Measurement, Uncertainty, and Data Analysis
Activity #3: Detecting a Signal Through the Noise
Student Worksheet

Objective

Determine how scientists are able to detect very weak signals in the midst of noise.

Background

The signal from mesospheric ozone is very weak compared to the noise accompanying it. Still, we are able to make a measurement of the signal by detecting the ozone from many locations over long periods of time. This allows us to integrate a signal that can be detected above the noise by adding together the signal from multiple detectors or over extended periods of time. The key to this is the fact that the noise, unlike the signal, is random, and therefore, tends to average away when many measurements are combined.

The "signal" in this case is the spectral line, which we expect to have a peak around channel 32 of the spectrometer and a Gaussian-like shape, with a width of a few channels.

Method: Data Access and Interpretation

Go to www.haystack.mit.edu/ozone and create a spectrum for any single spectrum and any 5 day period you like. Just make sure your spectrometer was operational on the days that you select! (It should be obvious if it wasn't, since you won't get any data!). Use the filter for "nighttime data only."

Once you have plotted your spectrum, download the avout.txt file containing the data in your plot. To do this, right mouse click on the link and select "save link as" to save the file in a convenient location on your computer.

Then, open the avout.txt file in Excel. You may need to ask Excel to look for "all files" when opening it, rather than just "excel files." The avout.txt file is a fixed width text file, which Excel should know what to do with.

For more information on the file you have opened, you may want to read the info.text file available on the same page as the graph you produced.

The first thing you are likely to notice when you download the data into Excel is how much of it there is. You are likely to have about 450 rows of data and 80+ columns. The heart of the data is the last 64 columns. Each row of data corresponds to a different 10 minute block of time, and each of the columns corresponds to a different frequency, with channel 32 corresponding to the central frequency for the ozone signal, at 11.0725 GHz.

Method: Analysis and Graphing

The first thing you need to do with the data is delete the measurements that did not correspond to any (or very many) measurements of the ozone. To do this, inspect the column corresponding to num_rec. This contains a number, usually 8 or 9, that corresponds to the number of measurements the system collected in each 10 minute segment being reported in the data. If this number is not 8 or 9, it means that some of the data was automatically discarded for being too noisy or coming from some malfunction. The bad

data is already gone, but in order to average together equivalent chunks of data, you will need to delete the rows corresponding to a num_rec of 5 or less. You can do this by scrolling through the rows, deleting the entire row whenever you see a num_rec value of 0, 1, 2, 3, 4, or 5. Or, you can use the search feature in Excel and find them that way. In either case, make sure you delete the entire row of data, not just the num_rec value.

Once you have deleted the signals corresponding to smaller numbers of measurements, you are ready to start averaging and graphing. You should use Excel to help you average, since you will need to average each of 64 channels over different numbers of rows. Average the spectra over 1 hour (about 6 rows), 5 hours (about 30rows), 1 day (about 70 rows, but will depend on time of year), and all 5 days of the data you've downloaded.

When you have created your averaged spectra, create column graphs of your results. You should create 5 different column graphs: 1 row of data, 1 hour of data, 5 hours of data, 1 day of data, and 5 days of data.

Data Analysis

Once you have created your bar graphs, copy them into a single document so that you can compare them on the same piece of paper. As you compare the graphs, identify which graphs look the most and least like the expected shape of the spectral line.

Discussion

Summarize your findings in one paragraph, including an explanation of why the graphs had the appearance they did. Be sure to refer to the objective and include a clear statement of your results.

Measurement, Uncertainty, and Data Analysis
Activity #3: Detecting a Signal Through the Noise
Teacher Notes and Sample Results

This activity is intended to show students how data that appears noisy over a short amount of time does can be averaged together to discern the signal. Students should see that for the MOSAIC, an integration time of at least 1 day is necessary to have any confidence in the peak value of the spectrum or even the central frequency. More time improves the signal.