

## Project Details

**ROSES ID:** NNH21ZDA001N-LWSTM

**Selection Year:** 2021

**Program Element:** Data, Tools, & Methods

**Project Title:**

Discovering Micro Events in AIA using Machine Learning

**PI Name:** Michael S Kirk

**PI Email:** michael.s.kirk@nasa.gov

**Affiliation:** Atmospheric & Space Technology Research Associates

**Project Member(s):**

- Thompson, Barbara J.;Co-I;NASA Goddard Space Flight Center
- Attie, Raphael;Co-I;George Mason University

**Summary:**

The catalogue of SDO AIA EUV images is now more than 6 PB and growing every day. Each of these images has 16 million pixels and the database is over 151 million images. Can we search the 2.5 quadrillion pixels to find single pixel eruptions? Does AIA regularly detect coronal 'campfires,' seen by Solar Orbiter? What other small dynamic phenomena are prevalent in the corona?

It is impossible to use conventional methods to analyze such an immense amount of data. This proposal utilizes the cosmic ray spikes database (so-called AIA spikes) and unsupervised machine learning (ML) techniques to search for single pixel eruptions. The cosmic ray spike removal image-cleaning algorithm runs automatically on the level-0 AIA images and removes isolated bright points which are usually cosmic rays but occasionally captures some small bright coronal features. Through using the spikes catalogue, we reduce our data science problem by over three orders of magnitude from a total AIA archive of 1015 pixels to 1012 pixels - still a very large data science problem.

Through reposing our science question into a data science framework, we further reduce the scope of the parameter space. Previous research has identified that different types of solar eruptions have a distinct evolution when observed in each of the 7 different AIA EUV passbands. For example, a coronal jet regularly has a different peak intensity than a solar flare in the 131 and 171 channels. Thus, passband intensity abstracts events into a 7-dimensional space where each dimension represents the intensity in an EUV channel.

A recent publication by Young et al. (2021), explores the relationships between the physical properties of AIA spikes in the 171 channel and their coronal environments in 126 cases. They found that 96% of the physical features identified in the AIA spikes have a diameter of less than 2 arcsec, can occur anywhere on the solar disk, and typically evolve over a period of less than 5 minutes, demonstrating that spikes are compact in both time and space. Combining the two spatial and one temporal coordinates with the 7-EUV channel coordinates yields a 10-dimensional space defining a spike.

This work will further develop temporal and spatial filters for linking spikes to their physical origin. Spatial-temporal intersections between spikes in different wavelengths give us insight into how these small features are linked together. Over a two-day sample set, we observe that about 6.5% of the remaining physically interesting spikes occur in all 7 AIA filters. Mapping these spikes back onto the coordinates of the solar surface, there is a concentration in the same location as larger solar features such as coronal holes and active regions. Where individual spikes are coincident in time and location between different AIA wavelengths in this sample set will further refine our physical filters. We eliminated over 99% of the spikes which are genuine cosmic rays using this technique in a two-day demonstration period in 2011.

**Methodology and Data**

This effort will build a database of every AIA spike in 10-dimensional space (EUV wavelength plus space and time). Abstracting the spikes from its physical origin will allow us to use unsupervised ML to further investigate trends and segment our dataset of over 1010 detections. Utilizing an innovative mix of Bayesian and clustering methods, we will identify clusters of spikes and label commonalities between groups of detections. Once this labeling is complete, we can return to our science question and ask if a group is in fact a 'campfire' observation or some other coronal eruption through extracting a subset of individual spikes, embedding them into a context region, and matching the identified label to the clusters in the database - a process described in Young et al. (2021). This effort delivers a high-value database of AIA spikes that have a coronal origin with labels of likely physical phenomena they represent.

## Publication References:

no references