STA6934, Section 1A24: Spatial Statistics, Fall 2012 MWF Period 7 (1:55 -2:45 PM), Room: Weil Hall 0273

Instructor: Dr. Nikolay Bliznyuk (nbliznyuk@ufl.edu), 406 McCarty C Office Hours: TBA after 29-Aug; before that -by appointment. Course web page has been set up on E-learning/Sakai at lss.at.ufl.edu. Anonymous feedback can be submitted at https://forms.hush.com/ufl-2012-spatial.

Course overview: This course is an introduction to spatial statistics, with a focus on methods that are relevant for public health applications, as well as earth and environmental sciences. It is primarily intended for two audiences: (i) statisticians who want to get exposed to methods and applications and (ii) researchers from other fields with some training in statistics that routinely work with spatial data and would like to learn appropriate statistical models and methods. Many of the methods and models are applicable to analysis of general -not necessarily spatially -dependent data, and may appeal to students not immediately interested in spatial statistics per se.

Methods, models and some theory behind will be motivated by applications and supplemented with discussion of implementation of various spatial methods using statistical environment R and, possibly, WinBugs (both freely available). The main goal is to disect methods and models in order to understand them, rather than to apply black-box procedures in existing software to carry out standardized analysis of data.

Lectures will cover the three main areas of spatial statistics: geostatistical (point-referenced) data, lattice (areal) data, and point process data. Tentative list of topics (with possible reordering at the discretion of the instructor) is provided below:

- some theory for point-referenced stochastic processes (e.g., temporal, spatial, general Gaussian)
- spatial point process models (overview)
- areal data models (overview)
- variogram and covariance function estimation
- spatial prediction and kriging
- spatial regression: mixed models, semiparametric models, additive and generalized additive models
- advanced topics: hierarchical models, spatial misalignment, non-stationary models, spatio-temporal processes, multivariate random fields

Objectives of the course are that students be able, by the end of the course, to

- describe the three main areas of spatial statistics: methods for geostatistical, areal, and point process data;
- read and discuss new methods in spatial statistics in the literature based on an understanding of the basic spatial statistics approaches, principles and main assumptions,
- evaluate which methods to use for spatial datasets that may arise in their research; and
- implement various methods using statistical software

Caveat: since I am a statistician, my goal is to cover in depth the areas that would help one to produce work at the level of statistical sophistication appropriate for modern methodological or applied stat/biostat/epi/geostat journals. For a survey course from a geographer's perspective with an emphasis on application of existing methods implemented in modern software, please consider GEO6938 regularly offered by Dr. Tim Fik; check out his syllabus.

Prerequisites/preparation:

- sufficient background: first-year required masters-level coursework in statistics at UF
- minimal background: a solid graduate course in regression (such as STA6207) with exposure to matrix notation; a solid course in inference at the level of STA5328; basic statistical computing skills; motivation (particularly, to pick up R)
- ideal background: in addition to the minimum background above, proficiency with matrix algebra and basic numerical linear algebra (STA6329); statistical computing using R; exposure to linear mixed models, generalized linear models and generalized linear mixed models; masters level sequence in probability and inference STA 6326-6327; interest (and/or need!) to apply methods learned in this course in your research work.

A homework assignment distributed during the first week of classes will give students a good idea of expected background and whether they should be taking this course.

Textbook: no required texts as lecture notes and readings will be posted on the course web page. A number of recommended supplementary texts are on course reserve at Marston Science Library. Some of the texts are listed below

- Schabenberger and Gotway (2004), Statistical Methods for Spatial Data Analysis. Quite readable; we'll cover a number of topics from the first 6 chapters of this text.

- Cressie (1993), Statistics for Spatial Data. Used to be the bible of spatial statistics. Somewhat outdated now but still extremely useful.

- Cressie and Wikle (2011), Statistics for Spatio-Temporal Data. A nice modern complement to Cressie's bible.

- Banerjee, Gelfand, Carlin (2003), Hierarchical Modeling and Analysis for Spatial Data. Good for a Bayesian viewpoint on spatial modeling and fitting using Markov Chain Monte Carlo (using WinBugs). Has separate chapters on areal models and on more advanced topics. Many code examples.

- Wood (2006), Generalized Additive Models: An Introduction with R. A superb text for additive, generalized additive and generalized additive mixed models. A nice complement for the R package mgcv developped by Wood as it has many code examples.

. . . and a few other texts

Grade Policy: tentatively, the final grade will be assigned based on the standard UF scale (90+%: A-or higher; 80+%: B-or higher, etc). Here are the components of the grade:

Homework -40%

Project -40% (tentatively, as follows: proposal 5%, report 20% and presentation 15%)

Final Exam -20%

Participation -up to 10%

There are no typos above, the numbers do add up to 110%, and you can end up with an "AAA" as the course grade. Please keep in mind that the assessment of participation is largely subjective -do not count on getting the whole bonus "for free". I may loosen up the policy as the course moves on, but not make it more stringent. My hope is that there will be plenty of well-deserved As assigned in this course, but that earning an A is not one's primary motivation for taking this course.

Homework: There will be 5-7 homework assignments, with greater frequency in the first half of the course. Some assignments will be more analytical, others will deal with data analysis and implementation of procedures in R. Unless noted otherwise, (i) assignments must be completed by students individually, but group discussion is permitted; (ii) all work that you submit must be your own and you should be able to justify your work (failure to do so typically makes instructors very upset). Additional guidelines (particularly, regarding submission and deadlines) will be provided along with each assignment.

Project: choose one of the following options (or propose an alternative to the instructor)

1 Data analysis option: perform a statistical data analysis illustrating the application of methods discussed in class, or related spatial statistics methods, to real data. Ideally, the data will come from your own line of work.

2 Methods option: study a method or methods of spatial statistics from modern literature not presented in class, implement and/or study/illustrate the method on real or simulated data.

3 Theory option: study and briefly summarize a paper on material not covered in class, critically review and relate to other literature.

In addition to carrying out the project work, a student will need to

1 submit a project proposal (1-2 pages long; must outline your goals and a proposed strategy of accomplishing them), which must be approved by the instructor

2 write a short report in the form of a scientific paper (i.e., containing abstract, introduction, methods, results succinctly summarized using text, tables and figures, discussion, references and supplements including code) summarizing the project; length -8-10 double-spaced pages (not counting supplements)

3 present your work in class (tentatively, during the last 2 weeks of classes)

Depending on the size of the class, work in small groups may be permitted. It is not unreasonable to expect that students working in a group will attempt a more ambitious project than soloists (hint, hint!). You will be assessed on the clarity and correctness of your presentation and quality of your report. Other details will be released after the first week of classes.

Final exam: details will be announced during the first few weeks of classes; potentially, it could be a longer-than-typical-homework take-home assignment.

Participation: I do not want to curb students' creativity in interpreting “participation”; some of the proven ways to participate include being alert and getting involved in class (for which attendance is necessary but not sufficient), making use of the course forum to discuss course material and organizational matters, providing anonymized or open feedback about the course, etc. Participation in these activities will allow the instructor to make a more objective assessment of your participation.

What follows needs to be here according to UF regulations. Make sure you read it carefully.

Academic Honesty: The University of Florida requires all members of its community to be honest in all endeavors. Cheating, plagiarism, and other acts diminish the process of learning. When students enroll at UF they commit themselves to honesty and integrity. Your instructor fully expects you to adhere to the academic honesty guidelines you signed when you were admitted to UF. As a result of completing the registration form at the University of Florida, every student has signed the following statement: I understand that the University of Florida expects its students to be honest in all their academic work. I agree to adhere to this commitment to academic honesty and understand that my failure to comply with this commitment may result in disciplinary action up to and including expulsion from the University. Furthermore, on work submitted for credit by UF students, the following pledge is either required or implied: On my honor, I have neither given nor received unauthorized aid in doing this assignment. It is to be assumed all work will be completed independently unless the assignment is defined as a group project, in writing by the professor. This policy will be vigorously upheld at all times in this course.

Students with Disabilities Act: The Dean of Students Office coordinates the needed accommodations of students with disabilities. This includes the registration of disabilities, academic accommodations within the classroom, accessing special adaptive computer equipment, providing interpretation services, and mediating faculty-student disability related issues. Dean of Students Office, 202 Peabody Hall, 392-7066. www.dso.ufl. edu

Campus Helping Resources: Students experiencing crisis or personal problems that interfere with their general well-being are encouraged to utilize the university's counseling resources. Both the Counseling Center and Student Mental Health provide confidential counseling services at no cost for currently enrolled students. Resources are available on campus for students having personal problems or lacking clear career and academic goals, which interfere with their academic performance. The Counseling Center is located at 301 Peabody Hall (next to Criser Hall). Student Mental Health is located on the second floor of the Student Health Services in the Infirmary.