Spring 2020
Department of Mathematical Sciences
Program in the Atmospheric Sciences

Seminar: ATM SCI 950 (Class # 54512–001)

TOPICS IN STATISTICAL ANALYSIS AND INTERPRETATION OF GEOPHYSICAL DATA SETS. PART II: SIGNAL DETECTION

Class meets: TR 11:00–12:15, EMS W434
Instructor: Professor S. Kravtsov

Objectives and method. This course will cover fundamentals of signal detection in noisy multivariate space-time data sets. One lecture per week will be devoted to theoretical considerations, followed by a lab, in which students will analyze actual climatic data sets using MATLAB software. Throughout the semester, each student will also work on individual research project — either an original study (ideally, related to this student's thesis work), or otherwise a thorough examination of a published study; in either case, actual data analysis using the techniques covered in class will be absolutely required. The students will report their results in conference-style presentations at the end of the semester (20-min talk + 5-min Q&A).

Pre-requisites. This course is designed for graduate-student level. Basic knowledge of statistics, linear algebra, calculus and MATLAB is required (in general, it is sufficient to successfully complete part I of this two-part lecture series). Knowledge of fundamentals of Atmospheric Science is a plus. For further information contact Sergey Kravtsov, kravtsov@uwm.edu, EMS W441.

Text. Not required — all necessary materials will be provided. A selected bibliography is given below. Wilks (1995) and von Storch and Zwiers (1999) discuss in depth applications of various statistical methods to problems in atmospheric and climate science. A good summary of basic statistics, linear matrix operations, spectral analysis and regression techniques can also be found in Numerical Recipes (Press et al. 1994). A number of very good related courses can also be found on the web. A notable example is "Objective Analysis" class by Prof. D. L. Hartmann (http://www.atmos.washington.edu/~dennis: go to ATMS 552 and click on "Class Notes").

Select Bibliography:

Box, G. E. P., G.M. Jenkins, and G.C. Reinsel, 1994: *Time Series Analysis, Forecasting and Control*. Prentice Hall, Englewood Cliffs, NJ, 3rd Edition, 592 pp.

Elsner, J. B., and A. A. Tsonis, 1996: *Singular Spectrum Analysis: A New Tool in Time Series Analysis*. Plenum, 164 pp.

Ghil M., R. M. Allen, M. D. Dettinger, K. Ide, D. Kondrashov, M. E. Mann, A. Robertson, A. Saunders, Y. Tian, F. Varadi, and P. Yiou, 2002: Advanced spectral methods for climatic time series. *Rev. Geophys.*, 40(1), pp. 3.1-3.41, 10.1029/2000RG000092.

Preisendorfer, R. W., 1988: *Principal Component Analysis in Meteorology and Oceanography.* Elsevier, Amsterdam, 425 pp.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 1994: *Numerical Recipes*. 2-nd edition. Cambridge University Press, 994 pp.

Von Storch, H., and F. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press, Cambridge, United Kingdom, 484 pp.

Von Storch, H., and A. Navarra, eds., 1999: *Analyses of Climate Variability — Application of Statistical Techniques*. 2nd Edition, Springer-Verlag, Berlin.

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences*. (International Geophysics Series, v. 59), Academic Press, San Diego, 467pp.

Evaluation:

- Weekly lab assignments (due before the next lab, points taken off afterwards) — 70% of the grade

- Final project REPORT (concise problem description, method of solution, results w/figures, appendix w/code) + PRESENTATION — 20%+10%

- **Grading Scale**: Minimum cutoffs

A	A–	B+	B	B–	C+	C	C–	D–	F
93	90	88	83	80	78	73	70	60	<60

Office Hours (EMS W441)**:** TR1:30–3:00pm. Please feel free to drop by at any other time — no appointment necessary. *E-mail* inquiries are also welcome and in fact *encouraged*: to formulate a concrete question in your e-mail will force you to think about the problem at hand and may by itself get you much closer to the solution.

Tentative schedule:

<u>Part I: Patterns in space</u>

Jan. 21, 23: Introduction. Review of background material: Eigenvalue problems. Matrix operations. Data sets as two-dimensional matrices. Dispersion matrix of a data set.

Jan. 28, 30: Empirical Orthogonal Functions (EOFs) computed via eigenvalue analysis of the data set's dispersion matrix. *Lab*: Data preparation and visualization. Techniques for filtering out seasonal cycle.

Feb. 4, 6: Principal components (PCs) as projections of EOFs onto the data. Orthogonality of PCs' time series. Manipulation of EOFs and PCs. *Lab*: EOFs as efficient representations of data sets.

INDIVIDUAL-PROJECT SELECTION DEADLINE (*Please see me ASAP during the first two weeks of the semester to decide on your research topic and work out a detailed research plan. Ideally, I would like to be meeting with each of you individually once a week for about half an hour or so [and longer if necessary] to check on your progress and discuss further work*)

Feb. 11, 13: Review of Singular Value Decomposition (SVD). EOF analysis via SVD of the data matrix. *Lab*: Regional EOF analyses and teleconnection patterns.

Feb. 18, 20: Factor analysis. Rotation of EOFs and PCs. Introduction to Maximum Covariance Analysis (MCA). *Lab*: Statistical significance of EOF modes: Monte-Carlo approach.

Feb. 25, 27: No regular classes: Catch up with labs, start working on individual projects.

Mar. 3, 5: Maximum Covariance Analysis (MCA) and Canonical Correlation Analysis (CCA). *Lab*: Combined EOF analysis of a pair of climatic data sets: Looking for coupled modes.

Mar. 10, 12: Summary: Applications and interpretation of EOF, MCA and CCA analyses. *Lab*: Application of CCA analysis to a pair of climatic data sets: Further search for coupled signals.

Mar. 17, 19: NO CLASSES: SPRING BREAK.

Part II: Patterns in time

Mar. 24, 26: Discrete Fourier transform. Continuous power spectrum. *Lab*: Power spectrum via least-square fitting of harmonic predictors. Plotting the power spectrum.

Mar. 31, Apr. 2: Data windows and window carpentry using single and multiple tapers. *Lab*: Computing power spectrum in MATLAB: Welch's method and Multi-taper method (MTM).

Apr. 7, 9: Statistical significance of spectral peaks. *Lab*: Monte-Carlo testing of spectral peaks.

Apr. 14, 16: Singular-spectrum analysis (SSA) and its multivariate version (M-SSA). *Lab*: SSA and M-SSA application to climatic time series.

Apr. 21, 23: More on the methods for detection of space-time signals: Extended EOFs (EEOFs), Frequency-domain EOFs (FDEOFs), complex EOFs (CEOFs). Summary and outlook.

April 28, 30: WORK ON INDIVIDUAL PROJECTS.

May 5, 7: WORK ON INDIVIDUAL PROJECTS. FIRST-DRAFT REPORTS ARE DUE.

May 12, 14: PRESENTATIONS OF INDIVIDUAL PROJECTS.

May 15: FINAL TERM-PROJECT REPORTS ARE DUE.