ABE 6933 Data Diagnostics

ABE 6933 Section 2F86
Time: Monday 2nd-4th Period (8:30-11:30)
Location: Frazier Rogers Hall 283
Fall 2018

Instructor

Ray Huffaker
Agricultural and Biological Engineering
Frazier Rogers Hall 281
rhuffaker@ufl.edu
(352)392-1864 x281
Office Hours: Mondays 1-4pm or by appointment

Prerequisites

calculus through differential equations, knowledge of linear algebra, experience with computer programming, or consent of instructor

Course Description

Application of nonlinear time series analysis to detect, characterize, and model deterministic structure in real-world time series data. Topics include signal processing, phase space reconstruction, surrogate data testing, causal network analysis, and phenomenological modeling.

In the process of data analysis, the investigator often observes highly volatile and random-appearing data. A common assumption is that observed volatility is due to underlying stochastic processes, but this is not necessarily the case. Nonlinear time series analysis (NLTS) allows researchers to test whether observed volatility conceals systematic nonlinear behavior, and to rigorously characterize governing deterministic dynamics. Behavioral patterns detected by NLTS, along with scientific principles and other expert information, guide the specification of mechanistic models that serve to 'ground-truth' and explain real-world behavior.

Course Objectives

The tools of NLTS were developed in mathematics and physics. This course helps non-mathematicians in the applied sciences, engineering, economics, and other social sciences to become operational with NLTS. Students acquire background knowledge of nonlinear dynamics required to apply NLTS in a sophisticated manner. Students gain hands-on experience with NLTS so that they can apply it confidently to diagnose the dynamic forces driving volatile real-world data.

These objectives will be accomplished through:

1. A 'workshop' classroom format emphasizing 'knowledge through discovery': Students read assigned introductory material on scheduled topics before class. The instructor begins class by reviewing this material, punctuating it with intuitive examples and real-world applications, and answering questions. Students then spend the majority of class time running prepared computer experiments under the instructor's direct supervision to gain hands-on experience with NLTS.
2. Detailed **R** code provided for computer experiments: The **R** code used in computer experiments is explained in detail both by the instructor in the classroom, and by required readings. This allows students to adjust the code for use in their own work.
3. An explicit framework for applying NLTS methods to real-world time series data: The framework is condensed from sound empirical practices recommended in the literature. Students become `data detectives', accumulating hard empirical evidence directing scientific inquiry.

4. Homework projects that apply NLTS diagnostics to real-world time series data: Classroom computer experiments are supplemented with homework projects giving students increased hands-on experience with real-world data diagnostics.
5. Evaluation of student skills with hybrid written and oral final presentation: Examination tests the extent to which students can apply NLTS methods to real-world data, and correctly interpret diagnostics results. Each student provides a written report on data diagnostics, and further presents an oral defense of diagnostics and conclusions to the class (15 minute presentation).

Required Textbook

- R. Huffaker, M. Bittelli, and R. Rosa. Nonlinear Time Series Analysis with R. Oxford University Press.

The required textbook is coauthored by the instructor. The instructor will receive no financial benefit from sales to students in the class. There are no additional fees for this course.

Course Materials

Course materials are available to students in Canvas. Materials include lecture notes and slide presentations, the **R** code used in classroom computer experiments, homework assignments, and examinations. Students are required to bring their personal laptops to class, download the most recent version of **R**, and dedicate time outside of class to familiarize themselves with programming basics in **R**.

Course Schedule

Aug 27: What is phase space? Why study nonlinear time series analysis?
- Read Huffaker, Bittelli, and Rosa (HBR), Chapter 1
- Phase Space Reconstruction: Read HBR, Section 1.2

Sep 3: Holiday

Sep 10: Data Preprocessing
- Introduction: Read HBR, Section 6.1
- Regular Behavior of Linear ODE Models: Read HBR, Sections 6.2-6.3, Appendix C

Sep 17: Data Preprocessing continued
- Signal Processing with Singular Spectrum Analysis: Read HBR, Section 6.4
- HOMEWORK 1

Sep 24: Testing for Nonlinear Stationarity in Time Series Data
- Change Point Analysis, Read HBR, Sections 6.6, 8.6
- HOMEWORK 2

Oct 1: Phase Space Reconstruction with Time-Delay Embedding, Read HBR, Section 6.6.1

Oct 8: Surrogate Data Testing
- Introduction: Read HBR, Sections 7.1-7.2
- Surrogate Types: Read HBR, Section 7.3
- Discriminating Statistics: Read HBR, Section 7.4
- Rank-Order Statistics: Read HBR, Section 7.5

Oct 22: Surrogate Data Testing continued
- R-code for Surrogate Data Testing: Read HBR, Sections 7.6-7.7
- HOMEWORK 3

Oct 29: Empirically Detecting Causality
- Convergent Cross Mapping
- Read HBR, Sections 8.1-8.3

Nov 5: Convergent Cross Mapping continued
- Network Plots: Read HBR, Section 8.4
- Application to disease epidemics: Read HBR, Section 8.5

- HOMEWORK 4

Nov 12: Holiday

Nov 19: Measuring/Characterizing Causal Interactions with S-Mapping

- HOMEWORK 5

Nov 26: Empirically Detecting Causality with Heterogeneous Data

- Change-point Detection: Read HBR, Section 8.6
- Tipping-point Detection: Read HBR, Sections 8.7-8.8
- HOMEWORK 3

Dec 3: Project Presentations

Homework Assignments and Examinations

There are five homework exercises that are due the week after assignment. Students are encouraged to collaborate on homework assignments, but must turn in their own work. The final examination requires a written report applying course methods to a set of data selected by the student, and an oral presentation to the class (15 minutes). Students can consult with each other on the written report, but must do their own work.

Evaluation of Grades

Assignment	Total Points	Percent of Grade
Homeworks (5)	100 (20 points/homework)	50%
Final (Written)	60	30%
Final (Oral)	40	20%
Total	200	100%

Grading Policy

A (269-300 points, 90-100%,); A- (254-268 points, 85-89%); B+ (239-253 points, 80-84%);
B (224-238 points, 75-79%); B- (209-223 points, 70-74%); C+ (194-208 points, 65-69%);
C (179-193 points, 60-64%); C- (164-178 points, 55-59%); D+ (149-163 points, 50-54%);
D (134-148 points, 45-49%); D- (119-133 points, 40-44%); E (0-118 points, 0-39%)
More information on UF grades and grading policies is located at:
http://catalog.ufl.edu/ugrad/current/regulations/info/grades.aspx

Class Attendance and Make-Up Policy

Class attendance is essential for students to benefit from the classroom workshop approach. Students should arrange with instructor for make-up material. General UF policy can be found at: https://catalog.ufl.edu/ugrad/current/regulations/info/attendance.aspx.

Online 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 two or three 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.

Academic Honesty

As a student at the University of Florida, you have committed yourself to uphold the Honor Code, which includes the following pledge: "*We, the members of the University of Florida community, pledge to hold* *ourselves and our peers to the highest standards of honesty and integrity."* You are expected to exhibit behavior consistent with this commitment to the UF academic community, and on all work submitted for credit 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*."

It is assumed that you will complete all work independently in each course unless the instructor provides explicit permission for you to collaborate on course tasks (e.g. assignments, papers, quizzes, exams). Furthermore, as part of your obligation to uphold the Honor Code, you should report any condition that facilitates academic misconduct to appropriate personnel. It is your individual responsibility to know and comply with all university policies and procedures regarding academic integrity and the Student Honor Code. Violations of the Honor Code at the University of Florida will not be tolerated. Violations will be reported to the Dean of Students Office for consideration of disciplinary action. For more information regarding the Student Honor Code, please see: http://www.dso.ufl.edu/sccr/process/student-conduct-honor-code.

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.

Services for Students with Disabilities

The Disability Resource Center coordinates the needed accommodations of students with disabilities. This includes registering disabilities, recommending academic accommodations within the classroom, accessing special adaptive computer equipment, providing interpretation services and mediating faculty-student disability related issues. Students requesting classroom accommodation must first register with the Dean of Students Office. The Dean of Students Office will provide documentation to the student who must then provide this documentation to the Instructor when requesting accommodation:
0001 Reid Hall, 352-392-8565, **www.dso.ufl.edu/drc/**

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.

- *University Counseling & Wellness Center, 3190 Radio Road, 352-392-1575,* **www.counseling.ufl.edu/cwc/**
  Counseling Services
  Groups and Workshops
  Outreach and Consultation
  Self-Help Library
  Wellness Coaching
- U Matter We Care, www.umatter.ufl.edu/
- Career Resource Center, First Floor JWRU, 392-1601, www.crc.ufl.edu/

Background Readings

- Nonlinear Time Series Analysis

Kantz, H. and Schreiber, T. (1997). Nonlinear Time Series Analysis. Cambridge University Press, Cambridge, England.

Schreiber, T. (1999). Interdisciplinary application of nonlinear time series methods. Physics Reports, 308, 1-64.
Williams, G. (1997). Chaos Theory Tamed. John Henry Press, Washington, DC.

- Why Study Nonlinear Time Series Analysis?

Huffaker, R., Berg, E., and Canavari, M. Reconstructing deterministic economic dynamics from volatile time series data, in the Handbook of Agricultural Economics, A. Schmitz and G. Cramer, edt, Routledge Press, Francis and Taylor Publisher, Oxford, England.

- Data Preprocessing
  - Signal Processing with Singular Spectrum Analysis

Elsner, J. and Tsonsis, A. (2010). Singular Spectrum Analysis. Plenum Press, New York.
Golyandina, N., Nekrutkin, V., and Zhigljavsky, A. (2001). Analysis of Time Series Structure. Chapman & Hall/CRC, New York.
Hassani, H. (2007). Singular spectrum analysis: methodology and comparison. Journal of Data Science, 5, 239-257.

  - Testing for Nonlinear Stationarity in Time Series Data

Schreiber, T. (1997). Detecting and analyzing nonstationarity in a time series with nonlinear cross predictions. Physical Review Letters, 78, 843-846.

  - Phase Space Reconstruction with Time-Delay Embedding

Kot, M., Scha_er, W., Truty, G., Graser, D., and Olsen, L. (1988). Changing criteria for imposing order. Ecological Modeling, 43, 75-110.

- Surrogate Data Testing

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., and Farmer, J. (1992). Testing for nonlinearity in time series: The method of surrogate data. Physica D, 58, 77-94.
Schreiber, T. and Schmitz, A. (2000). Surrogate time series. Physica D, 142, 346-382.
Small, M. and Tse, C. (2003). Detecting determinism in time series: The method of surrogate data. IEEE Transactions on Circuits and Systems, 50, 663-672.

- Empirically Detecting Causality
  - Convergent Cross Mapping

Sugihara, G., May, R., Hao, Y., Chih-hao, H., Deyle, E., and Fogarty, M. (2012). Detecting causality in complex ecosystems. Science, 338, 496-500.
Ye, H., Deyle, E., Gilarranz, L., and Sugihara, G. (2015). Distinguishing time-delayed Causal interactions using convergent cross mapping. Scientific Reports (Nature), 5, Article 14750.

  - Change-Point Detection

Ide, T. and Inoue, K. (2005). Knowledge discovery from heterogeneous dynamic systems using change point correlations. In Proceedings of 2005 SIAM International Conference on Data Mining (SDM 05), Newport Beach, CA (ed. H. Kargupta, J. Srivastava, C. Kamath, and A. Goodman), pp. 571-576. SIAM, Philadelphia.

  - Tipping-Point Detection

Lenton, T. and Livina, V. (2016). Detecting and anticipating climate tipping points. In Extreme Events: Observations, Modeling, and Economics (ed. M. Chavez, M. Ghil, and J. Urrutia-Fucugauchi), AGU Geophysical Monograph Series, Volume 214, pp. 51{62. Wiley, Hoboken, NJ.

- Phenomenological Modeling

Brunton, S., Proctor, J., and Kurtz, J. (2016). Discovering governing equations from data by sparse identification of nonlinear dynamical systems. Proceedings of the National Academy of Sciences of the USA, 113, 3932-3937.
Dong, L. Wei, Ren, K., Cluzel, S. Meunier Guttin, and Gouesbet, G. (2015). Global vector-field reconstruction of nonlinear dynamical systems from a time series with SVD method and validation with Lyapunov exponents. Chinese Physics, 12, 1366-1373.

- Extreme Value Statistics

Katz, R. (2010). Statistics of extremes in climate change. Climate Change, 100, 71-76.
Katz, R., Brush, G., and Parlange, M. (2005). Statistics of extremes: modeling ecological disturbances. Ecology, 86, 1124-1134.

- Applications

Huffaker, R. (2015). Building economic models corresponding to the real world. Applied Economic Perspectives and Policy, 37, 537-552.
Huffaker, R., Canavari, M., and Munoz-Carpena, R. (2016). Distinguishing between endogenous and exogenous price volatility in food security assessment: An empirical nonlinear dynamics approach. Agricultural Systems. doi:10.1016/j.agsy.2016.09.019.

Huffaker, R., Munoz-Carpena, R., Campo-Bescos, M., and Southworth, J. (2016). Demonstrating correspondence between decision-support models and dynamics of real-world environmental systems. Environmental Modeling and Software, 83, 74-87.