Spring 2-1-2019

BIOB 595.07: ST - Ecological Models and Data

Robert O. Hall
University of Montana, Missoula

Ecological Models and Data
BIO 595 Spring 2019 3 Cr. Location ISB 103b.

W 1-2 pm First 3 weeks of semester, T, 15:00-19:00 and W 7:00-10:00

“Models without data are fantasy, data without models are chaos” (R. Hirsch, Found on Twitter, Dec. 2017)

Instructor: Robert Hall, Flathead Lake Bio Station. bob.hall@flbs.umt.edu Office hours: Wednesday 10-12, 416A Health Science

Learning outcomes

1 Learn a diversity of approaches in the context of ecological modeling, ranging from analytical, deterministic, probabilistic.

2 Create your own ecological models, ideally in the context of your graduate research.

3 Learn to confront these models with data.

4 Continue a never ending quest to perfect your written and graphical display of scientific information and understanding.

Course audience
I will steer his course to beginning graduate students with varying ranges of modeling skills and interests. It will not be an advanced class in discipline-specific models or statistics. If you are in your 4th year of a epidemiology Ph.D. you may already have the skills you need.

Course requirements and grading
Course participation, short exercises 50%
Semester-long project 50%

Course description
This class will take a generalist approach to describing, creating, and testing ecological models. Ask 10 ecologists what a “modeling” class entails and you will get 10 different answers ranging from analytical population models (Lotka-Volterra), to physiological-based models, “small” ecosystem / food web models, “large, complex” ecosystem models, simple linear models (regression), complex empirical models, and combinations of all of the above. The goal of this class is to explore general principles behind these models and then you will explore the suite of modeling approaches that are most relevant to your graduate research. Thus the course will comprise two parts:
1 An early semester overview of modeling principles
2 Weekly course meetings for ad hoc topics, readings, and project work.

This course will have focused time early in the semester. This will allow you and me to put all of our efforts in 7 h/wk so that you will have time to apply these methods to your own work later in the semester. Additionally, we will meet weekly for readings, discussion, and work together. I will also allocate time for 1:1 time helping with research projects. Your project will be one, hopefully related to your research interests, where you pose a hypothesis, build a model, and confront with data. If you have no data, then find data from a published study. There is so much data out there that one can be very successful (if not bored) by never going outside to collect more. For example, we have modeled primary production for >300 streams (>100,000 days) without ever leaving the computer. You will have no problem finding data!
Caveat. I make no claim whatsoever to be either an ecological modeler or a statistician. Nearly all I know comes from 20 y or working with people much more knowledgeable than myself and then hundreds of hours of self teaching. But this observation brings me to the 3 tenets of the course:

1 The tools and instruction for effectively using and testing ecological models with small to giant datasets have increased greatly in the past few years.

2 Given these tools, one does not need to be a specialist in ecological modeling to confront models with data.

3 Through instruction from me, teaching yourself, working with others, reading lots of good books, and assiduously practicing coding and math you can fearlessly learn new methods that will transform your abilities as a researcher.

Books I hesitate to recommend a single book because there are a zillion texts in ecological modeling and statistics and your favorite book will depend on your field and your approach. That said, there are two required texts: One is Bolker (2008). This book is step by step approach to understanding simple models and data. It is brilliant. The second Hobbs and Hooten (2015) It is a new book on Bayesian confrontation for ecologists. It focuses on theory and not computation. I love this book. I did not ask students to get it last year, but I found it by far the most useful for me as a teaching resource. If it works for me to teach, then I hope it will work for you too. I ask you to read parts two books because I find having two teachers is better than one. Here are my other favorites: Hilborn and Mangel (1997). The book that started ecologists down the path of confronting models with data. Excellently written, mathematically tractable, but dated. computationally. Kruschke (2014) The introductory stats book every grad class would use in an alternative universe where computers were developed before methods of statistical inference. Gelman and Hill (2007). The book on hierarchical regression modeling. New edition supposedly on the way. I can't wait. McElreath (2016). Great new book on Bayesian analysis. Focuses too much on his method of computation. Beautifully written.

More specific ecological subject matter books include: Caswell (2001) Matrix population models Agren and Bosatta (1998) A rare one on ecosystem models Gotelli et al. (1995) Classic on mathematical population biology Grimm and Railsback (2005) A method you may need Haefner (2012) Standard of the industry for deterministic model methods. Some stochastic also. Soetaert and Herman (2008) Have not used it, but new and uses R

Books that live near my computer: Wickham and Grolemund (2016) No escaping the Tidyverse. Murrell (2016) Handy reference for graphics.

Course prerequisites

1 Some understanding of a programming language. I use R as do many other ecologists, and it is fast becoming the lingua franca for many quantitative sciences. There is a lot of ink spilled on why R and I won't repeat it here. If you know some other language like Python, Matlab, etc., go ahead and use that; if you know these then you probably don't need my programming help anyway. But you should know enough R before this course to be able to load data, write a simple function, use a for() loop, make a scatterplot.

2 Enough statistics to understand things like regression, sums of squares, binomial probability, normal probability distribution. Enough math to explain $\frac{dN}{dt} = r_{max}N(1 - \frac{N}{K})$ and solve for its equilibrium.

3. Use of LaTeX and/or RMarkdown for your project and occasional assignments.. These are much better for writing technical literature. MSWord works great for letters and 3-page papers for classes. But for serious writing (papers, proposals, dissertations etc.) LaTeX is far better if you want them to look good and also not spend valuable time formatting. There is no better way for writing equations.You can grab the template for your favorite journal from the web, e.g., American Naturalist, AGU journals have nice ones. PNAS makes a draft that looks just like a PNAS paper, if you want to submit your project there! Template are nice because you let someone else worry about the formatting.

Learning disabilities

The University of Montana assures equal access to instruction through collaboration between students with disabilities, instructors, and Disability Services for Students. If you have a disability that adversely affects your academic performance, and you have not already registered with Disability Services, please contact Disability Services in Lommasson Center 154 or 406-243-2243. I will work with you and Disability Services to provide an appropriate modification.

Course schedule

Date	Tuesday (1500-1900)	Wednesday (0700-1000)	Reading
15Jan.	Model philosophy and types of models	Deterministic model examples	Bolker 1, 3, HH 1,2 Wiegert, Levins
22	Probability, Likelihood	Likelihood to Bayesian	Bolker 4, 6, 7, HH 3, 4, 5

Course will meet for 21 h early in the semester to cover most of the material. We will meet weekly during the semester to work on projects, discuss further reading, and address topics of student interests.

A note about the reading: Since we are moving quite fast, I put down the relevant chapters from the books. I do not expect you to commit this information to memory before the class. For example, there is no way to read 6 chapters for the 22 Jan class, and in fact they are the guts of all we will do this semester. Skim the chapters so that you know what will be coming, and then revisit after class and throughout the semester.

Date	Tuesday (1500-1900)	Wednesday (0700-1000)	Reading
15Jan.	Model philosophy and types of models	Deterministic model examples	Bolker 1, 3, Levins
22 Jan	Probability, Likelihood	Likelihood to Bayesian	Bolker 4, 6
29 Jan	Probability distributions	More Bayesian	Bolker 4, H

Below is a ROUGH schedule for the semester. I am going to keep the semester portion of the class flexible so that we can go at a pace suitable to you and your research interests. The below is merely a starting point.

Week	Topic	Reading
6Feb.	Bayesian wrap up	HH7
13 Feb.	Multilevel models	Clark, Cressie, HH 6
20 Feb.	More multilevel models	
27 Feb.	Model assessment via forward simulation	HH8
6 Mar.	The statistical crisis	"statistical crisis" folder
13		

Project

A central part of this course is a project where you will build a model(s) and confront this model with either data you collected or harvested from some repository. There are huge amounts of data on the web (see, for example hubbardbrook.org or the Hawaii Ocean Time Series or NEON or Dryad) if you do not already have some data of your own. I am going to keep the requirements of this part flexible so that you do what you want to do. But at a minimum, projects should test an ecological question by confronting models with data using likelihood or Bayesian approaches. Models are wide open and can be of regression type, SEM, dynamic models, time series, or anything else you can dream up. They should be written as a scientific paper, albeit a short one (10 pp text and ~5 display items should be ok). Each paper should have its code and output as a supplement written in RMarkdown to enable easy duplication and understanding (and grading) of the work. Bring your idea for a project to me by 6 Feb as a 1 page document (Rmarkdown or LaTeX please). I will be around to assist with your project all semester You will submit a draft of your project to me by 10 April. In the last 2 weeks of class you will give a 10-min oral presentation of your project to the class, focusing on the technical aspects of the model and data confrontation.. Please make this project as relevant to your graduate work as possible.

References

Agren, G. I. and E. Bosatta. 1998. Theoretical ecosystem ecology: understanding element cycles. Cambridge University Press.

Bolker, B. M. 2008. Ecological models and data in R. Princeton University Press.

Caswell, H. 2001. Matrix population models. Wiley Online Library.

Gelman, A. and J. Hill. 2007. Data analysis using regression and multilevelhierarchical models, volume 1. Cambridge University Press New York, NY, USA.

Gotelli, N. J. et al. 1995. A primer of ecology. Sinauer Associates Incorporated.

Grimm, V. and S. F. Railsback. 2005. Individual-based modeling and ecology, volume 2005. BioOne.

Haefner, J. W. 2012. Modeling biological systems: principles and applications. Springer Science & Business Media.

Hilborn, R. and M. Mangel. 1997. The ecological detective: confronting models with data, volume 28. Princeton University Press.

Hobbs, N. T. and M. B. Hooten. 2015. Bayesian models: a statistical primer for ecologists. Princeton University Press.

Kruschke, J. 2014. Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan. Academic Press.

McElreath, R. 2016. Statistical Rethinking: A Bayesian Course with Examples in R and Stan, volume 122. CRC Press.

Murrell, P. 2016. R graphics. CRC Press.

Soetaert, K. and P. M. Herman. 2008. A practical guide to ecological modelling: using R as a simulation platform. Springer Science & Business Media.

Wickham, H. and G. Grolemund. 2016. R for Data Science. O'Reilly, Sebastopol, CA. http://? r4ds.? had.? co.? nz.