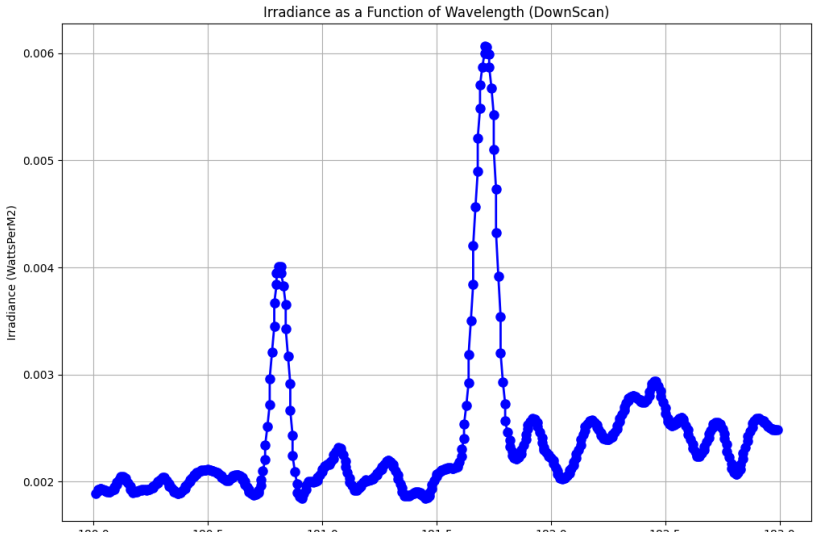
Interview Coding Challenge - Paul Swanson

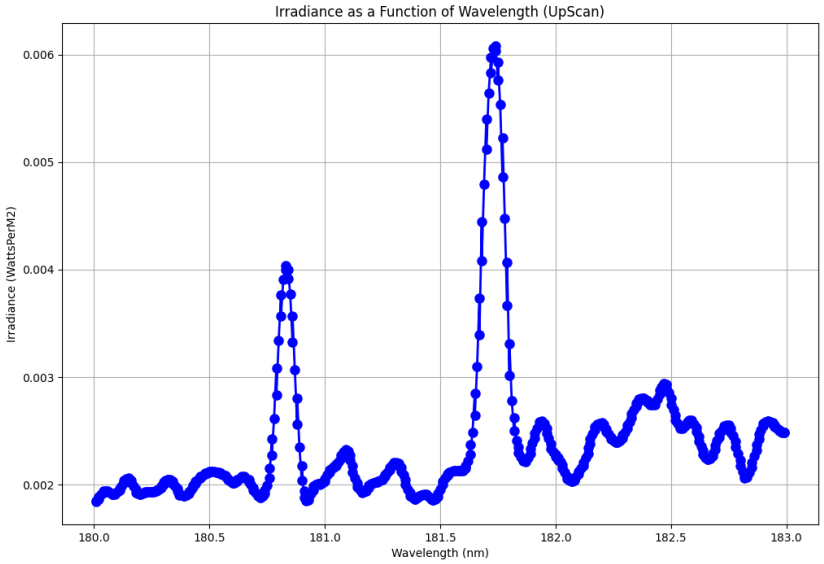
Processing data from SORCE mission from 2009-11-28

October 12, 2023

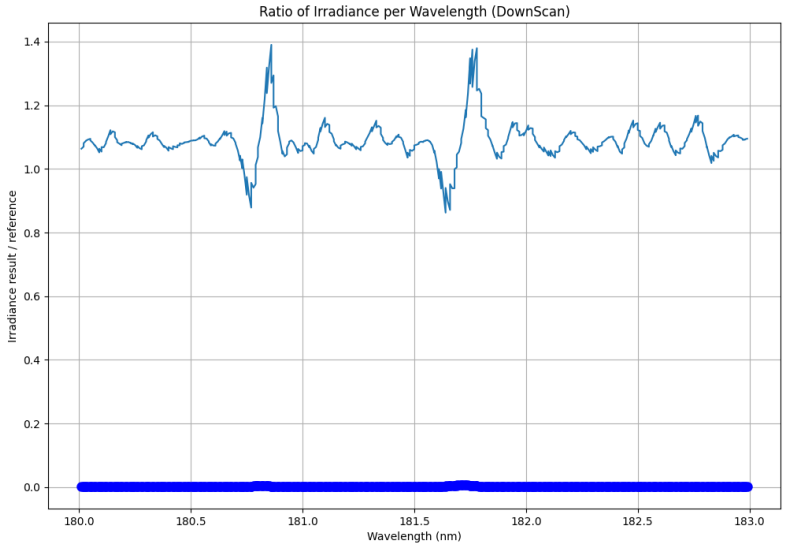
Results

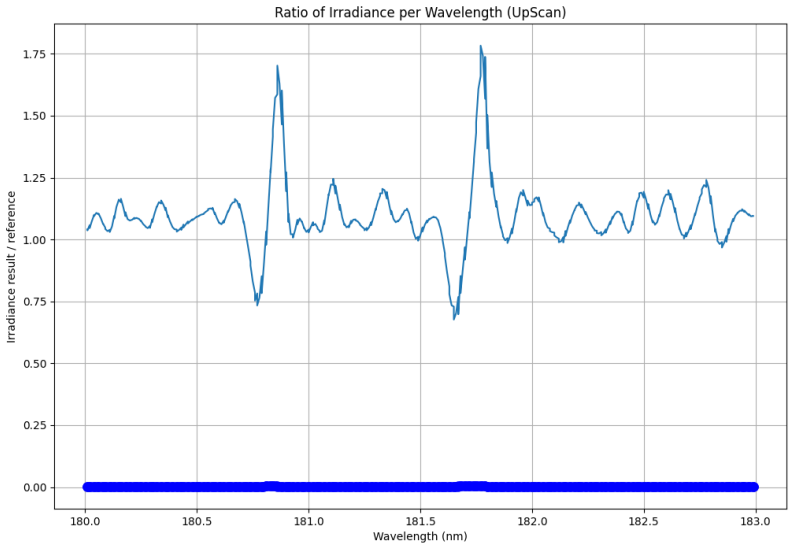
From the six data files a python application was created to process the data. This section includes the data plots, a description of the application development process, whether the results can be considered reliable. Plots were made for the Irradiance as a function of Wavelength and plotted separately for the down scan and the up scan.





The ratios at each wavelength were also plotted. It shows the result irradiance over the accurate reference spectrum data at each wavelength.





After reading the background and a description of the data files, it was decided to start with the plans.txt document. This file provided the start and end times for the

Down, Up and Dark scans. This would make the code more efficient by ignoring telemetry data outside of the scan periods. The measurement data were further observed for trends that could be used for processing. At the Downscan start time, the grating position in instrumentTelemetry goes from 0 to non-zero. At the Downscan end time, the grating position goes back to zero. The same was true for Upscan where the end was at the end of the telemetry file.

The integrationTime interval changes from 1000 to 1750 at the beginning of the Downscan. Reading the integrationTime file is slightly easier to read since it is per second, so there is only one number difference for each row, meaning you can see the 0-9 pattern, which is not true for the other files which later were discovered to be at an interval of 1750 per line during the scans so the interval wasn't obvious.

There was an exact match on the timestamps when the scans stopped and started which was not true for each line between the files.

With the counts not being per second, the counts were multiplied by the interval ratio to get them per second. The interval during the scans was always 1750 even though there were other values for interval (e.g. 250.0). A full version of the application would use the actual interval value, but it was hard coded in this case.

The Astronomical Unit was also added to the irradiance calculation. The initial value was used throughout, but a full application would check for a change based on a match in the distanceAndDoppler.txt file for each telemetry row iteration. The AUs change less frequently (only ~700 rows). The sunObserverCorrectionDistance from row one was used throughout. It's generally ~1.0273.

The provided calculations and constants were used to calculate the fields required for plotting.

Several considerations were made along the way to prove the data were being processed as expected.

The two plots irradiance plots appear nearly identical on a large scale when including all of the data, even thought the results files had different data. Sample runs of only the first 30 rows reveal obvious differences in the plots showing it was being done correctly.

The differences between the start and end indices for the downscan was found to be 2528, which was the same number of rows in the results file for the scan.

The difference between the timestamps of two rows during the downscan was found to be exactly 1.75 seconds which is to be expected due to the interval being 1750 during the scans.

Checks like these can be added to unit tests.

Containerization

The application is written in python 3.8.5 and can be found here

https://github.com/pwalswan/solar

From a terminal

git clone https://github.com/pwalswan/solar.git

docker build -t solar-image .

At a terminal command line type, to get your current directory

pwd

Run the following docker command and substitute you pwd for <absolute>

docker run -v <absolute>/solar:/app solar-image

Mine for example

docker run -v /Users/paulswanson/work/solar3/solar:/app solar-image