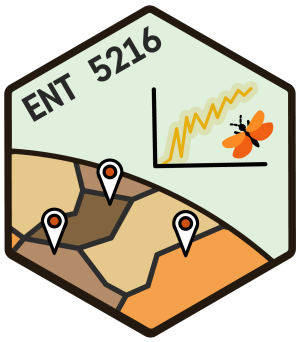

Entomology 5126

Spatial and Temporal Analyses (of Bio/Ecological Data)

Course stewardship

Instructor Jeremy Hemberger, PhD | iBUG Lab
Office 408 Hodson Hall
Email jhemberg@umn.edu

Course details

Lectures Wednesday and Friday, 9:30-10:20am, ROOM
Labs Friday, 10:30-12:30pm, ROOM
Course webpage website link goes here
Course Discord server discord server link goes here

Description

Anyone attempting to analyze data from lab, field, or existing datasets are frequently confronted with a stark reality: organisms rarely stay in the same place in space and time. It is well known that spatial and temporal dependence in biological/ecological data can affect inference; but how can we (1) detect potential problems; (2) determine their severities; and (3) apply the appropriate methods so that our analyses robustly and accurately represent reality and allow clear and unbiased inference? Attempts to locate suitable tools often encounter mathematical jargon inaccessible to the average practitioner. *The goal of this course is to provide a gentle, but thorough foundation to the concepts and tools needed to analyze, interpret, and communicate spatially and temporally dependent data.*

The first half of the semester starts with a brief refresher on using R for data management and linear models (regression and ANOVA). In the second half of the semester, we will extend these foundational frameworks to include an introduction to temporal data analysis and three major areas of spatial statistics: spatial point processes, lattice data, and geostatistics. By bridging basic theories and practices, students will leave this course understanding why it is important to contend with spatiotemporal dependence and the skills needed to confidently apply these practices to their own data.

Student learning outcomes

Upon successful completion of this course, students will be able to:

1. **Apply the appropriate techniques to characterize spatial and temporal dependence in their data**
2. **Understand how spatial and temporal dependence can affect statistical inference**
3. **Identify and choose the appropriate analytical methods to contend with such dependence in widely-used statistical methods like regression and ANOVA**
4. **Critically evaluate and understand the strengths and limits of these approaches**
5. **Present analytical results in a format befitting a scientific presentation or publication**

Meeting times

Lecture: We will meet for two lectures per week, with a lab session following after the second lecture. Attendance is, obviously, strongly encouraged! For some of you, the material in weeks 1-3 might mostly be review from previous courses. Despite this, I still encourage you to attend to (1) make sure your memory serves and (2) get to know your fellow students (i.e., collaborators). We'll also be re-enforcing the expectations for future lab assignments.

Lab: Because campus does not have a computer lab available for our class, please bring your own laptop with R and RStudio installed to class. If you don't have access to one, please let me know and I'll facilitate furnishing one for you! The lab session each week will emphasize hands-on analytical skills and statistical methods relevant to the week's lecture content. Each lab will have an accompanying assignment/report that will be graded.

Student assessments

Lecture: Lecture-related components will comprise 45% of the course grade and include: (1) a midterm exam (20%); and (2) a comprehensive final exam (25%). These will include concepts covered in the lectures and may include relevant examples or reference to skills practiced in the labs, as well.

Lab: There are a total of 13 labs, each with a lab assignment/report that will be evaluated. The first three periods will be devoted to a primer of using R/RStudio for managing, wrangling, and interpreting data, and the reports associated with these will receive feedback but will not be included in the course grade. The remaining 10 lab assignments/reports will be graded and count toward 40% of total course grade (4% per

report). There will also be a lab final worth 15% of the total course grade. In this, students will be given a real data set to analyze and develop a report for during the last lab period.

Lab reports will be due one week following the lab exercise at the beginning of the lab. Please upload these into Canvas. Please don't email these to me! There's no grade penalty for late submissions of lab reports. That said, I *highly* recommend staying caught up. The level of feedback I'm able to give will decline exponentially with each passing day, so that's effectively the only punishment.

At the end of the course, letter grades will be assigned to the quantitative point totals from student performance in the class. Pretty standard stuff here!

Point Breakdown

Assignment	Points	Percent
Lab reports (10 × 16)	160	40%
Lab final	60	15%
Lecture midterm	80	20%
Lecture final	100	25%
Total	**400**	**100%**

Final Grade Breakdown

Grade	Range	Grade	Range
A	93–100%	C	73–76%
A–	90–92%	C–	70–72%
B+	87–89%	D+	67–69%
B	83–86%	D	63–66%
B–	80–82%	D–	60–62%
C+	77–79%	F	< 60%

Adjustments to final grades based on class participation or instructional interaction demonstrating mastery may be made on a case-by-case basis. Similarly, adjustments to the entire class based on comparisons to previous offerings may be made.

Learning material

Textbooks/Readings

This course does not have a single, assigned text, however we'll make use of material from several reference books, journal articles, and websites throughout the semester. All of these materials will be posted on the course website. These are for reference only and serve as more in-depth treatments of some of the course material.

Software

The purpose of this class is to teach **concepts** over **software** - - this ensures that after this course, students will be able to implement the methods covered in their software of choice. That said, we need to choose some computational tool to do our work, so we will be using R and its integrated development environment, RStudio, in this course. R is the most widely used language and statistical software in the life sciences. It's available on all major operating systems and is 100% free and open-source ensuring that students will retain access to the toolkits developed in this course long after the semester ends. We'll cover all of the basics of installing and using R to load, manipulate, and visualize data, assuming that folks may have some familiarity with R, but with enough detail that novices will be able to pick up the necessary skills as well. However, if you'd like to get a head start on learning R, some great, free online resources are included on the course website.

Policies

Attendance

Attendance in lectures and labs is ***strongly*** encouraged. Can you learn these skills on your own? Sure. But you'll get a better outcome and enjoy the process a whole lot more by doing it with your friends and colleagues! This is especially the case when learning a new skillset like R - - being able to quickly bounce a question off of someone IRL is always preferable to spending hours trawling StackExchange for an answer. Please give me a heads-up if you will be away for a week or more. If you will miss any assessment activity, it is your responsibility to contact me in advance so suitable arrangements can be made for a make-up exercise given valid circumstances (time-sensitive field work, bereavement, medical, etc.). Absence of a good faith attempt to let me know of your anticipated or unexpected absence may result in a grade of zero for missed assessments.

AI, LLMs, and Bullshit

I highly recommend **NOT** using ChatGPT or other large language models (aka LLMs) in this (or any…) class.

These tools, while highly functional and time-saving in certain situations, are at direct odds with the practice of learning and developing competence in concepts and skills - be they coding, writing, or other artistic pursuits. I'm not opposed to using these tools in a variety of circumstances, and I use ChatGPT for a variety of tasks to provide a quick review of my own work. These tools are phenomenal resources for coding ***when you already know what you are doing***. Using LLMs requires careful skill, attention, practice, and they tend to be useful only in specific limited cases.

Using ChatGPT to write all of your code will not help you learn how to master R or understand the statistical concepts that we are learning in this class. What these models do is essentially amalgamate all of the publicly-available code on GitHub and spit out a plausible looking answer. This is otherwise known as bullshitting. Sure, bullshitting can sometimes get you part of, or, in rare instances, all the way to the right

answer. But more often than not, the code it spits out will contain components that are outright wrong or have stuff you don't need. These models don't care if the "answer" provided is wrong or right (nor can the models actually know). It doesn't care if your code runs or not. It just cares that the crap it spits back at you looks plausible. A hilarious, and slightly terrifying example of this comes from Andrew Heiss:

> Unset
>
> *"A good analogy for this is with recipes. ChatGPT is really confident at spitting out plausible-looking recipes. A few months ago, for fun, I asked it to give me a cookie recipe. I got back something with flour, eggs, sugar, and all other standard-looking ingredients, but it also said to include 3/4 cup of baking powder. That's wild and obviously wrong, but I only knew that because I've made cookies before. I've seen other AI-generated recipes that call for a cup of horseradish in brownies or 21 pounds of cabbage for a pork dish. In May 2024, Google's AI recommended adding glue to pizza sauce to stop the cheese from sliding off. Again, to people who have cooked before, these are all obviously wrong (and dangerous in the case of the glue!), but to a complete beginner, these look like plausible instructions."*

The purpose of this course is to outfit y'all with the concepts and tools needed to accurately and honestly portray the results of hard-won data that we use to understand the way the world works so that we can be better stewards and citizens of our shared planet. I want y'all to succeed in this. I don't want to read a bunch of bullshit generated by an LLM. It's not useful to you. It's not useful to me. It's a waste of time.

All that said, I'm not going to play Sherlock and try and guess whether or not what you're turning in is AI-generated. As I mentioned above, there are legitimate and good use-cases for LLMs in the work that we do such as helping debug some code, interpreting an error message that you don't understand, or helping translate what individual lines of code may be doing if you can't understand from the documentation. However, there's a fairly good chance that I will notice if what you're turning in came from a LMM. There are tells in the way machines write code versus the way most humans do. While this may not impact your grade, it will impact your learning. And it will make me sad. Please don't make me sad.

Phones and electronic devices

Obviously we'll need access to computers for this course, and taking notes during lectures on your computers is fine. However, I would suggest taking real, physical notes given the abundance of research that links this to improved comprehension and learning outcomes. As for cell phone use, please silence them and put them in a bag.

Student conduct

The University seeks an environment that promotes academic achievement and integrity, that is protective of free inquiry, and that serves the educational mission of the University. Similarly, the University seeks a community that is free from violence, threats, and intimidation; that is respectful of the rights, opportunities, and welfare of students, faculty, staff, and guests of the University; and that does not threaten the physical

or mental health or safety of members of the University community. As a student at the University you are expected to adhere to the Student Conduct Code as approved by the Board of Regents. The Student Conduct Code can be found in the UMN Policy Library: https://regents.umn.edu/policy/allboard-of-regents.

Academic integrity

Academic integrity is essential to a positive teaching and learning environment. All students enrolled in University courses are expected to complete coursework responsibilities with fairness and honesty. The University Student Code of Conduct defines scholastic dishonesty as:

> Unset
>
> *"'...submission of false records of academic achievement; cheating on assignments or examinations; plagiarizing; altering, forging, or misusing a University academic record; taking, acquiring, or using test materials without faculty permission; acting alone or in cooperation with another to falsify records or to obtain dishonestly grades, honors, awards, or professional endorsement.'"*

Rarely - especially among upper division courses - have I found this to be a problem. I ***encourage and expect*** students to work together on laboratory assignments; in real practice, you will collaborate or discuss with colleagues the best analytical route to choose. However, ***the work you submit must be your own***. Plagiarism (from classmates or the internet), cheating during tests or exams, or misrepresenting the nature of your involvement in any assigned work may result in disciplinary action, including, but not limited to, a penalty up to and including an "F" or "N" for the course.

Student accommodations

Occasionally, the need arises to employ different testing techniques in assessing student mastery of the subject matter. Determining appropriate disability accommodations is a collaborative process. You as a student must register with Disability Services and provide documentation of your disability; I will then work with Disability Services to determine appropriate accommodations. See http://disability.umn.edu.

Stress management

We live in continually unprecedented times. On top of all the external stressors in our lives, strained relationships, anxiety, depression, substance addiction problems, and more can impede learning, reduce your ability to participate in daily activities, and reduce your academic performance (in addition to just making your life more difficult and miserable!). Please realize that you are not alone and that a broad range of confidential mental health services are available on campus to assist you. Reaching out for help is challenging – but it is a smart and courageous thing to do. http://www.mentalhealth.umn.edu