AOSC-630 STATISTICAL METHODS IN METEOROLOGY AND OCEANOGRAPHY
Updated January 2012 (Eugenia Kalnay/Huug vanden Dool)

Prerequisite: STAT 400 or equivalent, or approval by instructor.

Grading: Homeworks, 30%; Mid-term take home exam, 30%; Final take home exam: 40%.

Syllabus: Review of the most widely used statistical methods currently used in meteorology and oceanography including tests of significance; time series analysis; linear multiple regression; neural networks; wavelets; Model Output Statistics; statistical weather and climate prediction.

Note: Some of these lectures are presented by world leaders in the field. The class notes prepared for this course by Dr. Huug van den Dool became the basis for the textbook "Empirical Methods in Short Term Climate Prediction" (Oxford University Press, 2007).

Required Text: Daniel Wilks: Statistic Methods in the Atmospheric Sciences (2006) Academic Press. (3rd edition). You can use the first or second editions as well).

Recommended reference textbooks: Empirical Methods in Short-Term Climate Prediction. Huug van den Dool (2007) Oxford University Press ISBN 0-19-920278-8

Hans von Storch and Francis Zwiers (1999): Statistical Analysis in Climate Research. Cambridge University Press (now in paperback).

Approximate schedule:

1. Introduction, probability distributions, tests of significance (3 weeks)
   (1) Introduction - concepts of probability, random variables and probability distributions. Wilks: Chapters 2, 3, 5.1
   (2) Probability distributions, discrete, continuous - the normal distribution, Central Limit theorem, χ^2 -distribution, t-distribution, and Fisher's F-distribution. Gumbel, Gamma and other distributions. Wilks: 4.1-4.5; 4.7
   (3) Tests of hypothesis - Type I error, Type II error, level of significance, one tailed tests and two tailed tests. Parametric tests of significance against non-parametric tests and Monte Carlo methods. Bootstrapping. Wilks: 5.1-5.4

II. Statistical Weather forecasting (3.0 weeks) Wilks: Chapter 6; other refs.
   (1) Simple regression - estimation of regression line, analysis of variance, confidence interval for regression coefficients, and confidence band for regression line.
   (2) Multiple regression - estimation of regression plane, partial correlation, and multiple correlation.
   (3) Screening regression - explained variance and incrementally explained variance, all possible regression, forward selection, stepwise regression, and stagewise regression.
   (4) Model Output Statistics, Perfect Prog, Adaptive Regression (Kalman Filtering). **Guest Lecturer: Mark Antolik**.
   (5) Nonlinear regression, neural networks (**Guest lecturer: Vladimir Krasnopolsky**).
   (6) Probabilistic forecasting and verification from ensembles (Guest lecturer: **Malaquías Peña-Méndez**).

III. Time series (3.0 weeks) Wilks: Chapter 8, plus additional refs.

(1) Introduction - definitions of stochastic processes: purely random process, stationary process, auto-regressive process and non-stationary process.
(2) Analysis of discrete time series - harmonic analysis, smoothing and filtering, frequency response of smoothing and filtering functions, and construction of low-pass, band pass and high-pass filters.
(3) Power spectrum analysis - Methods of estimating power spectra: Lag-correlation, Fast Fourier Transform, Maximum Entropy. Aliasing.
(4) Cross-spectrum analysis - Estimation of co-spectrum, quadrature-spectrum and coherence.
(5) Wavelets (Guest lecturer: **Andy Tangborn**)

IV. Statistical methods for climate prediction (5.0 weeks, lecturer: **Huug van den Dool**)
(1) Introduction: Empirical orthogonal functions (principal components) - rotated and complex empirical orthogonal functions. Coupled fields: Singular Value Decomposition, Canonical Correlation Analysis. Clustering (Wilks, Chapter 9)
(2) Applications developed at CPC: Empirical Wave Propagation; Natural analogues; Constructed analogues; Empirical Basis Functions; Teleconnections; Empirical Orthogonal Teleconnections: examples from reanalysis; Empirical Orthogonal Functions; Compact representation of data sets

V. Forecast verification (0.5 week)
(1) Currently used operational forecast scores