**AGR 6932 (Special Topics): Hierarchical Bayesian Models for Agricultural Sciences in Stan**

Running Title: Hierarchical models in Stan

Fall 2019: 3 Credits

Friday, Periods 6-8 (12:50-3:50)

McCarty Hall D, G062

**Instructor: Chris H. Wilson, Ph.D.**

2089 McCarty Hall B

[chwilson@ufl.edu](mailto:chwilson@ufl.edu)

**Office Hours:** By appointment

**Course Description**: The purpose of this course is for graduate students to build statistical fluency within the framework of Bayesian Hierarchical Models (BHM), as they apply to solving diverse challenges in the analysis and modeling of ecological and agronomic data. We develop the foundation of Bayesian analysis as an extension of concepts in probability theory, and then proceed from the application of basic linear models to mechanistic, non-linear models that characterize cutting-edge scientific research in most fields of the natural and social sciences. An emphasis on hands-on skill building is present throughout the course, and the ultimate goal is for students to analyze their own research data with some of the techniques and frameworks discussed in this class. Students will learn and consolidate skills in using the Stan probabilistic programming language to express and fit their statistical models ([www.mc-stan.org](http://www.mc-stan.org)).

**Course Prerequisite**: Basic statistics, basic skills in R.

**Course Format**: We will combine short, focused lectures, with discussions of textbook and other readings, and interactive labs. The class material will be available on Canvas, and may include lectures, readings, discussions, websites and videos. All of our applied Bayesian analysis will be conducted within the Stan statistical language/environment ([www.mc-stan.org](http://www.mc-stan.org) ).

**Textbook:** *Bayesian Models: A Statistical Primer for Ecologists.* Hobbs and Hooten (2015), Princeton University Press.

**Additional Reading List (incomplete):**

-Chapters from-

* Statistical Rethinking (McElreath)
* 10 Great Ideas About Chance (Diaconis and Skyrms)
* Ecological Models and Data in R (Bolker)
* Bayesian Data Analysis, 3rd Edition (Gelman et al.)

-Peer-reviewed papers by Hobbs and Hooten, James Clarke, Michael Dietze, Kiona Ogle, others.

-Case-studies hosted by the MC Stan group and Michael Betancourt.

-Video lectures by Richard McElreath (freely available via Github).

**General Course Objectives**:

* Understand basic theory of Bayesian analysis and its relation to maximum likelihood estimation and other classical statistical methods
* Describe and analyze the elements that comprise a full Bayesian hierarchical model as found in the applied literature
* Develop original Bayesian analyses in Stan
* Evaluate and interpret the output of statistical models fitted in Stan

**Course Modules**:

1. *Revisiting the basics*. What is statistical inference anyway? How do we understand and cope with uncertainty and variability in Science? Probability theory as a framework for answering these questions. Basic analysis of probability distributions and their properties. Practical skills: set up and solve basic probability problems, writing tidy scripts in R, setting up and managing a Github repository.
2. *Phenomenological versus Mechanistic Models*. What is a scientific model? What kinds of questions do we answer with linear models and when do we need something more sophisticated? Practical skills: write down and analyze linear models, and several commonly encountered non-linear models. Perform basic simulations combining stochastic and deterministic components.
3. *Bayesian Models: Likelihoods, Priors and Hyperpriors.* How do we link data to our models? What is a likelihood? What is a prior? What happens when we put them together? Practical skills: write down basic statistical models using correct notation, translate models encountered in past or in literature into notation, understand and identify conjugate priors, maximum likelihood analysis as a special case of simplified Bayes.
4. *Models within models: Hierarchical Models.* How do we model multiple levels of variation within our data? What are “random effects” and “fixed effects”? Practical skills: extend fluency developed in previous module to the general case, quantification of “partial pooling”, and application of various metrics of model fit to hierarchical models, ability to write down and estimate hierarchical linear models.
5. *Putting it all together:* State-space modeling framework. How do we combine the tremendous generality and power of Hierarchical Models with the mechanistic, often non-linear models explored previously? Practical skills: write down, fit and evaluate a basic state-space model.
6. *A closer look at MCMC:* Wait, how do we get inferences from models that have no analytical solution? What are the pitfalls to watch out for? Practical skills: describe the mechanics of MCMC, interpret output and diagnostics from Stan.
7. *Inference from single models:* How do check and evaluate a model? What is a good model versus a not-good model? Practical skills: posterior predictive distributions,metrics of model fit (e.g. R2, information criteria, etc.).
8. *Inference from multiple models:* How do we compare and combine models? What is the relationship between predictive ability and scientific insight? Do we ever have the “true” model, and if not, what do we do? Practical skills: model selection, model combination using Bayesian model averaging or predictive stacking.
9. *Case studies:* The rest of course is dedicated to application of all this material to student’s own research.

**Course Grading Scale**: A = 100-94% C = 76.9-73 A- = 93.9-90 C- = 72.9-70 B+ = 89.9-87 D+ = 69.9-67 B = 86.9-83 D = 66.9-63 B- = 82.9-80 D- = 62.9-60 C+ = 79.9-77 E < 60

**Grade Point Distributions:** 8 Homework assignments (1 per content module) (5 pts each)

+ 1 final project (60 pts)

The final project will consist of a complete Bayesian analysis of a dataset of the student’s choosing. It will consist of the following components:

1. A properly notated hierarchical statistical model,
2. Simulation analysis prior to model fitting
3. A fitted model in Stan (including reproducible R and Stan code)
4. Evaluation of model diagnostics and MCMC output
5. Predictive evaluation of model using a variety of metrics
6. Graphical display of fitted model and any inferences
7. Inference on model and interpretation of scientific insight

The final assignment shall be completed entirely within R using Markdown for document creation, and Github as a repository for all text, data, graphics, code, etc.

**Course Evaluation Process:** Student assessment of instruction is an important part of efforts to improve teaching and learning. At the end of the semester, students are expected to provide feedback on the quality of instruction in this course using a standard set of university and college criteria. These evaluations are conducted online at https://evaluations.ufl.edu. Evaluations are typically open for students to complete during the last weeks of the semester; students will be notified of the specific times when they are open. Summary results of these assessments are available to students at <https://evaluations.ufl.edu/results>.

**University Honesty Policy:** UF students are bound by The Honor Pledge which states, “We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of honor and integrity by abiding by the Honor Code. On all work submitted for credit by students at the University of Florida, the following pledge is either required or implied: “On my honor, I have neither given nor received unauthorized aid in doing this assignment.” The Honor Code (https://www.dso.ufl.edu/sccr/process/student-conducthonor-code/) specifies a number of behaviors that are in violation of this code and the possible sanctions. Furthermore, you are obligated to report any condition that facilitates academic misconduct to appropriate personnel. If you have any questions or concerns, please consult with the instructor or TAs in this class.

**Software Use**: All faculty, staff and students of the university are required and expected to obey the laws and legal agreements governing software use. Failure to do so can lead to monetary damages and/or criminal penalties for the individual violator. Because such violations are also against university policies and rules, disciplinary action will be taken as appropriate.

**Campus Helping Resources:**  Students experiencing crises or personal problems that interfere with their general well-being are encouraged to utilize the university’s counseling resources. The Counseling & Wellness Center provides confidential counseling services at no cost for currently enrolled students. Resources are available on campus for students having personal problems or lacking clear career or academic goals, which interfere with their academic performance. www.counseling.ufl.edu/cwc/University Counseling & Wellness Center, 3190 Radio Road, 352-392-1575, Counseling Services