**STOR 890: Topics in Time Series Analysis Spring, 2017**

**Course goals and key learning objectives:** The main goal of this course is to introduce students to a range of both classical and more modern topics in time series analysis that can be used as a foundation for and understanding of both theoretical and applied research work in a range of fields (STOR, Computer Science, Economics, Psychology/Neuroscience, and others). Time series analysis is understood in a broad sense as the analysis of data collected in time, and is related to Signal Processing, Machine Learning, Stochastic Processes, and other areas. No less important goal is also to have fun and a good time, in an environment that is conducive to learning!

**Target audience:** Graduate students in STOR, Biostatistics, Computer Science, Economics, Psychology/Neuroscience and other departments.

**Place and Time**: Hanes 125, Tuesday & Thursday, 12:30-1:45PM.

**Instructor**: Vladas Pipiras, Professor in the Department of Statistics and Operations Research; office: Hanes 305; e-mail: pipiras@email.unc.edu; phone: 843-2430.

**Instructor office hours:** Office hours will be held by appointment: If you need to meet with me, feel free to email me and we will set up a time. I am usually in the department every weekday. If many of you would like me to have regularly scheduled hours, I can arrange this too. I very much encourage you to ask questions on anything that may be unclear during classes as well.

**Prerequisites**: There is no official prerequisite for the course but I will assume that you have seen basic linear algebra (matrix manipulation, eigenvectors, eigenvalues, etc.), basic probability (multivariate random variables, covariance, correlation, etc.) and statistics (multivariate regression, maximum likelihood estimation, hypotheses testing, etc.). Familiarity with real analysis (especially Fourier transforms) might also be helpful for certain topics. From the computing side, I expect that you are familiar with R or Matlab (or can pick up their basics quickly).

**Course and lecture format:** The classes might be a bit different from what you are usually used to, certainly in my previous courses! The focus will be on understanding the basic principles and ideas, as well as on data, as opposed to technicalities of the theory. Here are several goals/highlights of the approach:

* I will be using R Markdown (and perhaps other project) files while teaching, where approaches will be illustrated (mostly) using R. In particular, you will have access to the data and R Markdown files after lectures, with which you can play around or experiment with any ideas or questions you have.

* There will ideally be at least one motivating data set for each topic. It may not necessarily be the most exciting or "best" for the considered topic (keep in mind that I am teaching like this for the first time), but techniques will be illustrated on the data set and the topic will conclude with something achieved for the data set, even if it appears of "limited" significance.

* My main goal is that at the end of each topic, you would have a good idea of what the topic is all about (maybe without necessarily understanding whatever technical details I happen to present or are out there), and ideally that you would be able to explain it in your own words to someone else.

It is the first time that I am experimenting with this teaching approach, and I am not completely sure how well (or badly!) it will work. Any of your feedback and suggestions (to implement during the semester or for future generations) will be greatly appreciated as we move along.

**Topics:** Below is a list of topics that I would like ideally to cover, separated into four rough themes (univariate, multivariate, nonstationary, and nonlinear). In a good scenario, each topic will be covered in one class, and they can be understood somewhat separately.

*Univariate:*
Classical decomposition
Differencing and unit roots
Structural modeling
Spectral perspective

*Multivariate:*
General VARs
Variable selection, LASSO, etc.
Cointegration
Graphical models and structural VARs
Factor and related models
Clustering/Classification
Spatio-temporal models
Functional data perspective

*Nonstationary:*
Periodic/seasonal models
Change point detection
Time frequency perspective, spectrogram, short FT
Wavelets
Other transformations and representations
Locally stationary processes
Anomaly detection

*Nonlinear:*
Integer-valued time series
Threshold, bilinear and other nonlinear models
ARCH and heteroscedasticity
Hidden Markov models
Neural networks
Deep learning

I already have something prepared for the univariate and multivariate themes, feel quite strongly about the topics in the nonstationary theme, but the topics in the nonlinear theme might still change. I have thought about a number of these topics in my own research, but certainly not about all of them. But I would like to think (wishfully?) that because of my age and hence longer research exposure, I can give you a good description of any of the topics above after looking through the literature.

**Course website**: Though I will keep the course website "published" at http://sakai.unc.edu, most of the material will be posted on Dropbox, which is more convenient for me to use. A Dropbox link for download will be posted on Sakai. You do not need to have a Dropbox account.

In particular, the lecture "slides" will be posted on Dropbox at least one hour before they are given in class, in case you want to make notes on your tablet, print them, follow them on your own computer, etc.

**Textbook**: There are no required textbooks for the course. The lectures will be based on a number of sources, whose references will be provided and some of which will be posted on Dropbox.

**Course requirements and grades:** Your performance in the course will be assessed based on "homework" or/and project, which are described in greater detail below. The grades will be given as follows:

H: Both satisfactory homework and project must be submitted.
P: A satisfactory project must be submitted.
L: Attendance.
F: Hopefully unnecessary.

If the work is deemed unsatisfactory, this will lower the grade by one scale per homework and project. There will be no exams in this course.

**Homework**: Each lecture will have a few questions, most often formulated in a somewhat open-ended fashion. As the homework requirement, I will ask to answer one of these questions per theme, thus the total of four questions (four homework assignments) for the whole course. If you wish, you can also pursue your own questions but with my permission. I expect that answering each question will require for you to do a bit of "digging" on your own. Your answers to questions can come in various forms: you can use a computer software like R, provide a technical argument or proof, etc. I expect your answers to be in a project form like R Markdown or typed up in Latex. Each of the four homeworks will be due about 1-2 weeks after the corresponding theme is concluded.

**Project:** This is a more substantial assignment asking you to report on a particular topic or/and a paper or/and time series data. In either case, the focus should be on data, simulations which are supplemented with theory, and the final "product" should be an R Markdown file, similar to what is used for lectures. For your projects, you can work in teams of up to 2-3 people (with my permission if 3 people). I might ask you to present some of the projects at the end of the semester. Below is some other information related to projects:

*Important dates/deadlines:*
February 14: Decide on your project and get my approval.
March 6-10 week: First discussion with me of your progress.
April 3-7 week: Second discussion with me of your progress.
April 20: Final submission.

Obviously, earlier "compliance" with the deadlines is perfectly fine and is encouraged.

*Possible topics and their sources:*
I am quite open to different suggestions for projects, as long as they are related to the course. For one, I am looking for volunteers to look into the last three listed topics, Hidden Markov models, neural networks and deep learning, as three different projects.

Otherwise, a project could be on a topic related to your own work, or something you find/found interesting in the course. Below is a list of several large conferences that has had papers related to time series analysis:

Joint Statistical Meetings; High-Dimensional Time Series in Macroeconomics and Finance; International work-conference on Time Series (ITISE); Joint Statistical Meetings (JSM); NBER NSF Time Series Conference; NIPS

A few journals where papers related to time series analysis can be found:

Annals of (Applied) Statistics; International Journal of Forecasting; JASA; Journal of Econometrics; Journal of Financial Econometrics; Journal of Machine Learning Research; Journal of Time Series Analysis; Journal of Time Series Econometrics

If you want to find some time series data, google and be creative (e.g. try "Google Trends", ask nicely your classmates doing more applied work, and so on).

Finally, if you do not know what to pursue, I may also suggest a few topics/papers in discussion with you.

**Syllabus changes:** I reserve the right to make changes to the syllabus, due dates and other information, when circumstances demand. These changes will be announced as early as possible so that students can adjust their schedules. (See also the "last modified" date at the bottom of the syllabus.)

**Acknowledgements:** The course preparation has benefited from the support of Data @ Carolina, the STOR PhD student Dylan Glotzer and the UNC IT computing support person Sandeep Sarangi.

*Syllabus last modified: January 12, 2017*