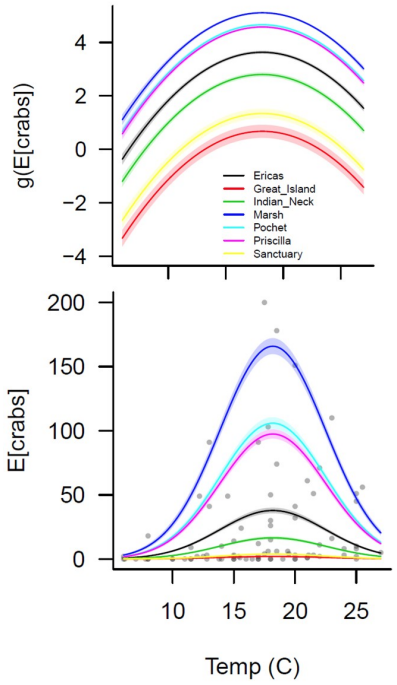
ECO 363: Applied Ecological Statistics
Spring 2018

Instructor
Dr. Chris Sutherland
Room 118 Holdsworth Hall
Office hours: *available on request*
csutherland@umass.edu

Teaching Assistant
Ben Padilla
Ag Eng 109
Office hours: Th 2:30–4*pm*
bjpadilla@umass.edu

Lecture: Tu & Th 1.00–2.15*pm*
Lab: We 4.00–6.00*pm*
Where: Holdsworth 301

Credits: 4



Course Overview

This graduate level course is an introduction to *modern regression methods*, arguably the most commonly applied statistical methods in ecology. The entire focus of the course is on understanding and applying linear and generalized linear models.

The course begins with an overview of data exploration and of the basic features of *linear models* before switching focus to developing a thorough understanding of the structure and application of linear models, and the considerations that need to be made in each step of a typical analysis. We start by describing in detail the *normal* linear model, the foundation of everything else in this course, and see how familiar statistical tests such as the *t-test*, *linear regression*, *ANOVA* and *ANCOVA* are all special cases of a general linear model. We will explore the utility and flexibility of this model in great detail as much of what comes later can be seen as a natural extension of this linear model.

Having mastered the structure and application of the *normal* linear model, we will move to *generalized* linear models (GLMs). These models are are used to analyze data-types commonly encountered in ecology: first count data using a Poisson and negative binomial

GLMs, and then binary data using Binomial GLMs. We will finish by discussing another family of models commonly applied in ecology - *generalized additive models* (GAMs).

This course will be taught using a combination of lectures and lab practicals. `R`-based practical sessions will be used to reinforce the theory covered in lectures, and provide supervised hands-on experience of analyzing ecological data. We will use a combination of real-world and simulated ecological data sets. Real-world examples will be drawn from research conducted by myself, ECO faculty, and ECO students. I encourage you to take this opportunity to use your own data where appropriate.

Table 1: Details of where and when to meet for lectures and lab practicals.

Day	What	Where	When
Tuesday	Lecture	Holdsworth 301	1.00 – 2.15 *pm*
Wednesday	Lab	Holdsworth 301	4.00 – 6.00 *pm*
Thursday	Lecture	Holdsworth 301	1.00 – 2.15 *pm*

Course structure & materials

All course materials will be made available on `moodle`. Lecture slides and the lab exercises will be available on `moodle` at least the morning before each session, where they will remain for the semester. The course page will be closed down after the course. I make an effort to create the course materials in the hope that they are useful stand-alone resources, so please download them all before the end of semester. The course has two components: *Lectures* and *Labs*.

Lectures

Lectures will present statistical background and context as well as demonstrations of how to implement these models in `R`. You will benefit most from lectures by:

1) attending *all* lectures (required)
2) take notes during lectures (recommended)
3) ask questions during lectures (recommended)

I *strongly* encourage you to ask questions at any point during any of the sessions (lab or lecture) - **there is no such thing as a silly question**! The schedule is structured in a very flexible way so that we can take as much, or as little, time on specific topics to ensure key conepts are understood.

Labs

Labs will be R-based practical sessions. You are encouraged to use your own data where possible, but **only if they are appropriate for the lab topic**. This time provides you with the opportunity to reinforced concepts covered in lectures through practical application.

Textbooks

There are many *many* terrific books devoted to the teaching of introductory statistics and (generalized) linear models. Of course, I prefer the ones directed at ecologists/environmental scientists, and that strike a balance between theory and application. Below are some books I have used to generate the material for this course (by permission), as well as some other books that are great resources (the majority of which are available free through the library).

- Crawley, M. J. (2012). The R book. John Wiley & Sons.
- Fox, G. A., Negrete-Yankelevich, S., & Sosa, V. J. (Eds.). (2015). Ecological Statistics: Contemporary theory and application. Oxford University Press.
- Kéry, M. (2010). Introduction to WinBUGS for ecologists: Bayesian approach to regression, ANOVA, mixed models and related analyses. Academic Press.
- Quinn, G. P., & Keough, M. J. (2002). Experimental design and data analysis for biologists. Cambridge University Press.
- Zuur, A., Ieno, E. N., & Smith, G. M. (2007). Analysing ecological data. Springer Science & Business Media.

Assessments

This course will be assessed through a combination of short quizzes, group take-home practicals and a final project (report and presentation). In addition, attendance will be taken and class participation will be noted.
Attendance will be taken by the instructor. Quizzes will be written short answer questions and grades will be uploaded to `moodle` within a week. For each assignment, groups will submit a single pdf document to `moodle` which will be graded within a week.

Quizzes (20%)

Quizzes will be given once a week at the the start of class and aim to assess your understanding of topics covered during lectures. Quizzes also help me identify topics that require more or less attention. Quizzes will be short and be based on material recently covered in lectures. There will be 10-12 quizzes, depending on our progress. Each quiz will be equally weighted and make up 20% of the final grade.

Practical assignments (40%)

Each take-home group assignment will be discussed in detail during the lab session, i.e., take-home practicals will be assigned on Wednesdays, and are due the **Friday of the following week**. Group assignment will provide opportunities for you to apply material from the lectures and labs unsupervised. These group assignments will require a single submission per group. There will be 5-7 assignments, depending on progress. Each assignment will be equally weighted and make up 35% of the final grade.

Final Project (35%)

The final project will be in the form of a presentation and an accompanying *Statistical Report* - a short paper that includes a brief introduction, a detailed methods and results section, a discussion, and an appendix containing the `R` code used to conduct the analysis. The focus will naturally be on the choice, use, and interpretation of the statistical methodology. Working in pairs, you will ideally analyze your own data, although data sets can be provided. The aim of the final project is to conduct an entire analysis from start to finish, describing in detail the choices made at each step of the analysis, thereby demonstrating an understanding of the material covered throughout the course. The final project report will make up 35% of the final grade (presentation: 10%, paper: 25%).

Attendance and participation (5%)

You are expected to attend all lectures and labs, you should complete all take-home assignments, and you are expected to make meaningful contributions to group exercises and discussions. Attendance will be recorded and each class missed will result in a 1% deduction from the available 5%. In addition, in-class participation will be monitored and students expressing low levels of participation or interest will incur point deductions.

Table 2: Grading structure for the course.

Letter	Grade Points	Percentage
A	4.0	94 - 100
A-	3.7	90 - 93
B+	3.3	87 - 89
B	3.0	84 - 86
B-	2.7	80 - 83
C+	2.3	77 - 79
C	2.0	74 - 76
C-	1.7	70 - 73
D+	1.3	67 - 69
D	1.0	64 - 66
F	0.0	0 - 63

Schedule

The course will follow this tentative schedule. I have built in enough flexibility in the timetable so that topics are well understood before we advance to the next topic. So, treat this as a *tentative* schedule. You will be notified of any major changes to this schedule as the course progresses.

Topic	**Material**
Course introduction	*Introductions & course outline*
Data Exploration	*Data exploration*
Linear Models	*The statistical linear model*
	Categorical predictors - t-test & ANOVA
	Continuous predictors - linear regression
	Continuous & categorical predictors - ANCOVA
	Interpreting output
Generalized Linear Models	*Generalizing the linear model*
Count data	*GLM with Poisson errors*
	Overdispersion
	Zero-inflated Poisson
	Interpretation
Binary data (Binomial errors)	*GLM with Binomial errors*
	Interpretation
Generalized Additive Models	*Introduction to GAMs*
	Interpreting GAMs

Local R user listserve

I have set up an internal (for ECO, OEB, etc.) `R` users group that will, hopefully, provide a friendly and safe environment for asking `R` related questions. As you continue to develop your proficiency in `R` you will realize that the environment provides a great deal of freedom and flexibility for handling, manipulating, visualizing and analyzing data. If you are struggling with some coding, it is *highly* likely that the problem you are trying to solve has been closely approximated, or even solved, by someone in the department. The point of this user group is to provide a space for you to ask questions about how to do such things so you can spend time on research, rather than coding.

- to subscribe, visit: https://list.umass.edu/mailman/listinfo/rusers
- you will be added within a few days
- only subscribers can email the group
- rusers@eco.umass.edu

NB: This is **NOT** a statistics forum. Students should use the *QSG* consulting sessions to seek statistical advice. This is meant to alleviate some of the frustrations of getting `R` to do precisely what you it to do, without taking time away from your core research.

Policy Statements

1. Accommodations

The University of Massachusetts Amherst is committed to providing an equal educational opportunity for all students. If you have a documented physical, psychological, or learning disability on file with Disability Services (DS), Learning Disabilities Support Services (LDSS), or Psychological Disabilities Services (PDS), you may be eligible for reasonable academic accommodations to help you succeed in this course. If you have a documented disability that requires an accommodation, please notify me within the first two weeks of the semester so that we may make appropriate arrangements (http://www.umass.edu/disability/).

2. Academic Honesty

Since the integrity of the academic enterprise of any institution of higher education requires honesty in scholarship and research, academic honesty is required of all students at the University of Massachusetts Amherst. Academic dishonesty is prohibited in all programs of the University. Academic dishonesty includes but is not limited to: cheating, fabrication, plagiarism, and facilitating dishonesty. Appropriate sanctions may be imposed on any student who has committed an act of academic dishonesty. Instructors should take reasonable steps to address academic misconduct. Any person who has reason to believe that a student has committed academic dishonesty should bring such information to the attention of the appropriate course instructor as soon as possible. Instances

of academic dishonesty not related to a specific course should be brought to the attention of the appropriate department Head or Chair. The procedures outlined below are intended to provide an efficient and orderly process by which action may be taken if it appears that academic dishonesty has occurred and by which students may appeal such actions. Since students are expected to be familiar with this policy and the commonly accepted standards of academic integrity, ignorance of such standards is not normally sufficient evidence of lack of intent. (http://www.umass.edu/dean_students/codeofconduct/acadhonesty/)