

MA/ST 810, FALL 2009
Mathematical-Statistical Modeling and Analysis of Complex Systems

Instructors:

H.T. Banks, 304 Cox Hall, 515-8968, htbanks@unity.ncsu.edu

Marie Davidian, 5124 SAS Hall, 515-1940, davidian@stat.ncsu.edu

Class meeting: T 3:30–6:00 pm, SAS Hall 1108

Instructors' office hours: TBA

Course Description: A current “hot topic” in biological, biomedical, and physical sciences; the pharmaceutical industry; engineering; epidemiology; and a host of other areas is the use of complex nonlinear dynamical systems models to characterize phenomena of scientific interest. These models involve (mostly) deterministic systems of ordinary and partial differential equations that are used to characterize interactions among various components of a biological, physical, or other system.

For example, there is much current interest in using these models to describe known and hypothesized mechanisms involved in the interplay over time between the HIV virus and the immune system taking place within organisms (like humans), understanding of which has the potential to suggest new treatment strategies. In general, such models allow a formal description of mechanisms in terms of parameters (some of which may be functions rather than real-valued). In the HIV example, parameters might include viral clearance rates, immune cell infection rates, and new virus production rates.

The solution of the system of differential equations implies an expression (that is likely not available in a closed form) for quantities of interest at any time point. For example, a system to model HIV dynamics determines expressions for the concentration of virus (or “viral load”), concentration of T-cells (an immune system component), etc., present in the organism at any time following administration of antiretroviral therapy. Given the values of the parameters that describe the system and some initial conditions, evaluating such expressions must in general be carried out via complex numerical techniques; this is referred to as “forward solution.”

As development of these models and computational strategies for evaluating them have improved, there is great interest in applying them to data. For example, data on viral load and measures of immune status may be collected longitudinally on one or more organisms; that is, values of the system at certain time points have been observed (possibly subject to variation due to measurement error and other sources). From such data, it is of interest to learn about the underlying mechanisms (values of parameters) for an individual organism and more generally how the mechanisms (values of parameters) vary across the population of organisms.

Thus, for application to data from the system corresponding to a single organism, the problem is the reverse of “forward solution:” given (possibly error-prone) values taken on by system components at known time points, can the values of the parameters be recovered, knowledge of which may provide insight into the underlying mechanisms? Recovering the values of the parameters is a form of what is known in this literature as an “inverse problem.” Statisticians refer to this as “parameter estimation,” and are concerned about taking proper account of the various sources of variation, development of valid inferential strategies, and efficiency of estimation. This requires that the complex dynamical system be embedded in a statistical framework appropriate to the data structure. To address the issue of how mechanisms (parameters) vary in the population requires an additional layer of statistical modeling to represent this variation.

An important outgrowth of this methodology that has engendered considerable recent interest in biomedical and pharmaceutical research is the use of these frameworks as a basis for simulation and experimental design. For example, the US Food and Drug Administration, through its Critical Path Initiative and more generally, has highlighted the urgent need to bring new technologies to bear on the development of new pharmaceutical and biologic entities. A key strategy that has been embraced by both the agency and industry is the integration of mathematical modeling of biological systems with statistical modeling of population behavior as a foundation for investigating the potential efficacy and effectiveness of new drugs in the development pipeline and for designing clinical confirmatory studies.

This course is meant to give applied mathematics, biomathematics, statistics, and other quantitatively oriented PhD students the necessary background to appreciate both the “big picture” and some of the specific mathematical and statistical challenges in this endeavor. For statistics students, the course will serve as an introduction to differential equation modeling, assuming no prior exposure, focusing on the aspects that are most important for statisticians to understand. For applied mathematics and biomathematics students, the course will serve as an introduction to statistical inference and considerations involved in statistical modeling (e.g. the need to take appropriate account of multiple sources of variation in complex data structures). For both groups, the course will serve as an introduction to the important area of modeling and simulation. These topics will be highlighted by discussion of subject-matter examples from numerous fields of application.

The diversity of the audience will provide a unique forum for students to learn about this area and relevant applications and to understand the perspectives of applied mathematicians and statisticians and the terminology each group uses. Students will also gain experience in interdisciplinary science and collaboration.

Prerequisites:

For mathematics and biomathematics students: Undergraduate-level differential equations (required), MATLAB or similar computing experience (required), Numerical analysis (desirable).

For statistics students: Statistical inference (at the level of Casella and Berger; required), Linear models theory (required), R, MATLAB, or similar computing experience (required).

(Mathematics and biomathematics students need not have statistics student prerequisites, and vice versa. Please contact one of the instructors if you are unsure about the suitability of your background.)

Class Web Page: Assignments and other course material will be available at

<http://www4.stat.ncsu.edu/~davidian/stma810c/index.html>.

Important announcements made in class will also be posted here.

Grading: Grading will be S/U. There will be no tests. Course work will consist of a series of assigned projects involving modeling and data analysis. Students will be placed in teams such that each team has representation from at least two of applied mathematics, biomathematics, statistics, and other quantitative disciplines. Each team will collaborate to complete each assignment and turn in a joint report. Grading will be based on active, satisfactory participation in team collaborations and contribution to and completion of all reports.

Text: There is no textbook. The instructors will present course material from their own notes, copies of which will be distributed in class.

Course Topics: The following is a tentative outline for the course; this is subject to adjustment. The initials of the instructor lecturing are in parentheses.

08/25 Introduction to mathematical modeling (HTB)
09/01 Inverse problems (HTB)
09/08 Integrating mathematical and statistical models (MD)
09/15 Statistical models and inference (MD)
09/22 Statistical models and inference (MD), Sensitivity analysis (HTB)
09/29 Generalized sensitivity and experiment design
10/06 Statistical inference (MD)
10/13 Inference with censored data (MD), Hierarchical statistical models (MD)
10/20 Hierarchical statistical models (MD)
10/27 Methods for aggregate data (HTB)
11/03 Methods for aggregate data (HTB), Modeling and simulation (MD)
11/10 Modeling and simulation and its use in study design (MD)
11/XX* Modeling/simulation case study (HTB/MD)
11/24 NO CLASS
12/01 Advanced topics (HTB/MD)

*The class meeting on November 17 will be moved to another date/time due to the travel schedules of both instructors.

Course Evaluations: Online class evaluations will be available for students to complete during the last two weeks of class. Students will receive an email message directing them to a website where they can login using their Unity IDs and complete evaluations. All evaluations are confidential; instructors will never know how any one student responded to any question, and students will never know the ratings for any particular instructors.

Academic Integrity: The instructors expect that students will abide by the University policy on academic integrity found in the Code of Student Conduct Policy (POL11.35.1), available at http://www.ncsu.edu/policies/student_services/student_discipline/POL11.35.1.php. Students are expected to uphold the standards of honesty set forth in this Code in completion of all class assignments.