Fall 2023

Program in the Atmospheric Sciences

ATM SCI 500/500G (Class # 15403/15404)

TOPICS IN STATISTICAL ANALYSIS AND INTERPRETATION OF GEOPHYSICAL DATA SETS
(STATISTICAL METHODS IN ATMOSPHERIC SCIENCES)

Class meets: TuTh 10:00–11:15, NW Quad Bldg D 1990
Instructor: Professor S. Kravtsov

Objectives and method. The purpose of this course is to introduce basic statistical concepts and develop a working knowledge of a number of statistical methods currently used for analysis, interpretation and modeling of weather/climate-related, observed or model-generated data sets. The lecture presentations will be based on a balance between mathematical rigor in derivation of various statistical techniques and the necessity to cover a fairly large (although by no means complete) set of analysis methods. A particular attention will be paid to the question of how to choose and apply (an) appropriate statistical method(s) depending on the nature of the phenomenon under consideration. Each topic covered in the lecture presentations will be complemented by exercises using synthetic and real data sets in practical classes. Course materials are available from Canvas.

Pre-requisites. This course is designed for advanced undergraduate and beginning graduate student level. Knowledge of Calculus and/or MATLAB is a plus, but not required. For further information contact Sergey Kravtsov, kravtsov@uwm.edu, GLRF 3003E.

Text. Not required — typed lecture notes will be provided. These notes rely heavily, in presentation of selected topics, on the work of Von Mises (1964), Press et al. (1994), and Prof. D. Hartmann's lecture notes (see the reference below). Some parts of these notes are in fact direct duplication, or minor rewording of the above texts. These notes should not, therefore, be considered as the original presentation; rather, the material from different textbooks has

been compiled here in a specific order and augmented by the author's comments. In principle, no other text, beside these notes, is required for the students to fully understand the material and succeed in this class. However, see the notes' bibliography for selected references. For example, 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). There also exist a number of online statistical texts (lecture notes, online courses, statistical manuals). Finally, MATLAB's statistics toolbox contains a brief description and illustration of its intrinsic functions and analysis methods.

Statistical software and practical exercises. Practical exercises will be done mostly using MATLAB. However, no prior MATLAB experience is necessary.

Course requirements and evaluation:

• Attendance and participation in class — 25%

• Homework assignments (solutions to theoretical problems can be handwritten, the results of numerical work are expected to be presented in the form of illustrated typed reports, with graphs created in MATLAB or by using any other software) — 50%

• Final project — 25%

Office Hours: TuTh 1:30–3:00PM. Please feel free to drop by at any other time — no appointment necessary.

Estimated workload

- 35 hours in the classroom (including lectures and in-class portions of labs)
- 100 hours for study and homework completion
- 10 hours for term project

Total number of hours: 145

Tentative schedule:

Sept. 5, 7: Introduction. Elementary statistical concepts. *Reading*: 1.1, 1.2.

Sept. 12, 14: The average and dispersion of a data sample. *Reading*: 2.1. *Exercise* 1. MATLAB set 1.

Sept. 19, 21: Central limit theorem. Comparing means using normal distribution. Student's *t*-test. *Reading*: 2.2, 2.3. *Exercise* 2. MATLAB set 2.

Sept. 26, 28: Binomial distribution. *Reading*: through 2.4.3. *Exercises* 3–6.

Oct. 3, 5: Nonparametric statistical tests. *Exercise* 7. Poisson distribution. Exponential and Gamma distributions. *Reading*: 2.4.4, 2.5. *Exercise* 8.

Oct. 10, 12: Chi-square distribution. *Reading*: 2.6.1–2.6.2.

Oct. 17, 19: Normal sampling theory: Tests of variance. *Reading*: 2.6.3.

Oct. 24, 26: Notes on statistical inference. *Reading*: 2.7.

Oct. 31, Nov. 2: Least squares as a maximum likelihood estimator. Chi-square fitting. Fitting data to a straight line. *Reading*: 3.1, 3.2A. *Exercises* 9–11.

Nov. 7, 9: Theory of correlation. Sampling theory of correlation. *Reading*: 3.2B, 3.3. *Exercise* 12.

Nov. 14, 16, 21: Autocorrelation. Red noise and white noise. Testing for trends. *Reading*: 3.4. *Exercise* 13.

Nov. 23: No classes (Fall Recess)

Nov. 28, 30: Multiple linear regression. *Reading*: through 3.5.2.

Dec. 5, 7: Singular value decomposition. *Reading*: 3.5.3, 3.5.4. *Exercise* 14.

Dec. 12, 14: Regression-based models as a means of forecasting. *Reading*: 3.7. *Exercise* 15 (**FINAL PROJECT, DUE DEC. 19**).