundance, state–space models can be used for explicit modeling of both the observation and the population dynamical processes (Stenseth et al. 2003, Jamieson and Brooks 2004).

Manuscript received 12 January 2008; revised 2 June 2008; accepted 12 June 2008. Corresponding Editor: M. Lavine.
[1] E-mail: jonas.knape@teorekol.lu.se

2008

Parameter estimation accounting for obs error Super easy in R…

Open up R and follow after me

- lecture_5_example_1.R estimation technically easy..
- Lecture_5_example_2.R replication
- Lecture_5_example_3.R ML on the edge

Estimating R is not so easy, but replication helps A LOT

Ecological Monographs, 76(3), 2006, pp. 323–341
© 2006 by the the Ecological Society of America

ESTIMATING DENSITY DEPENDENCE, PROCESS NOISE, AND OBSERVATION ERROR

Brian Dennis,[1,5] José Miguel Ponciano,[2] Subhash R. Lele,[3] Mark L. Taper,[4] and David F. Staples[4]

[1]*Department of Fish and Wildlife Resources and Department of Statistics, University of Idaho, Moscow, Idaho 83844 USA*
[2]*Initiative for Bioinformatics and Evolutionary Studies (IBEST), Department of Mathematics, University of Idaho, Moscow, Idaho 83844 USA*
[3]*Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta T6G 2G1 Canada*
[4]*Department of Ecology, Montana State University, Bozeman, Montana 59717 USA*

Ab
ence,
model
state-s

Ecology, 89(11), 2008, pp. 2994–3000
© 2008 by the Ecological Society of America

ESTIMABILITY OF DENSITY DEPENDENCE IN MODELS OF TIME SERIES DATA

Jonas Knape[1]

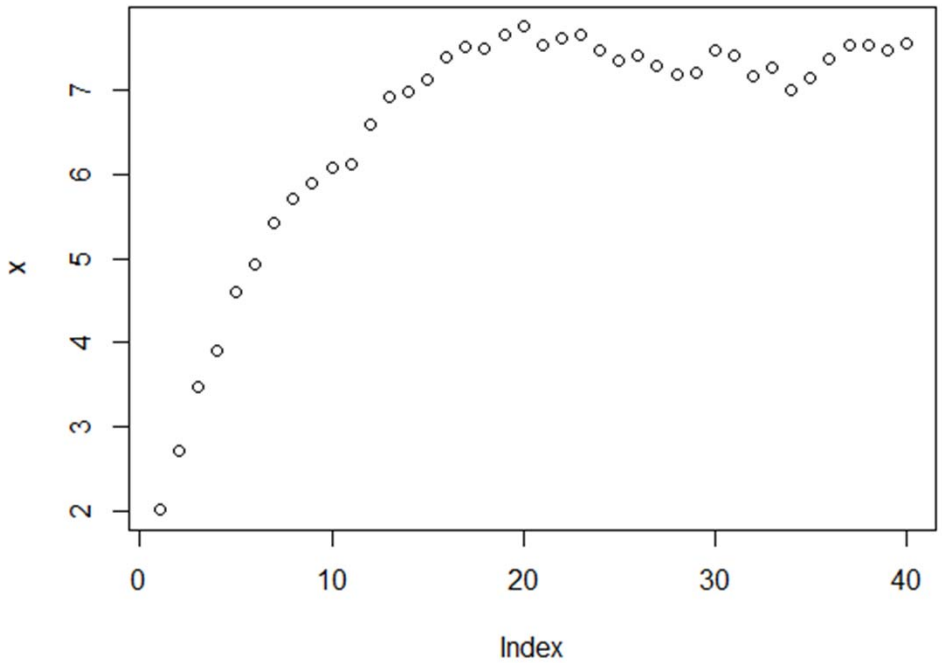
Department of Theoretical Ecology, Ecology Building, Lund University, SE-223 62, Lund, Sweden

Abstract. Estimation of density dependence from time series data on population abundance is hampered in the presence of observation or measurement errors. Fitting state–space models has been proposed as a solution that reduces the bias in estimates of density dependence caused by ignoring observation errors. While this is often true, I show that, for specific parameter values, there are identifiability issues in the linear state–space

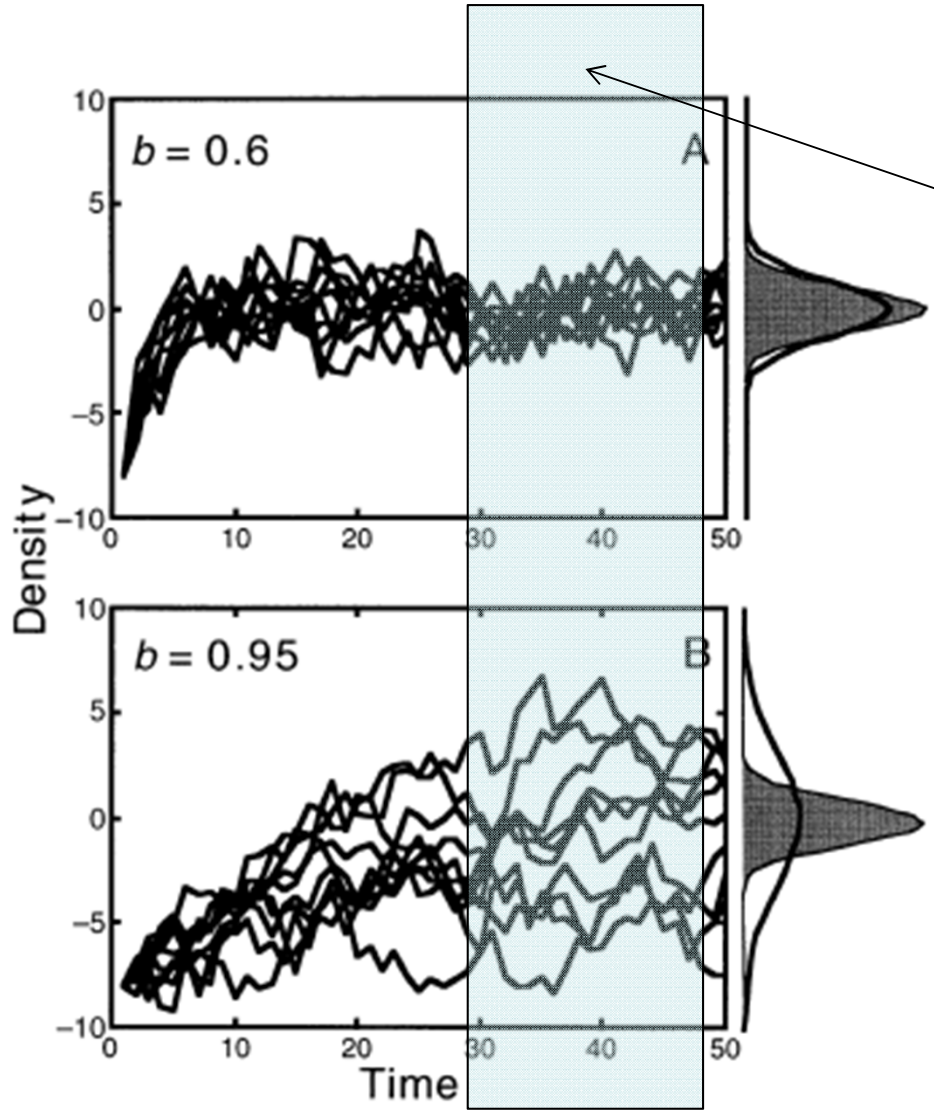
All the previous examples were NON-stationary

Open up R and follow after me

- lecture_5_example_0.R
- plot(x)



Estimation much easier if you can assume that your data are a sample from the stochastic equilibrium



Process has reached equilibrium

How to estimate b when you are willing to assume the data come from stoc. equil.?

Ecological Monographs, 76(3), 2006, pp. 323–341
© 2006 by the the Ecological Society of America

ESTIMATING DENSITY DEPENDENCE, PROCESS NOISE, AND OBSERVATION ERROR

Brian Dennis,[1,5] José Miguel Ponciano,[2] Subhash R. Lele,[3] Mark L. Taper,[4] and David F. Staples[4]

[1]*Department of Fish and Wildlife Resources and Department of Statistics, University of Idaho, Moscow, Idaho 83844 USA*
[2]*Initiative for Bioinformatics and Evolutionary Studies (IBEST), Department of Mathematics, University of Idaho, Moscow, Idaho 83844 USA*
[3]*Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta T6G 2G1 Canada*
[4]*Department of Ecology, Montana State University, Bozeman, Montana 59717 USA*

Abstract. We describe a discrete-time, stochastic population model with density dependence, environmental-type process noise, and lognormal observation or sampling error. The model, a stochastic version of the Gompertz model, can be transformed into a linear Gaussian state-space model (Kalman filter) for convenient fitting to time series data. The model has a

Idea #1 Impose the constraint that
W(t)=(y(t)-y(t-1)) and
W(t)-W(t-1)

- Have the variance-covariance structure of a stochastic Gompertz observed with error.
- Compute Q from the total **sample** variance and the **estimate** of b

How to estimate b when you are willing to assume the data come from stoc. equil.?

Idea #2 If you subtract E(x(t)) then U=0

- Use mean(data) as E(x(t))
- Has the added value of removing "a" too!
- Used in Ives et al 2003

- Lecture_5_example_4.R

Take home messages

U and B are confounded. Likelihood is banana-shaped, so we need to constrain u

- de-mean the data
- set u=0
- set tinitx=1

What happens when we add observation error?

- Estimation is more difficult.
- Replication will help us estimate R vs Q

2-species: Predator-Prey

Moose $$x_{m,t} = u_m + b_{m\to m} x_{m,t-1} + b_{w\to m} x_{w,t-1}$$

Wolf $$x_{w,t} = u_w + b_{w\to w} x_{w,t-1} + b_{m\to w} x_{m,t-1}$$

MAR(1): $x_t = Bx_{t-1} + u + w_t$

"spp" abundance

$$\begin{bmatrix} x_{m,t} \\ x_{w,t} \end{bmatrix} = \begin{bmatrix} b_{m\to m} & b_{w\to m} \\ b_{m\to w} & b_{w\to w} \end{bmatrix} \begin{bmatrix} x_{m,t-1} \\ x_{w,t-1} \end{bmatrix} + \begin{bmatrix} u_m \\ u_w \end{bmatrix} + \begin{bmatrix} w_m \\ w_w \end{bmatrix}$$

B = interaction matrix

Process variation
MVN(0,Q)

Meaning of the B matrix

$$\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & b_{13} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & \cdots & \vdots \\ b_{31} & \vdots & b_{33} & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{p1} & \cdots & \cdots & \cdots & b_{pp} \end{bmatrix}$$

Inter-specific effects
(can be set to zero)

Inter-specific effects
(can be set to zero)

Intra-specific effects
effect of spp i on itself,
aka density-dependence

Computer lab #2: moose and wolf

Adding covariates

$$\begin{bmatrix} x_{m,t} \\ x_{w,t} \end{bmatrix} = \underbrace{\begin{bmatrix} b_{m\to m} & b_{w\to m} \\ b_{m\to w} & b_{w\to w} \end{bmatrix}}_{\text{B = interaction matrix}} \begin{bmatrix} x_{m,t-1} \\ x_{w,t-1} \end{bmatrix} + \begin{bmatrix} u_m \\ u_w \end{bmatrix} + \underbrace{\begin{bmatrix} c_{1\to m} & c_{2\to m} & c_{3\to m} \\ c_{1\to 2} & c_{2\to w} & c_{3\to w} \end{bmatrix}}_{\text{covariate effect}} \overbrace{\begin{bmatrix} c_{1,t} \\ c_{2,t} \\ c_{3,t} \end{bmatrix}}^{\text{covariates}} + \begin{bmatrix} w_m \\ w_w \end{bmatrix}$$

Process variation not from covariates (“unexplained”)

Observation error causes

- **Spurious density-dependence, i.e. apparent stronger effect of self on self**
- **Spuriously low species interaction strengths, i.e. apparent lower effect of other on self**

$$\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & b_{13} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & \cdots & \vdots \\ b_{31} & \vdots & b_{33} & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{p1} & \cdots & \cdots & \cdots & b_{pp} \end{bmatrix}$$

Inter-specific effects (go to 0)

Intra-specific effects go to 0

R code #3: Lotka-Volterra predator-prey interactions

Simple 2-species system
Predator & Prey

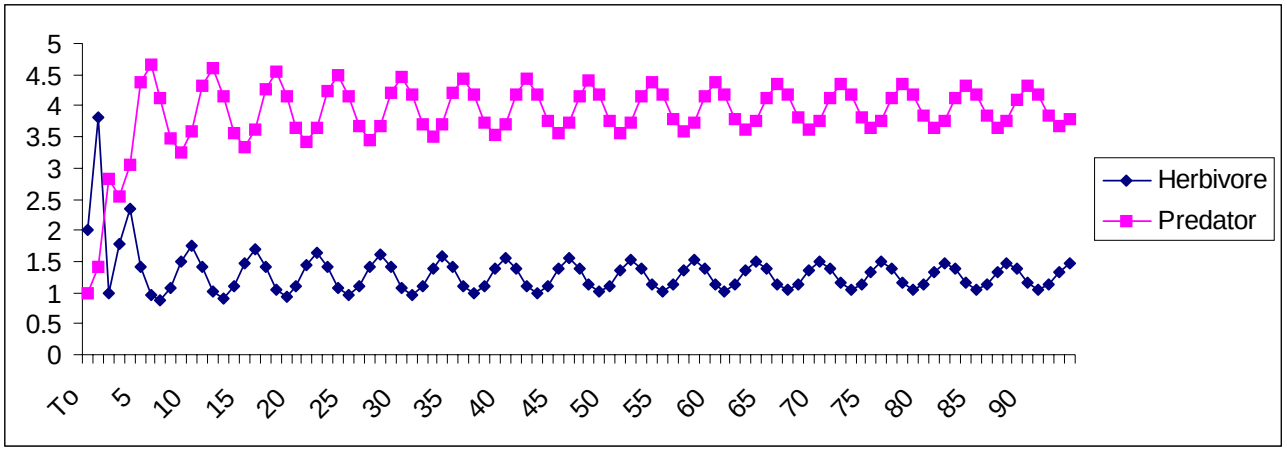
Data are simulated using a Lotka-Volterra model with density-dependence in the herbivore– easy to change interaction strength

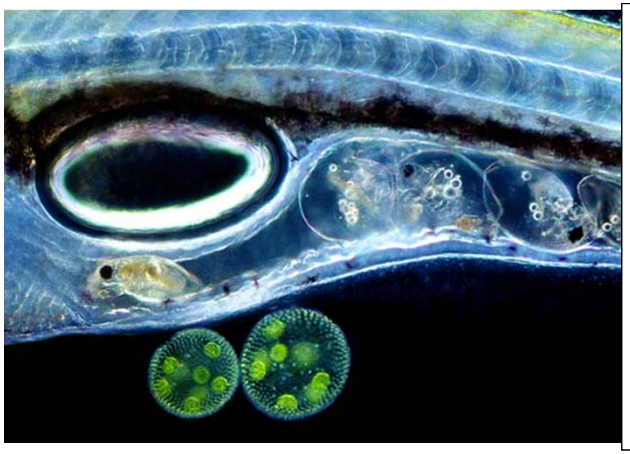
$$dH/dt = bH(1 - H/K) - aHP + \varepsilon$$

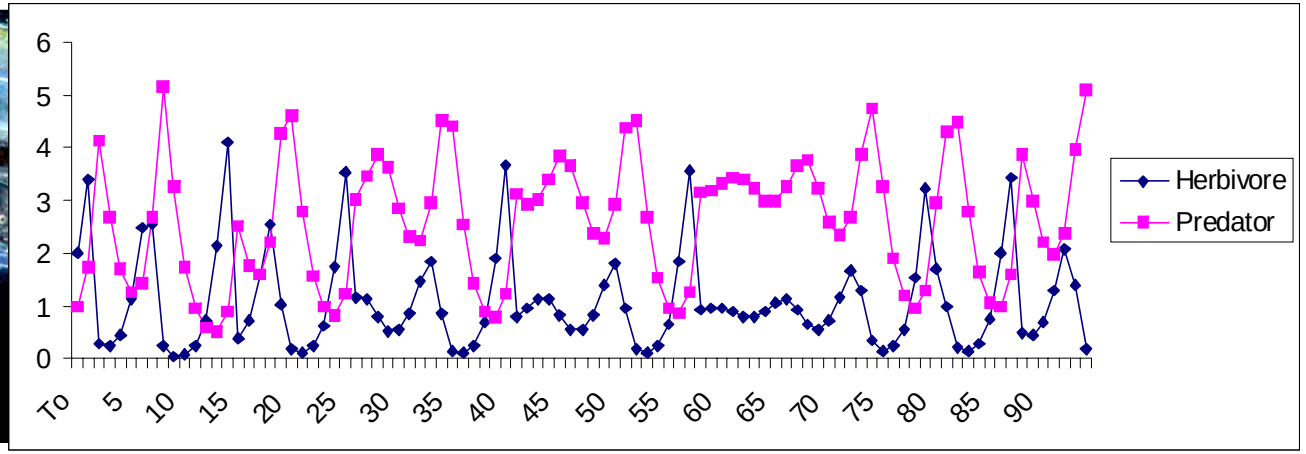
$$dP/dt = e(aPH) - sP + \varepsilon$$

H = herbivore sp
P = predator sp.
b = herbivore birth rate
K = herbivore carrying-capacity
a = per capita attack rate
e = conversion efficiency (consumed prey turning into new predators)
s = death rate for predators

This model can display a variety of dynamics







Estimate strength of density-dependence and interaction strength using MARSS

- Lecture_5_example_LV.R
  - change conversion efficiency of predator
- Lecture_5_example_LV_2.R
  - add observation error
- Lecture_5_example_LV_3.R
  - covariate affects K of herbivore
- Lecture_5_example_LV_4.R
  - covariate affects conversion efficiency of predator