Correlating Hubble Telescope Dark-Rates with Solar Activity

Justin Ely

605.462 Data Visualization

27 March, 2017

1 Description

The Hubble Space Telescope (HST) collects large volumes of data every day. In addition to scientific observations of stars and galaxies, the spacecraft also regularly gathers engineering data from the instruments on-board and a variety of internal sensors. This data is used by analysts and scientists to identify anomalies and improve performance of the telescope overall.

One of the instruments, the Cosmic Origins Spectrograph (COS), is particularly sensitive to interference from high-energy radiation in space. This interference often takes the form of a low-level of radiation detected even when the telescope's shutter is closed; commonly referred to as a dark-rate. In addition, the particular COS detectors permit measuring the exact arrival time of each detected photon. This millisecond-scale data stream can be combined with other measurement information to identify interesting behaviors and trends.

We hypothesis here that fluctuations in the observed dark-rate in the COS detectors is caused predominantly from direct solar radiation. Though dark-rate fluctuations can be caused by temperature fluctuations, distant supernovae, or eath's radiation belts, we believe that a thorough investigation of all available data will lead to the conclusion that the varying solar output is the leading factor.

2 Plan

To answer this question we will need to find possible physical causes, collect and reduce the relevant data, and create visualization tools to permit investigation. Finding physical underpinnings is a critical first step as observed trends mean little on their own. Though correlations and trends can be informative, physical understanding of a root cause provides validation. We will look through HST technical reports, scientific literature, and astronomy textbooks for possible sources.

The raw data will come from two sources. NASA, through STScI and the HST archive, will provide the entire suite of COS data taken with a closed shutter. This data focuses on the dark-rate without the added noise of a scientific observation. Since this data doesn't exist in the archive in the needed form, it will need to be first collected and collated. A single example dataset can be found at the following url: `https://archive.stsci.edu/cgi-bin/hst_preview_search?ne=on&imfmt=fits&name=LA7803FIQ`, though the entire dataset will contain thousands of such files. NOAA, through it's open data portal, will provide measurements of the solar flux and particle density measurements.

Specifically, these measurements will come from the Daily Solar Density ($*\_DSD.txt$) and Daily Particle Density ($*\_DPD.txt$) files found at `ftp://ftp.swpc.noaa.gov/pub/indices/old_indices/`. All necessary datasets will be collected, collated, and cleaned in Python before use in the visualizations

From the combined data, analytic visualizations will be put together to explore the collated dataset. From initial looks, correlations of the various detectors against attributes like time, orbital position, solar activity, and temperature will all need to be investigated. The investigation will focus on an interactive approach, as the high-dimensional data will be hard to summarize in just a few static charts. The visualizations will be done in Javascript and D3, and will focus on overlaying the data on the world in various transformations, selective 1D plots with regressions and correlations, and hovering to dig deeper on each individual point. The total suite will be hosted through Github Pages in an HTML/bootstrap webpage.

3 Timeline

Week Ending	Work
April 03	Locate, retrieve, and combine datasets
April 10	Locate relevant literature and turn in revised proposal
April 17	Perform exploratory analysis, turn in bibliographic report
April 24	Prototype visualizations, investigate visualization hosting
May 01	Begin final report writing
May 08	Finalize report and visualization, turn in project