Using Python to Store, Display, and Analyze Pressure Observations from Earthscope's USArray Network

Alexander A. Jacques and John D. Horel
Department of Atmospheric Sciences, University of Utah

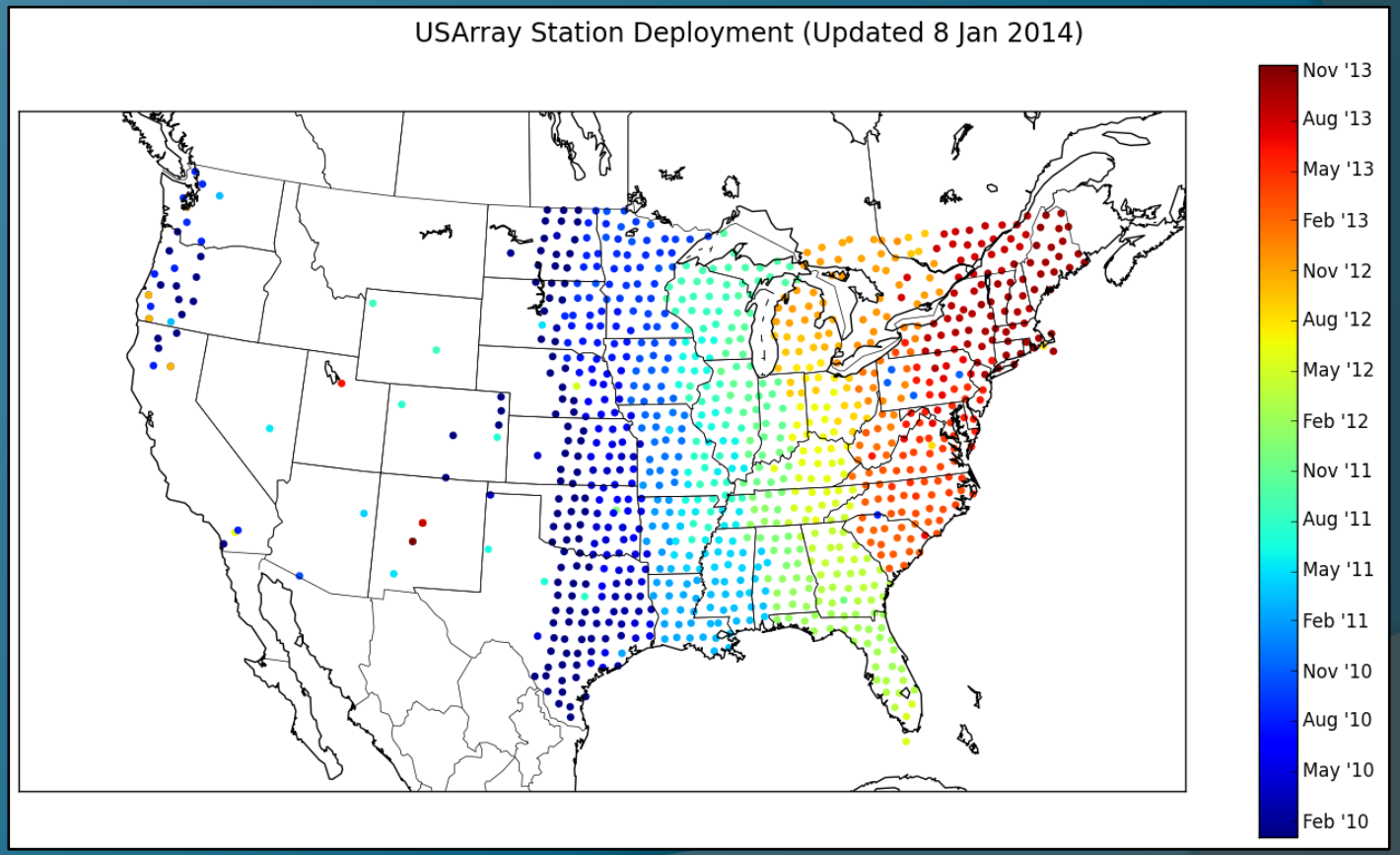
Fourth Symposium on Advances in Modeling and Analysis Using Python
2014 American Meteorological Society Annual Meeting
Atlanta, GA
February 3, 2014

THE UNIVERSITY OF UTAH

NSF

What is the US Transportable Array (USArray)?

- EarthScope-funded project of ~400 seismic stations traversing across the CONUS in an effort to catalogue and study seismic activity
- Stations outfitted with microbarometers in 2010 (1 Hz reporting frequency)



THE UNIVERSITY OF UTAH

NSF

Project Objectives

- Improve access for the atmospheric science community to both real-time and archived pressure observations through displays and other web-based products at varying sample intervals (1-second, 5-minute, etc.)

- Develop and analyze climatologies at each observing site of pressure perturbations and signatures in terms of frequency and amplitude

- Examine pressure perturbations and perturbation gradients in the context of local, mesoscale, and synoptic-scale flows (waves, terrain-flow interactions, thermally-forced systems, etc.)

THE UNIVERSITY OF UTAH

NSF

Motivation for Use of Python

- USArray pressure dataset provides a unique opportunity to catalogue meteorological events at multiple spatial and temporal scales
    - Semi-regular station spacing (~70 km between locations)
    - 1 Hz temporal resolution
    - Station deployment ~1.5-2 years

- Python chosen due to its numerous advantages in the following areas…
    - Storage and querying of large arrays of data
    - Object-oriented nature for fast manipulating
    - Time series filtering procedures
    - Leveraging of interactive web displays

THE UNIVERSITY OF UTAH

NSF

Python for USArray Data Storage

- Dataset produces a number of challenges due to its large size, spatial irregularities (non-gridded), and temporal irregularities

- PyTables chosen as mechanism to store observations in two "databases" based on station and time in compressed HDF5 format

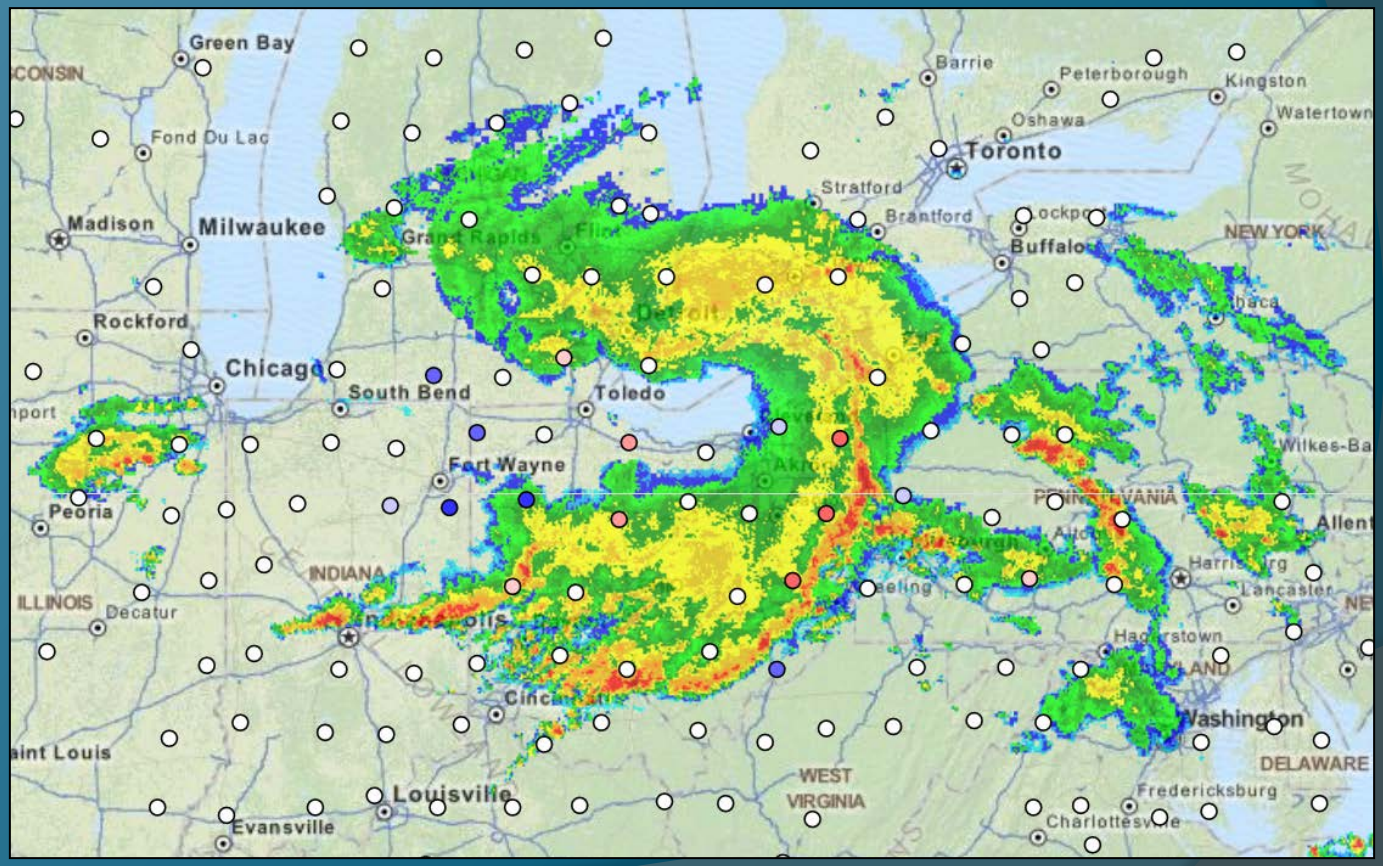
- Database #1: Individual station tables of all observations
    - Level-7 Zlib compression
    - New data appended on a daily basis
    - Statistics from 1 Jan 2010 – 31 Dec 2013
        - Total observations: 46,086,051,637
        - Disk Space Used: 58 GB
        - ~1.26 bytes per observation (time, station, value)

THE UNIVERSITY OF UTAH

NSF

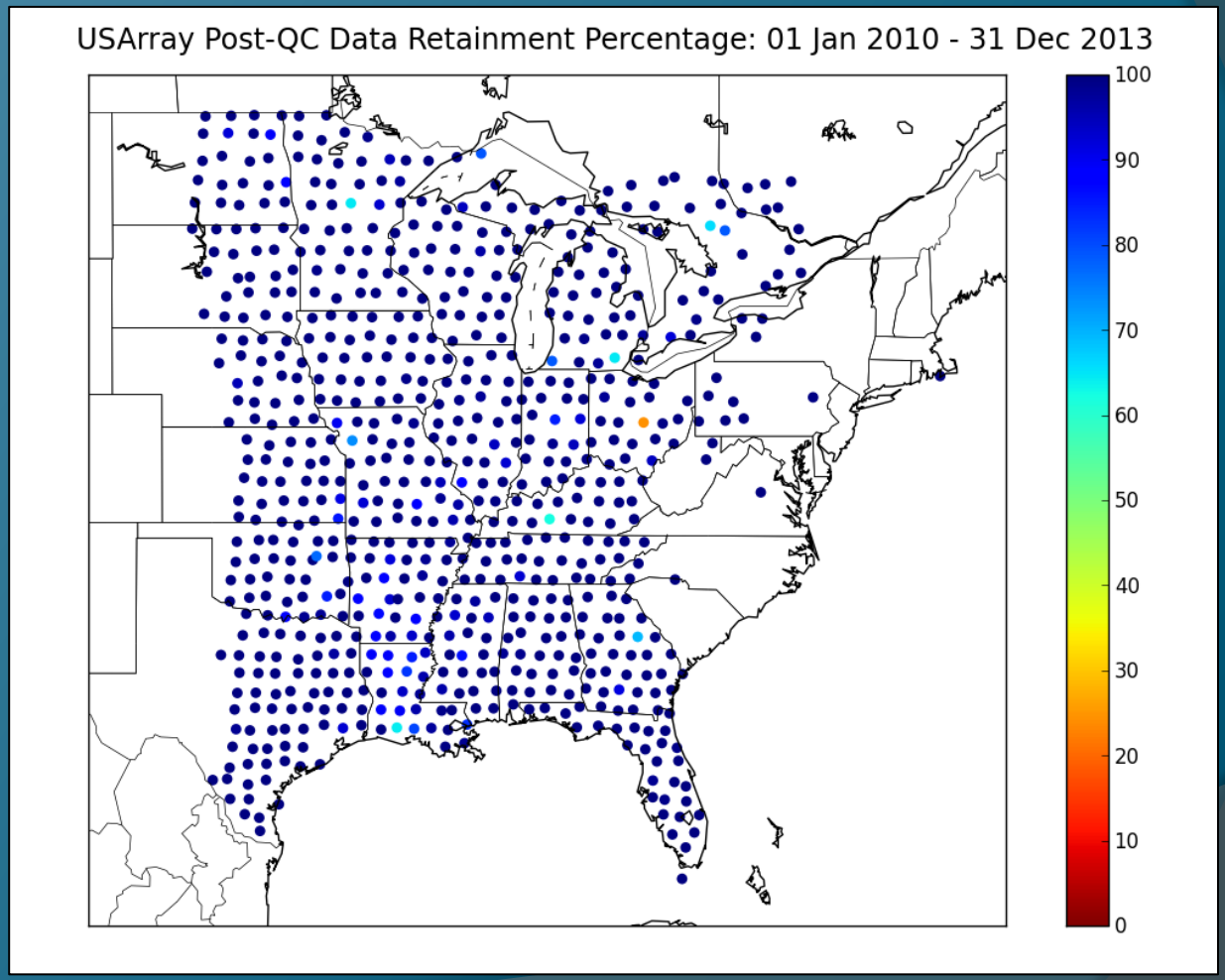
Python for USArray Data Storage

- Database #2: Daily files of all active USArray Stations
  - Nested tables used to store data in pseudo-gridded format
  - Supplemental metadata schema used to leverage dataset onto maps



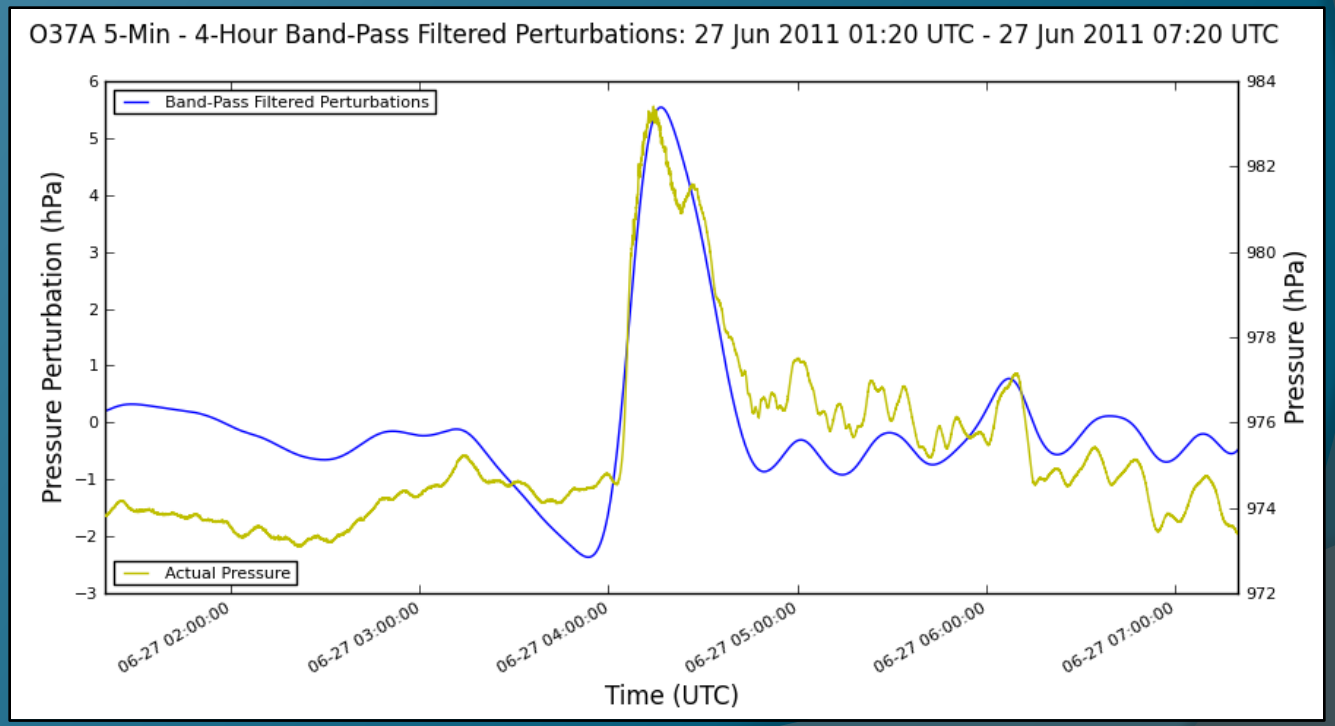
Python for USArray Data Analysis

- Quality Control: the ability of NumPy to efficiently parse large data arrays aided significantly in identifying missing/spurious data periods



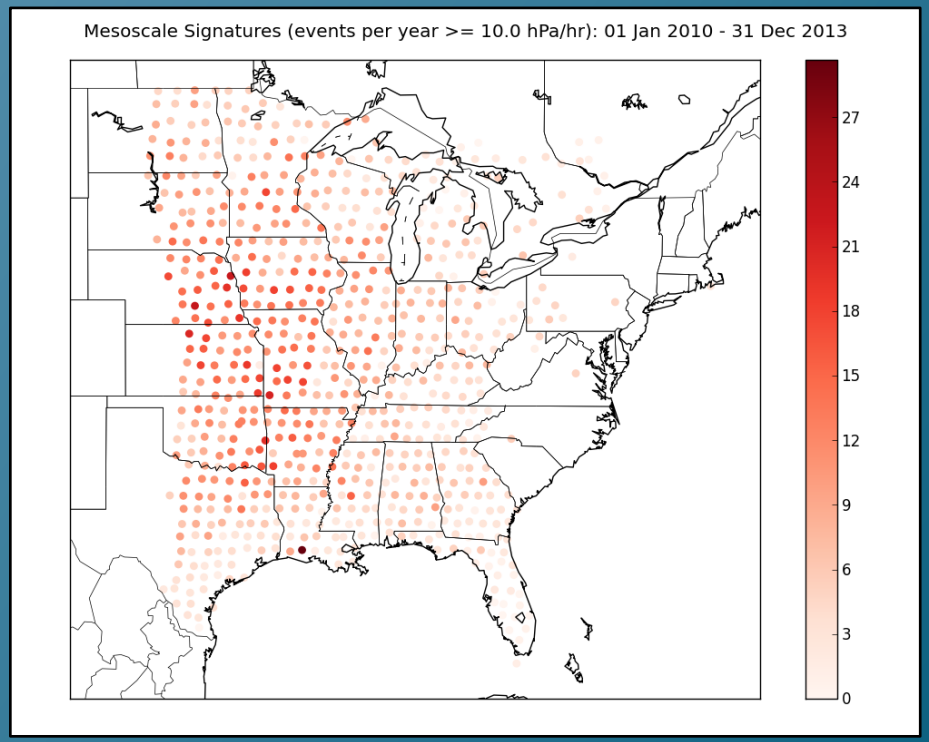
Python for USArray Data Analysis

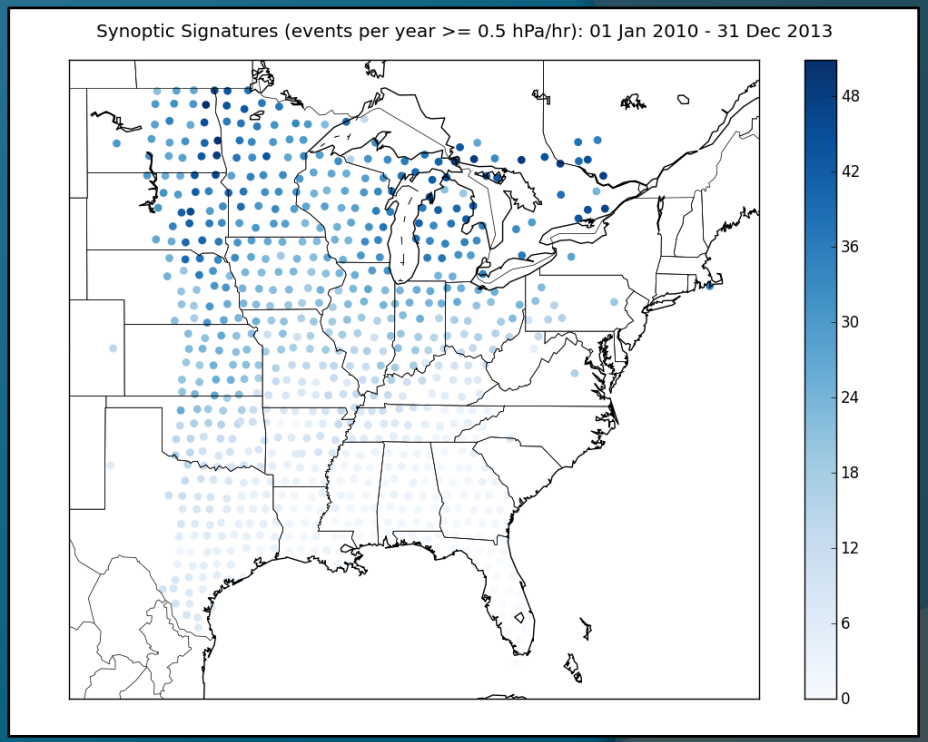
- Mesoscale/Synoptic Pressure Signature Identification Algorithms
  - SciPy Butterworth filtering techniques applied to generate perturbation time series from the original time series data
  - Large pressure signatures identified using local min/max functions



Python for USArray Data Analysis

- Descriptive quantities of the pressure signature fields stored in HDF5 PyTables file for fast querying based on user-defined preferences



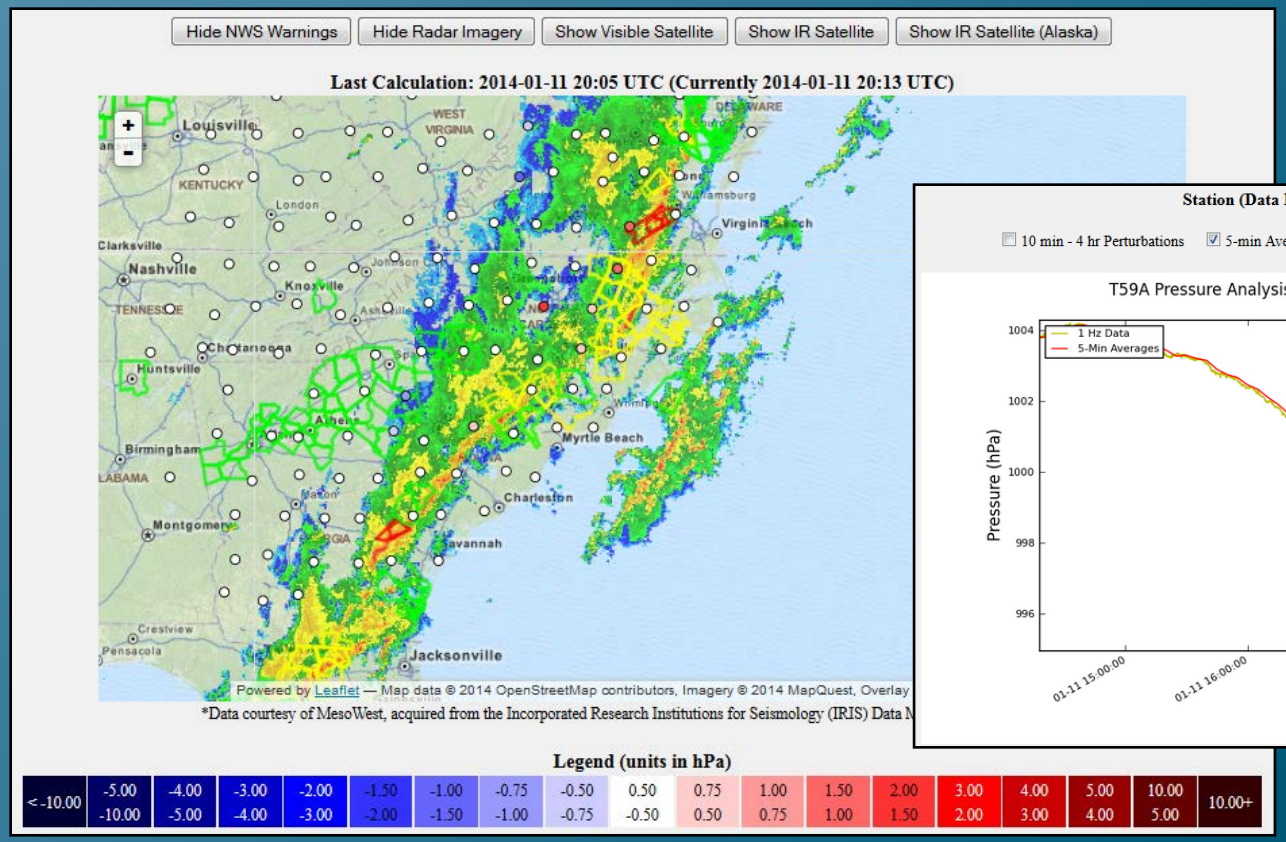


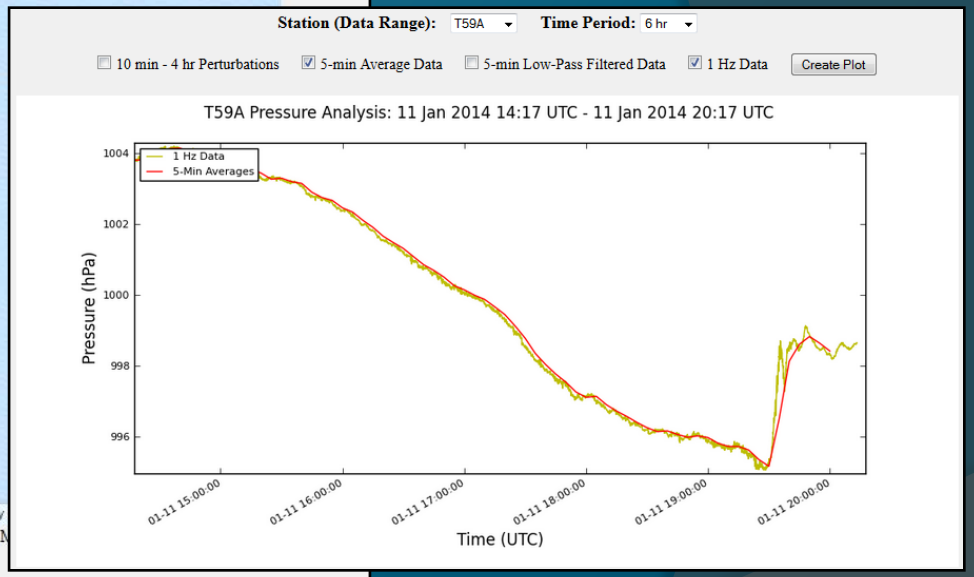
THE UNIVERSITY OF UTAH

NSF

Python for USArray Web Displays

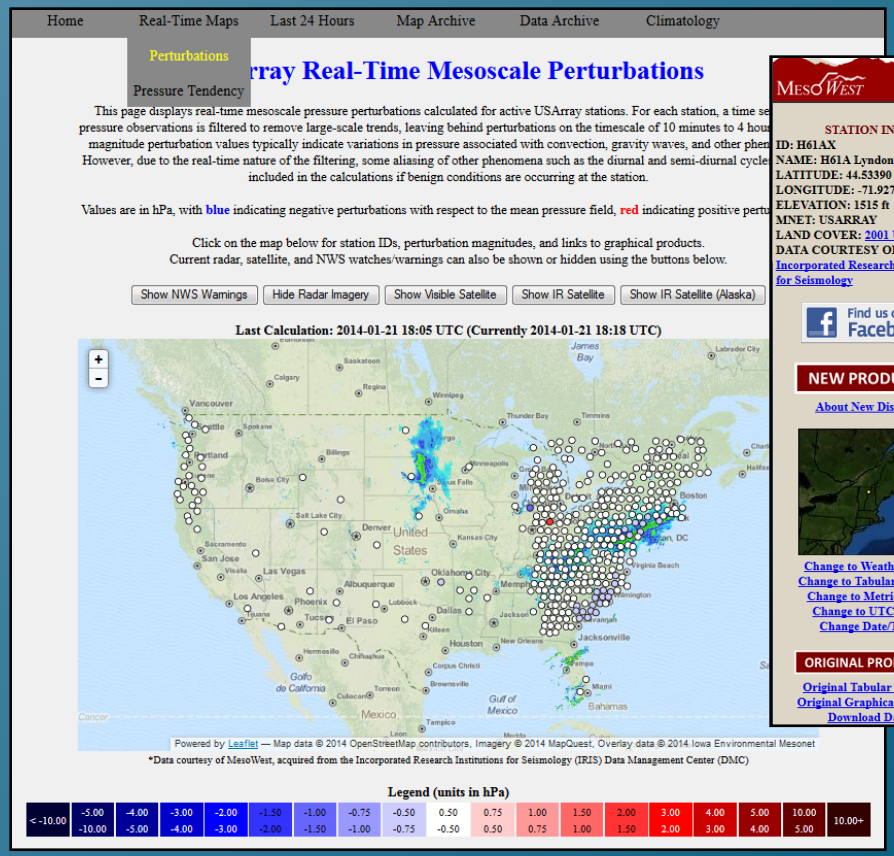
- Website (http://meso1.chpc.utah.edu/usarray) uses Python in conjunction with JQuery and Leaflet to generate interactive maps and plots

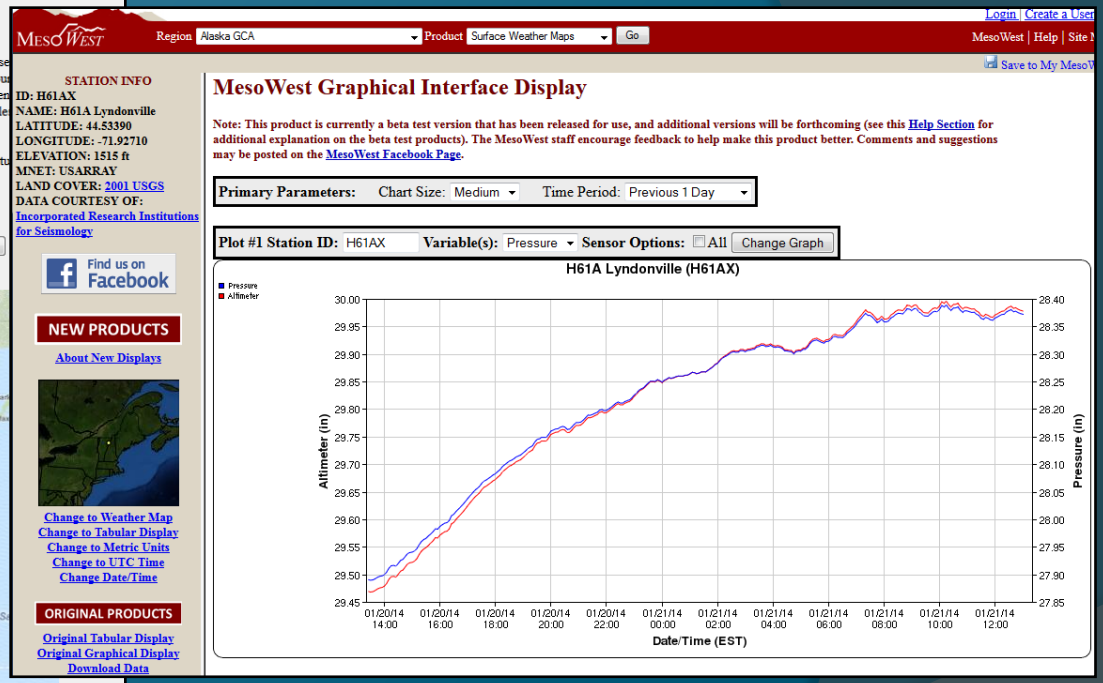




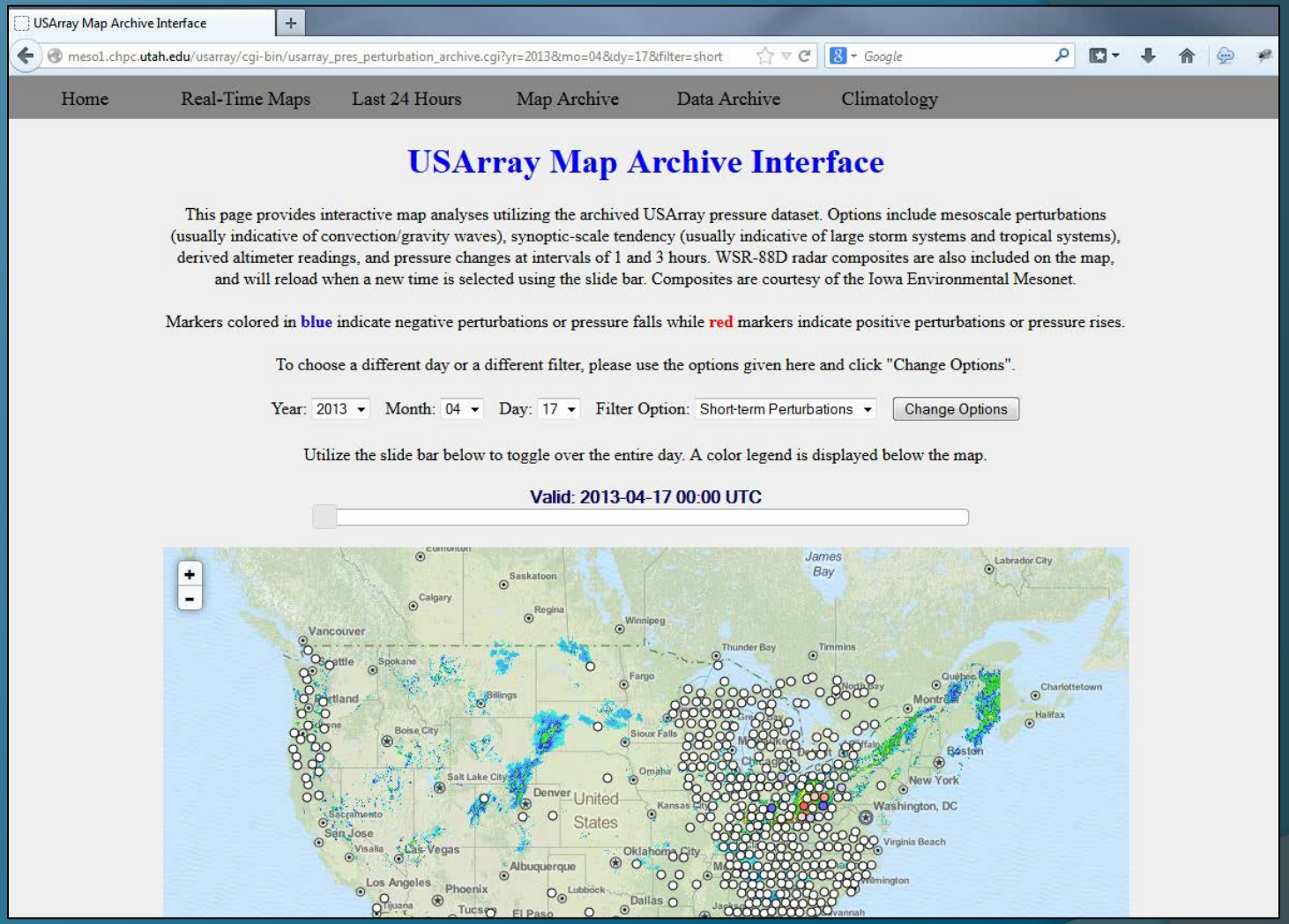
Python for USArray Web Displays

- Real-time displays utilize 5-minute observations courtesy of MesoWest (http://mesowest.utah.edu). MesoWest data is also forwarded onto MADIS.



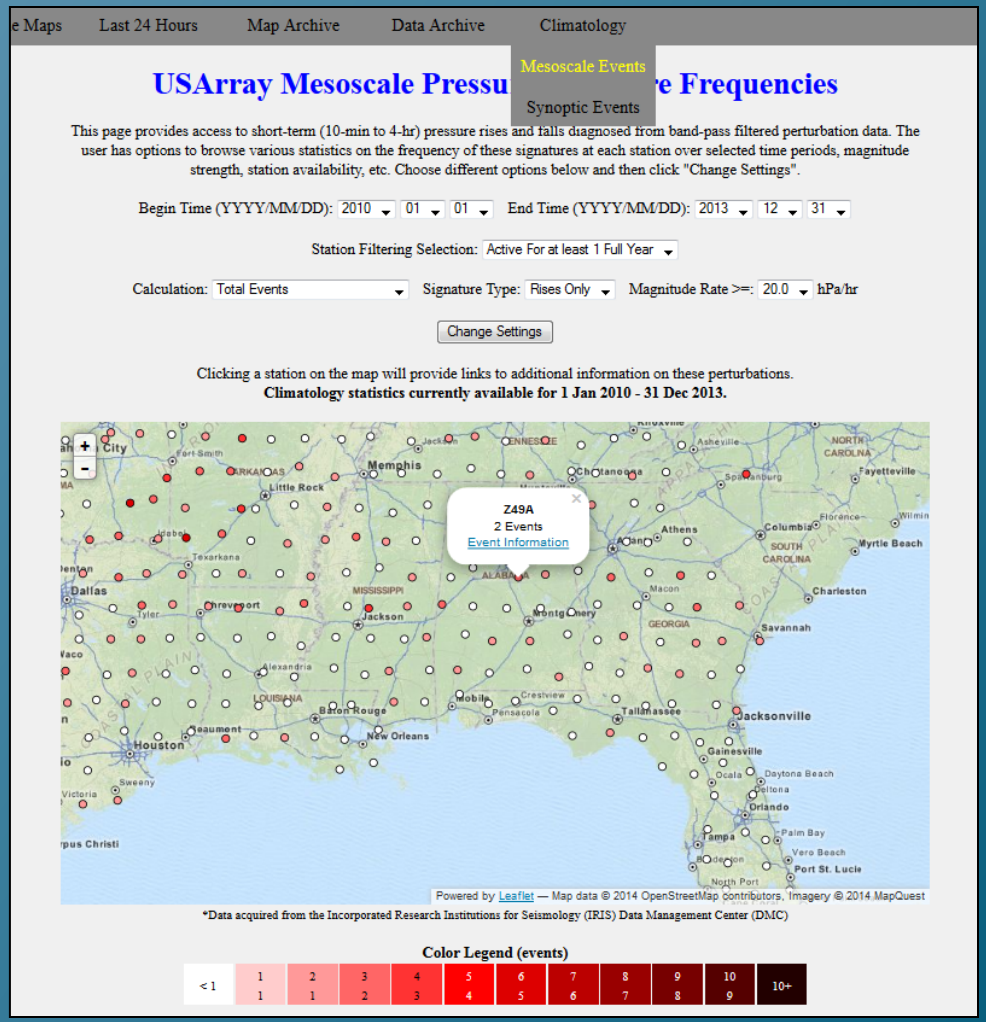


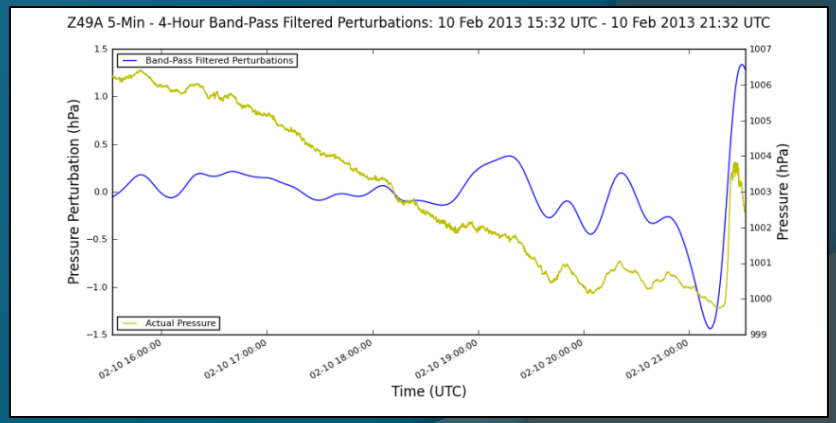
Python for USArray Web Displays



http://meso1.chpc.utah.edu/usarray

Python for USArray Web Displays





http://meso1.chpc.utah.edu/usarray

Future Tasks

- Continue developing climatologies of pressure perturbations and signatures for all USArray stations using Python storage/analysis techniques
- Explore wavelet analyses of time series using PyWavelets
- Develop algorithms to calculate pressure perturbation gradients
- Develop climatologies of the above pressure perturbation gradients for each USArray station in the database

THE UNIVERSITY OF UTAH

NSF

Acknowledgements and References

This research is funded by National Science Foundation Grant Number 1252315. We would also like to thank Dr. Frank Vernon of Scripps Institution of Oceanography, the USArray Array Network Facility (ANF), and the Incorporated Research Institutions for Seismology (IRIS) for providing access to live data streams for the USArray project.

Additional work on this project can be found in tomorrow's poster session…

Jacques, Alexander A. and J. Horel, 2014: Pressure Signatures of Extreme Weather Events Deduced from Earthscope's USArray Network
26th Conf. Weather Analysis and Forecasting - Poster #108

THE UNIVERSITY OF UTAH

NSF