ATM 693: Analysis Methods in Meteorology and Climate

Spring 2013

Professor: Igor Polyakov
Office: IARC 408g
Tel: (474) 2686
email: igor@iarc.uaf.edu

Class room: ARSC computer lab WRRB 009
Class times: MW: 9:45am to 11:15am
Office hours: TBD
3 credits

Prerequisites: Understanding of statistics and basic programming skills using fortran or some other form are required. Examples will be provided using fortran and IDL (the syntax of IDL is similar to fortran).

Introduction: It is your first day of work and your supervisor comes to you and says, "It's great to finally have on board someone who knows something about environmental data. What I'd like from you is a map showing maximum hourly 10-year return temperatures for the 1950-2005 period for all weather stations reporting in North America."

At the very least we can note the following about this task:

- There are almost half a million hourly observations for this period
- There are thousands of weather stations
- What data set(s) are there to handle this? – Where to get them?
- How to read them in?
- How is a return frequency analysis performed?
- How do we ramp it up to run for so many weather stations?
- How do we present the results?

But where/how do you start? This is what we will explore.

This course is an introduction to the access and analysis of the large data sets that are often encountered in atmospheric sciences, and many other allied disciplines, such as oceanography. We will cover these broad areas:

- where to find data
- how to get it into a form that can be analyzed (programming)
- various standard analysis methods, including extreme event work.

During this class we will find and use observational data in order to explore some real scientific problem. For example, for 2013 Spring semester we may explore "*Regional and temporal features of Arctic warming*." Each student will analyze data from a limited Arctic region. In the beginning of each class some theory for statistical methods will be provided which will be first applied to a simple synthetic time series or fields in order to understand pros and cons of the selected methods. At the end of the class and at home each student will apply these methods to real data for his/her region. At the end of the semester we will compose jointly a pan-Arctic picture of observed high-latitude atmospheric changes.

We do have a recommended text. Readings from the text will be regularly assigned. The course will follow these readings because the textbook was designed for an atmospheric analysis type of course, and you should keep up with them. Of course in class we will emphasize certain topics. Material will be conveyed by standard lecture, in-class discussions, in-class presentations by students, and lab work.

Course Objectives:

1. Learn about weather and climate data storage formats. This includes various regular and non-regular ascii formats for observational data and a range of self-describing gridded binary formats, including HDF, GRIB, netCDF. Understand how their dimensionality works. Develop creative ways to read in all kinds of data. Data I/O is the biggest stumbling block to analyzing data!
2. Learn to manipulate programming software to tackle large data sets. These software packages are fundamental tasks of analysis work; familiarity with their operation is essential.
3. Learn about statistical methods to reduce and analyze data sets. Analysis methods frequently utilized in weather and climate research will be explored. This includes standard aggregate reduction statistics, time series analysis, and eigen methods.
4. Learn about extreme event analysis. A frequently requested output is some idea of return intervals of extreme events. We will explore this topic.
5. Learn about model output analysis. Weather and climate models are heavily utilized. Assessment of their output is a critical first step before their results can be folded into research.
6. Results presentation. An important component to this business is presentation of your results, both in oral format as well as in journal format. We will work on issues of presentation via submitted work and class presentations.

Student learning objectives

By the end of this course students should be able to:

- Describe what a self-describing gridded binary format is and identify major formats currently in use.
- Know how to gain information about such a dataset when presented with one.
- Delve into a dataset to reduce and prepare it for further analysis.
- Perform a variety of standard types of data analysis (e.g. correlations/regressions/EOFs).
- Present project results in a clear and concise manner. This includes readying results for print publication.

Textbook/readings:

The students do not need to buy books. The course will draw primarily from the following books but will also draw from journal articles and other texts:

- Statistical Methods in the Atmospheric Sciences 3d ed., by Daniel Wilks (2005), Academic Press, ISBN 978-0127519661.
- Statistical Analysis of Climate Research by Hans von Storch and Francis W. Zwiers; Cambridge. University Press, Cambridge, 1999, x+484pp., ISBN 0-521-45071-3.
- Statistics of Extremes. Theory and Applications by J. Beirlant etc. Wiley Series in Probability and Statistics, 2004, ISBN 10: 0-471-97647-4.

Evaluation: The course grade will consist of the following components. Final letter grades will be based on a standard scale: A=90 to 100%, B=80% to 89%, C=70% to 79%, D=50% to 69%, and F≤50%. As of Fall 2006, UAF has instituted a +/- scale to the grades, so the bottom and top 3 percentage points will fall within the '-' and '+' ranges, respectively. For example: 90-92% will be an A-, 93-96% will be an A, and above 97% will be an A+. Note that tests will be graded on a curve, so the above scale may be modified.

This course is based on a combination of lectures, labs and homework. There will be approximately one homework assignment each 1-2 weeks. It is expected that your homework is returned in 1-2 weeks. Late problem sets will have grade lowered by 10% per day late. Exam will be in-class open-book exam; practical exercises will be its integral part. Missed exams will be given a 0 grade and make up exams will be given only under extenuating circumstances.

There will be a system of bonuses for lab/homework when creative additional in depth analysis is carried out (2% for each bonus, 10% maximum). For example, similarity between two time series may be established by several methods; use one-two methods in addition to those used in the lab for your homework. Or make investigation of sensitivity of your statistical method to errors using Monte Carlo approach (if it is not required by instructor). Or take active part in the concluding class devoted to synthesis of regional analyses to a pan-Arctic picture of warming.

	ATM693
Attendance/participation	40%
Homework	40%
Quiz	20%
Bonus points	10%

Disabilities Services:
The Office of Disability Services implements the Americans with Disabilities Act (ADA), and insures that UAF students have equal access to the campus and course materials. We will work with the Office of Disabilities Services (203 WHIT, 474-7043) to provide reasonable accommodation to students with disabilities.

CLASS CALENDAR

Wk	Day	Date	Topic
1	W	01/23/13	**0. Intro**: Goals/objectives, using lab's hardware/software
2	M	01/28/13	**1. Data structure**: ASCII, binary and self-descriptive formats
	W	01/30/13	Computer lab
3	M	02/04/13	**2. Probability and statistics:** Descriptive statistics, distributions
	W	02/06/13	Computer lab
4	M	02/11/13	**3. Time series**: Regression, correlation
	W	02/13/13	Computer lab
5	M	02/18/13	**4. Time series**: Autocorrelation
	W	02/20/13	Computer lab
6	M	02/25/13	**5. Time series**: Spectral analysis
	W	02/27/13	Computer lab
7	M	03/04/13	**6. Filtering**: Low-frequency signal and seasonal and weather "noise"
	W	03/06/13	Computer lab
	M	03/11/13	Spring break
	W	03/13/13	Spring break
8	M	03/18/13	**7. Wavelet analysis**
	W	03/20/13	Computer lab
9	M	03/25/13	**8. Statistical estimation derived from autocorrelated time series.**
	W	03/27/13	Computer lab
10	M	04/01/13	**9. Extreme value analysis**
	W	04/03/13	Computer lab
11	M	04/08/13	**10. Time series**: Parametric vs. non-parametric analyses. Bootstreping
	W	04/10/13	Computer lab
12	M	04/15/13	11. **Eigenmode analysis**: Empirical Orthogonal Functions (EOF)
	W	04/17/13	Computer lab
13	M	04/22/13	**12. Eigenmode analysis: (cont)**
	W	04/24/13	Computer lab
14	M	04/29/13	**13. Statistical forecasting**: Multivariate regression analysis
	W	05/01/13	Computer lab
15	M	05/06/13	**14. Overflow:** Building a large-scale picture from regional analyses