August 19, 2015
Daniel Parrott

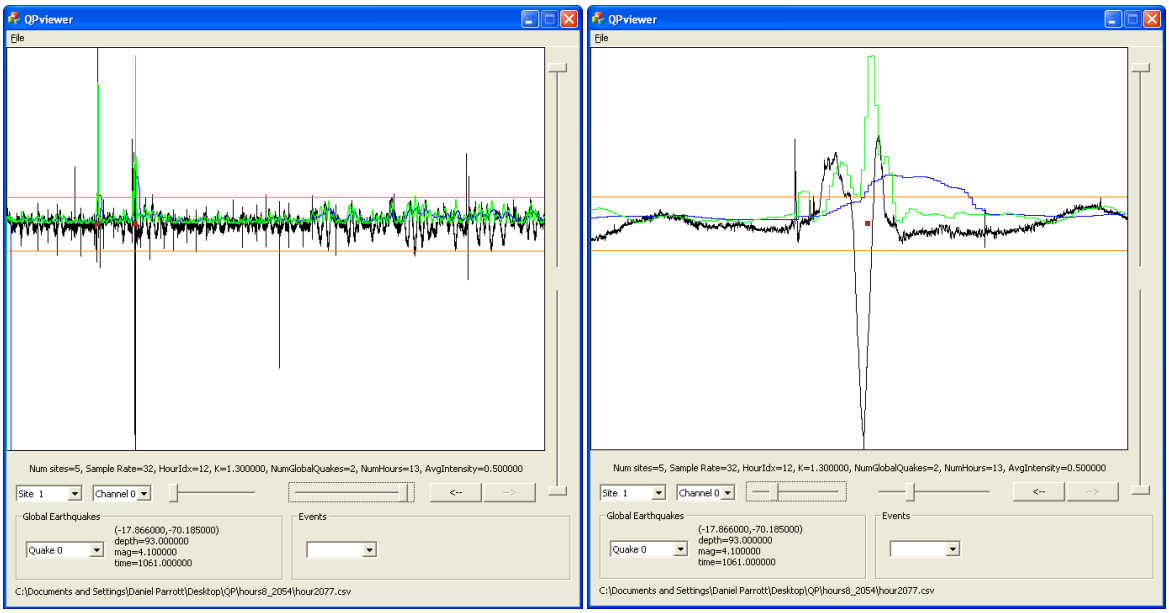
Description Document for QuakePredictor Submission

Background

The purpose of the project is to write a program that returns an array containing probabilities of a seismic event following the detection of pulses from a magnetometer signal. To this end, the first step I took in writing my program was to have some way of looking at the data to determine what these pulses look like and how I could detect them.

Data Viewer

From the project description, each test case is comprised of magnetometer data for up to nine sites, and the data has sample rates of either 32 or 50 Hz. Given the format of the data, I was able to write a program to aid in visualizing it in order to better understand how to build the pulse detector.



The above screenshots show the data viewer, with a full hour of data on the left and a zoomed-in shot of one of the pulses on the right.

Pulse Detector

After looking through hundreds of data files using the data viewer, I was able to get a pretty good idea of what constitutes a pulse of interest. To this end, I proceeded to write a detector that could analyze a given magnetometer signal and identify pulses that have the desired duration and amplitude. This was accomplished by writing a function that computes a running average of a given duration, and then using that function to compute two averages, known as short term average (STA) and long term average (LTA).

For my detector, I used an STA of 3.0 seconds and an LTA of 30.0 seconds. A pulse is identified whenever two main criteria are met: the ratio of STA to LTA exceeds a given threshold, and the difference between STA and baseline amplitude also exceeds a given threshold. When such a pulse is found, its time index (in seconds) is added to a vector of time indices.

August 19, 2015
Daniel Parrott

Pulse Counter

The program runs the above detector on each of the input sites and tallies up the number of pulses that it encountered for a given hour of data. Because some of the magnetometer data can include calibration pulses, the counter looks at how many pulses were detected in the first 800 seconds, and if that count exceeds a given threshold, they are discarded as being calibration pulses.

Quake Predictor

Magnetometer data leading up to previously encountered seismic events was used to establish a model for seismic events based on the input signal, as well as to determine the thresholds for the pulse detector. Then, for each hour of data, the program returns an updated forecast using the pulse counts derived from the past three weeks of signal data.

Answer for Follow-on Question

Because the solution relies on the magnetometer data to update the forecast, the highest score I could have reached without using that data would likely have been zero.