





Using Cosmic Rays Detected by HST as Geophysical Markers. I. Detection and Characterization of Cosmic Rays

Nathan D. Miles^{1,2} , Susana E. Deustua¹ , Gonzalo Tancredi³ , Germán Schnyder³, Sergio Nesmachnow³, and Geoffrey Cromwell^{4,5}

¹ Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA; ndmiles@ucla.edu, deustua@arsmetrologia.org

² Department of Earth, Planetary, and Space Sciences, University of California Los Angeles, Los Angeles, CA, USA

³ Universidad de la República, Montevideo, Uruguay

⁴ U.S. Geological Survey, California Water Science Center, Santa Maria, CA, USA

Received 2020 May 20; revised 2021 April 19; accepted 2021 April 20; published 2021 September 15

Abstract

The Hubble Space Telescope (HST) has been operational for over 30 years, and throughout that time it has been bombarded by high-energy charged particles colloquially referred to as cosmic rays. In this paper, we present a comprehensive study of more than 1.2 billion cosmic rays observed with HST using a custom-written Python package, `HSTcosmicrays`, that is available to the astronomical community. We analyzed 75,908 dark calibration files taken as part of routine calibration programs for five different CCD imagers with operational coverage of Solar Cycle 23 and 24. We observe the expected modulation of galactic cosmic rays by solar activity. We model the observed energy-loss distributions to derive an estimate of 534 ± 117 MeV for the kinetic energy of the typical cosmic ray impacting HST. For the three imagers with the largest nonuniformity in thickness, we independently confirm the overall structure produced by fringing analyses by analyzing cosmic ray strikes across the detector field of view. We analyze STIS/CCD observations taken as HST crosses over the South Atlantic Anomaly and find a peak cosmic ray particle flux of ~ 1100 particle $\text{s}^{-1} \text{cm}^{-2}$. We find strong evidence for two spatially confined regions over North America and Australia that exhibit increased cosmic ray particle fluxes at the 5σ level.

Unified Astronomy Thesaurus concepts: [Geomagnetic fields \(646\)](#); [Cosmic rays \(329\)](#); [Hubble Space Telescope \(761\)](#); [Galactic cosmic rays \(567\)](#)

1. Introduction

The Hubble Space Telescope’s (HST) four generations of instruments have enabled scientific research since 1990, providing key data for new astrophysical discoveries. Operating at its current orbital altitude of roughly 538 km (Figure 1) above the Earth’s surface, HST is not shielded by the terrestrial atmosphere, and so every image obtained with a solid-state detector is polluted with charged particle events. These charged particles originate in the solar wind, coronal mass ejections, and elsewhere in the Milky Way galaxy as a result of energetic astrophysical processes (e.g., supernovae, accretion-driven phenomena like jets). Observers design their programs to minimize the effect of cosmic rays on astronomical images, and significant effort goes into developing software to identify, flag, and remove these cosmic rays so that the acquired data are useful for astronomical science analysis. However, this process also throws away information that could be used for geophysical investigations.

For example, the incident flux of charged particles at the Earth is affected by the strength, stability, and morphology of the Earth’s magnetosphere. Upon entering the Earth’s magnetosphere, these particles are immediately subjected to the Lorentz forces associated with their motion through the Earth’s magnetic field (hereafter referred to as the geomagnetic field). Low-energy particles become bound to the field lines and follow helical trajectories toward the poles, while the high-energy particles penetrate deeper into the upper atmosphere. The nature of this interaction provides us with an opportunity

to examine the geomagnetic field in the orbital environment of HST by extracting and analyzing cosmic rays in HST observations.

The geomagnetic field is described by a dipole with higher-order terms that reflect short-term variations. These short-term variations occur on timescales of seconds to years, and their contributions produce nonuniform departures in direction and intensity from the dipole field (Thébault et al. 2015). The most well-known departure is a region known as the South Atlantic Anomaly (SAA), where the local magnetic field measured at the Earth’s surface is significantly weaker (Schaefer et al. 2016). The decreased magnetic field intensity of the SAA reduces shielding from high-energy charged particles. This increases the amount of radiation damage to spacecraft in low Earth orbit, which includes the International Space Station (ISS) and its human inhabitants. The orbital trajectory of HST regularly crosses the SAA, and observations along the boundary are used to help map the extent of SAA at an altitude of ~ 540 km, assess its effects on observations, and to provide a means of monitoring the efficacy of the SAA avoidance contours used in the scheduling of HST observations (Lupie 2002; Martel et al. 2009; Barker et al. 2010).

Over the course of the 11 yr Solar Cycle, fluctuations in solar activity modulate the flux of galactic cosmic rays (GCRs; Potgieter et al. 2013); increased solar activity results in more scattering and decreases the GCR flux at Earth and vice versa. The first evidence of this modulation was observed in data obtained with ion chambers at four stations (Forbush 1954) and then by a neutron monitoring system that was established in 1958 for the International Geophysical Year (Simpson 1958). The neutron monitoring system provided continuous and

⁵ Current Address.

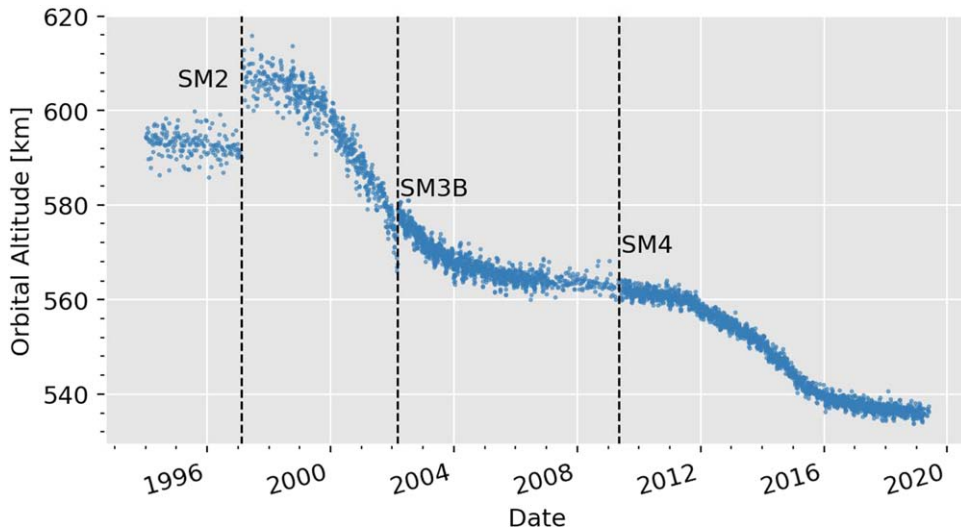


Figure 1. The orbital altitude of HST as a function of time computed from telemetry data. Vertical lines indicate servicing missions where instruments were added and/or removed. The observatory received two re-boasts, one in Servicing Mission 2 (SM2) and one in Servicing Mission 3B (SM3B), to compensate for orbital decay. The periods of accelerated orbital decay are driven by contemporaneous solar maxima in Solar Cycles 23 and 24. The increased solar activity increases the density in the thermosphere, which extends from 60 to 750 km and encompasses the orbit of HST. This leads to an increase in satellite drag and an accelerated orbital decay (Walterscheid 1989).

Table 1
Properties of the CCD Imagers Analyzed

Instrument	Detector Size (cm ²)	Epitaxial Layer Thickness (μ m)	Operational Period	Detector Type
WFPC2	5.76	~10	01/1994–05/2009	F
STIS/CCD	4.624	13.24–14.83	02/1997–08/2004, 05/2009–	B
ACS/HRC	4.624	12.49–16.03	03/2002–01/2007	B
ACS/WFC	37.748	12.60–17.10	03/2002–01/2007, 05/2009–	B
WFC3/UVIS	37.804	13.50–18.00	05/2009–	B

Note. For CCDs with multiple chips (i.e., ACS/WFC, WFPC2, and WFC3/UVIS), the detector size is the combined area of all the chips. For the detector type, “F” corresponds to thick, frontside-illuminated CCDs and “B” corresponds to thin, backside-illuminated CCDs.

standardized measurements of the GCR flux and confirmed the modulation (Lockwood 1958). By the 1990s, the 11 yr cycle and the interplanetary origin of the GCR intensity variation was firmly established when data were obtained from satellites and space probes (Simpson 1994). By analyzing charged particle rates in HST observations, we examine the effects of solar activity on the overall incident particle flux as a function of time in the orbital environment of HST.

As described in Schnyder et al. (2017a, 2017b), we deployed a cloud-based distributed computing platform for processing HST observations to identify and analyze cosmic rays. This early implementation was based on Python wrapped around the IRAF command language, and served as a proof of concept. However, the IRAF command language is not designed for batch processing of large data sets and is currently no longer supported. Because of this, we developed the open-source Python package, *HSTCosmicrays*, to identify and analyze cosmic rays in HST images.

In this paper, we describe our software and highlight key results that demonstrate our ability to reliably identify and analyze cosmic rays. In Section 2, we describe the ~ 4.6 TB of HST calibration data used in the analysis. In Section 3, we provide an overview of the pipeline, describe the cosmic ray rejection algorithms, and list the data extracted. In Section 4, we present the results of analyzing more than 1.2 billion cosmic rays observed over a period of ~ 25 yr, which includes the following:

cosmic ray morphology in Section 4.1, cosmic ray track lengths in Section 4.2, the observed cosmic ray particle fluxes in Section 4.3, cosmic rays as a proxy for detector thickness in Section 4.4, modulation by the solar cycle in Section 4.5, cosmic rays in the SAA in Section 4.6, spatially correlated cosmic ray “hot spots” in Section 4.7, and an estimation of the observed cosmic rays’ average kinetic energy in Section 4.8. A detailed geophysical analysis is beyond the scope of this paper.

2. Data Set

In this work, we analyzed images taken with five different CCD imagers on four instruments: Wide Field Planetary Camera 2 (WFPC2), Space Telescope Imaging Spectrograph (STIS), the High Resolution Channel (HRC) and Wide Field Channel (WFC) in the Advanced Camera for Surveys (ACS), and the Wide Field Camera 3 (WFC3) UVIS channel. In Table 1, we list the detector characteristics relevant to our analysis. In Table 2, we show the breakdown of the dataset across the five imagers.

The detector size and thickness determine the overall cross section for interaction between cosmic rays and the detection layer of the CCD substrate, and this allows us to probe the detector properties. The combined period of operation of the five imagers provides continuous coverage of Solar Cycle 23 and Solar Cycle 24 (Figure 2), enabling us to probe the effects of solar activity on the cosmic ray flux at HST’s orbital altitude.

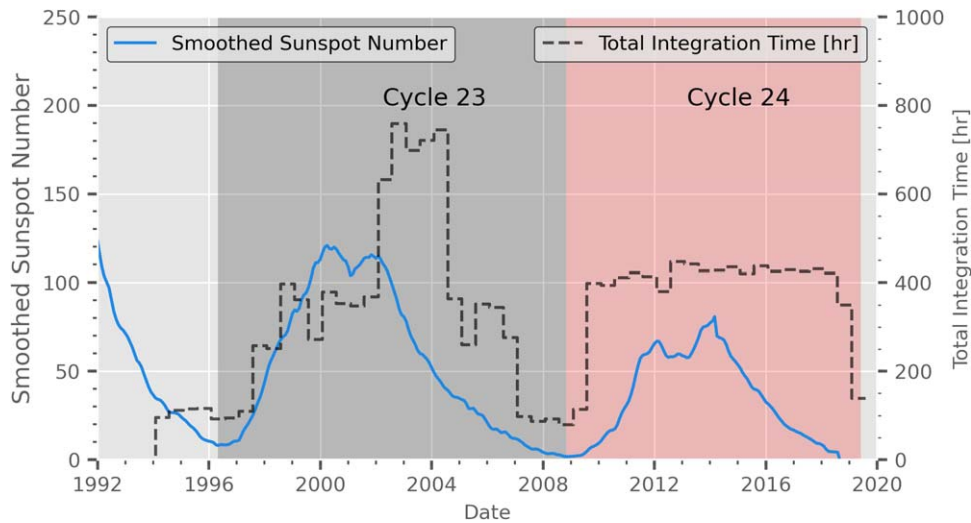


Figure 2. The 13 month smoothed, monthly sunspot number reported by the the WDC-SILSO, Royal Observatory of Belgium, Brussels (SILSO World Data Center 1991–2019) The gray shading indicates Solar Cycle 23, and the red shading indicates Solar Cycle 24. The dashed, black line is a histogram of the integration times for all instruments in 6 month bins. On average, the data set contains roughly ~ 1.8 hr of HST observations every day from 1994 to 2019.5, with complete coverage of Solar Cycles 23 and 24. The significant drop starting in 2005 is due to the STIS failure in 2004 August and the ACS failure in 2007 January (see Table 1).

Table 2
The Data Set for Each Imager

Instrument	Image Count	Data Volume (TB)	Total EXPTIME (hr)
WFPC2	13,317	0.131	5098
STIS/CCD	31,430	0.311	3765
ACS/HRC	5477	0.055	1462
ACS/WFC	13,311	2.130	3498
WFC3/UVIS	12,373	1.980	3040
Totals	75,908	4.607	16,863

We restrict our analysis to dark calibration frames (hereafter referred to as darks). Darks are images taken with the shutter closed, and they are used to quantify and remove the thermal noise (or dark current) present in the CCDs (Janesick 2001). As part of the standard calibration, each dark is gain-calibrated to convert from units of DN to units of electrons and has been bias-corrected and dark-subtracted. Since the shutter is closed, charged particles are the only external sources present. Thus, any signal above the background noise level is due to the interaction between charged particles and the silicon atoms in the epitaxial layer. This facilitates the identification process by completely eliminating any chance of confusion with transient astrophysical sources. Additionally, the entire data set of darks is available in the cloud as part of the HST Public Data set⁶ hosted on Amazon Web Services (AWS). This gives us the ability to leverage the compute resources and network infrastructure of AWS to boost the performance of our software.

3. HSTcosmicrays

3.1. Pipeline Overview

HSTcosmicrays is written entirely in Python and is available on Github.⁷ We optimize runtime with `dask` (Dask Development Team 2016), a Python parallelization framework.

⁶ <https://registry.opendata.aws/hst/>

⁷ https://github.com/nmiles2718/hst_cosmic_rays

We store the results for each data set in HDF5 (The HDF Group 2010) format and the package contains a module for reading and writing of data. In order to handle the 4.6 TB data volume associated with all 75,908 images, the pipeline was designed to be lightweight and modular, and the benefits of this are twofold.

First, the storage requirements are now tied to the data generated and not the data downloaded. This means that users who wish to reproduce this analysis will be able to do so, provided they have ~ 170 GB of hard disk space available to store the results—a requirement that is met by most laptops nowadays. Second, when leveraging AWS to perform the analysis, costs are minimized because there is no need to allocate additional Elastic Block Store volumes to accommodate all 4.6 TB of darks.

The pipeline consists of five distinct steps (Figure 3) that are applied to consecutive one-month intervals of darks. For each interval, we use `astroquery` (Ginsburg et al. 2018) to programmatically query and download the bias-corrected dark frames along with their engineering and telemetry files from the Mikulski Archive for Space Telescopes (MAST). Once downloaded, the images are processed through the cosmic ray rejection and identification steps. After processing, the identified cosmic rays are analyzed, resulting in a catalog of parameters describing their morphology. After the analysis has completed, the results are written to file and an (optional) email is sent to the user-defined email address with summary statistics for all the cosmic rays found in each image. Finally, all downloaded images and temporary files are deleted to prepare for the next one-month chunk of darks.

3.2. Cosmic Ray Identification

Unlike other external sources (e.g., stars, galaxies), cosmic rays are unaffected by the optics of the telescope. Hence, there is no prescription for determining *a priori* how the energy deposited by a cosmic ray is distributed among the pixels that it affects. For unresolved sources, one can use knowledge of the PSF to determine how much flux, from two (or more) blended sources, is present in the pixels they share in common. For

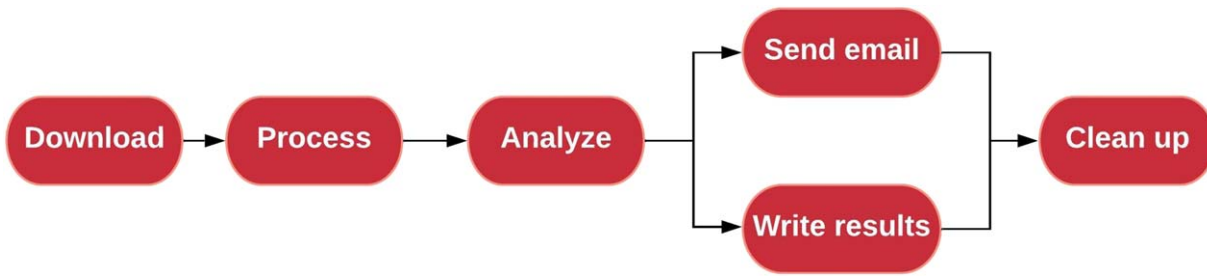


Figure 3. Schematic overview of the processing and analysis pipeline.

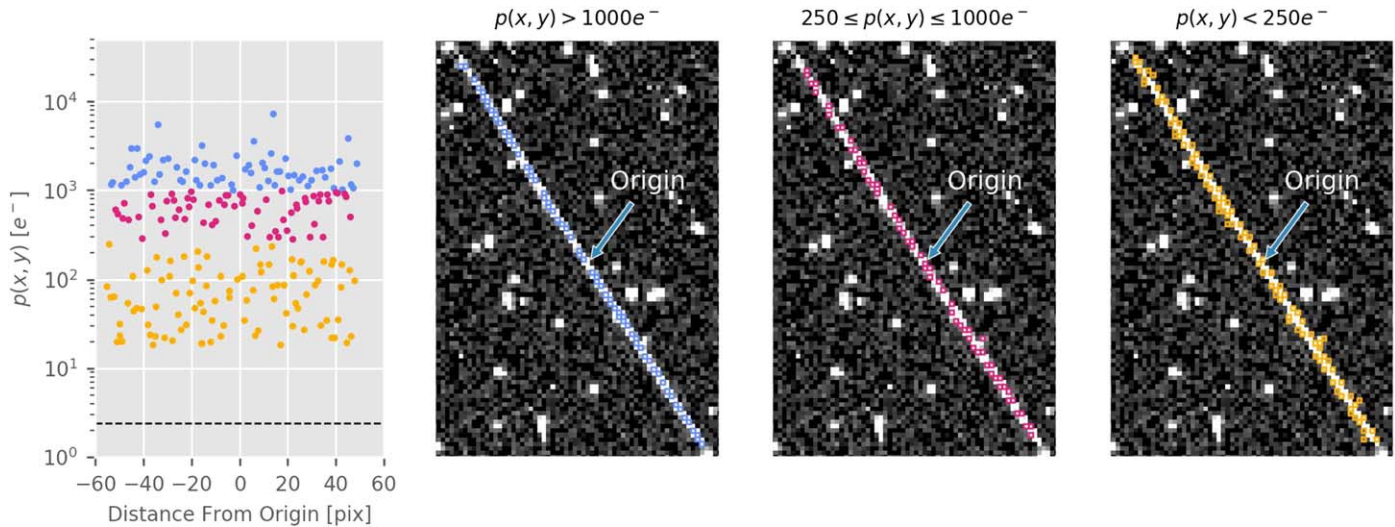


Figure 4. Cosmic ray track observed in a STIS dark image, o3s101pcq_fit.fits. The scatter plot on the left shows the pixel values as a function of their distance to the origin marked in the image cutouts. We color-code the points into three distinct groups; pixel values below $250 e^-$ are yellow, pixel values between $250 e^-$ and $1000 e^-$ are magenta, and pixel values greater than $1000 e^-$ are blue. In the three image cutouts, we show the location of pixels belonging to the three groups. The pixel values associated with this single cosmic ray span two orders of magnitude from tens of electrons to thousands of electrons, and there is no obvious “profile” for how the electrons are distributed. By comparing the three bins of pixel values and their locations along the cosmic ray track, it is easy to see why any deblending algorithm will incorrectly segment this single object into numerous sources.

charged particles (e.g., cosmic rays), the energy deposited in a given pixel is probabilistic and so the energy received by adjacent pixels from the same cosmic ray fluctuates. In Figure 4, we highlight an example of this by showing an elongated cosmic ray observed in a STIS dark, o3s101pcq_fit.fits, taken in 1997.

For very elongated cosmic rays, these pixel-to-pixel fluctuations can be large enough such that typical deblending software (e.g., *Source Extractor*) will mistakenly segment a single elongated cosmic ray into numerous smaller ones. In the same vein, spatially coincident cosmic rays lack the contrast required to reliably identify a local minimum in their combined source profiles, making it extremely difficult to separate the two objects. For these reasons, we do not attempt to deblend spatially coincident cosmic rays, as this does more harm than good by introducing a non-negligible number of false positives from oversegmentation, which artificially raises the observed cosmic ray particle flux.

In the following subsections, we elaborate on the two techniques used to identify and label cosmic rays. We describe the process for instruments that are currently operationally active on HST, hereafter referred to as the active instruments, and those that are no longer operationally active, hereafter referred to as the retired instruments. The active instruments have robust calibration software that is available for use by the astronomical community through Python, which greatly

simplifies the process of identifying cosmic rays. The retired instruments lack similar calibration software, and so a different method is utilized.

3.2.1. Active Instruments

For the active instruments, ACS, STIS, and WFC3, we use their Python packages, *acstools*, *stistools*, and *wfc3tools*, to run their cosmic ray rejection routines *ACSREJ*, *OCCREJECT*, and *WF3REJ*, respectively. As part of the cosmic ray rejection step, the Data Quality (DQ) extension of each input file is updated to indicate which pixels were affected by cosmic rays. Each routine implements a noise-based rejection model that looks for statistically significant outliers in observations made in sequence, at the same exact pointing. Because the images are taken at slightly different times and the same exact pointing, actual (nontransient) sources (stars and galaxies) will have the same detector position in each pointing, but cosmic rays and other transient artifacts will not. This allows for an easy identification of cosmic rays. Here, we provide a review of the underlying algorithm utilized by the active instruments to identify pixels affected by cosmic rays.

The algorithm works by first performing a global background subtraction using the mode for each image. Next, the initial guess for the cosmic ray cleaned image is computed by taking the minimum or median across the stack of background-subtracted images. For each image in the stack, a pixel-wise

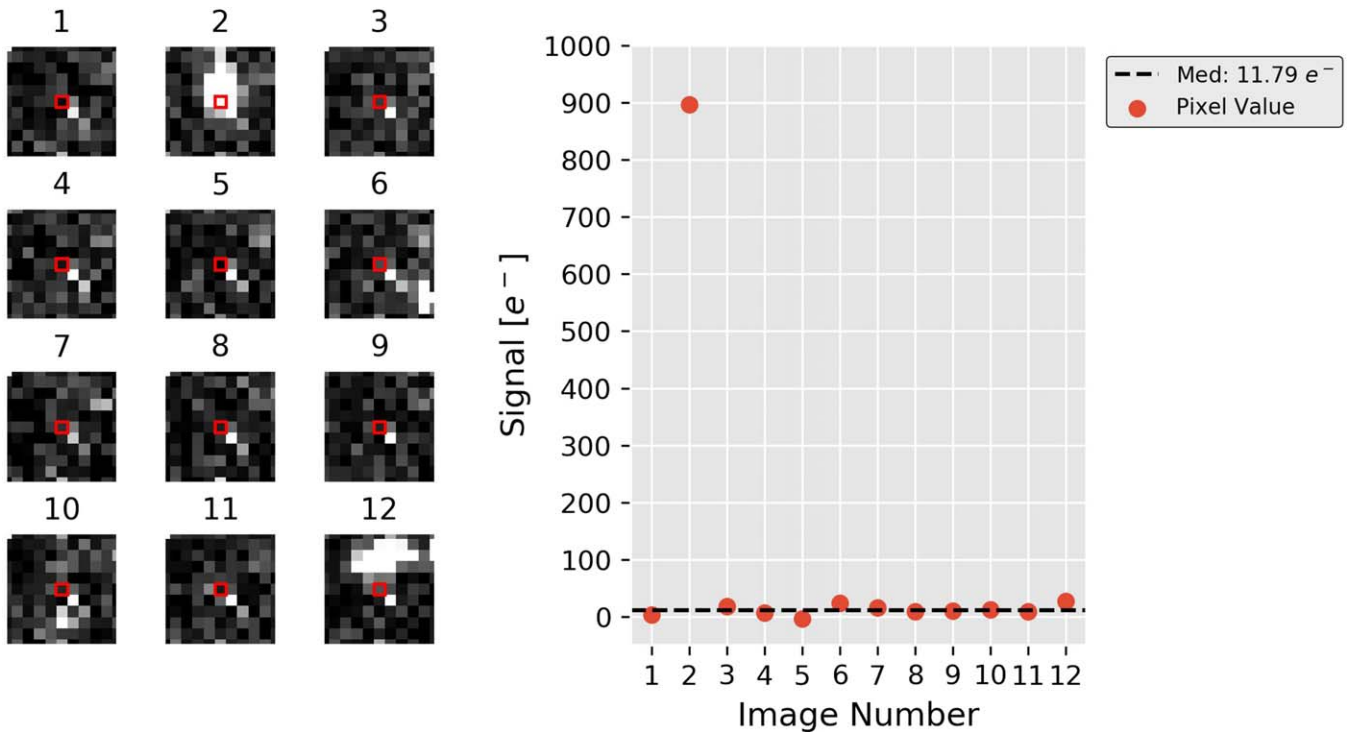


Figure 5. A visual representation of the cosmic ray rejection algorithm. Left: 10 pixel by 10 pixel cutouts from 12 different ACS/WFC images centered on the pixel marked by a red square. Right: The value of each pixel marked by the red square in the 12 different cutouts. The dashed vertical line is the median of all 12 values and represents the initial guess for the true value of the pixel marked by the red square. The only image where the pixel was marked as a cosmic ray by the ACS/WFC cosmic ray rejection routine, ACSREJ, is image number 2.

comparison with the initial guess is performed to determine which pixels have been affected by cosmic rays. If $p_n(x, y)$ is the value of the pixel at position (x, y) in the n th image, T_n is the exposure time of the n th image, sky_n is the sky background of the n th image, and $p(x, y)$ is the value of the same pixel in the comparison image, then the variance with respect to the comparison image is computed as:

$$\Delta_n(x, y) = \frac{[p_n(x, y) - (\text{sky}_n + p(x, y))]^2}{T_n^2}. \quad (1)$$

This value is then compared to the expected variance, $\tau_n(x, y)$, for the given pixel,

$$\tau_n(x, y) = \left(\frac{\sigma^2}{T_n^2} \right) [\sigma_{RN}^2 + \sigma_p^2(x, y) + \lambda(p(x, y) - \text{sky}_n)^2], \quad (2)$$

where σ is a number representing the required level of significance (e.g., 3 or 5), σ_{RN} is the read noise of the amplifier used to read out the pixel at (x, y) , $\sigma_p(x, y)$ is the Poisson noise in comparison image, and the last term, $\lambda(p(x, y) - \text{sky}_n)^2$, is used to accommodate the undersampled PSF of the HST imagers. In Figure 5, we show an example of the process on a set of ACS/WFC darks.

If $\Delta_n(x, y) > \tau_n(x, y)$, then the pixel at (x, y) is marked with a special bit flag in the data quality (DQ) extension to indicate that it has been affected by a cosmic ray. If a pixel is identified as cosmic ray contaminated, the rejection criteria is applied to the neighboring pixels with a stricter σ value. Finally, if

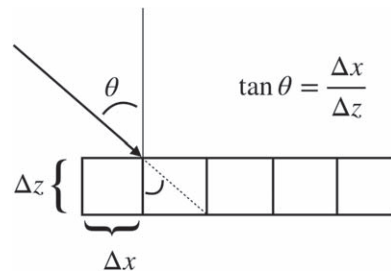


Figure 6. Cartoon depiction of the interaction between a cosmic ray and the pixel grid of a CCD. The trajectory of the cosmic ray is denoted by the arrow, Δx is the pixel width, and Δz is the pixel thickness. The angle of incidence, θ , is defined by the relationship in the figure.

multiple values are supplied for the σ parameter, the algorithm will be applied in an iterative manner.

After running the cosmic ray rejection routine, the DQ extension of each input file has been updated with a special BIT flag, 8192, to indicate which pixels are affected by cosmic rays. Key to our work is utilizing this information to construct a binary image suitable for connected-component labeling analysis. We perform a bitwise-AND comparison between the DQ extension and 8192 to generate a binary image where any cosmic ray affected pixel is marked by 8192, everything else is marked with 0. We use the 8-connectivity matrix (Equation (3)) to identify all groups of connected pixels affected by the same cosmic ray. Any object identified that affects 1 or 2 pixels is rejected. This allows for a robust rejection of any unstable, hot pixels (Borncamp et al. 2017) identified during the cosmic ray rejection step that have large

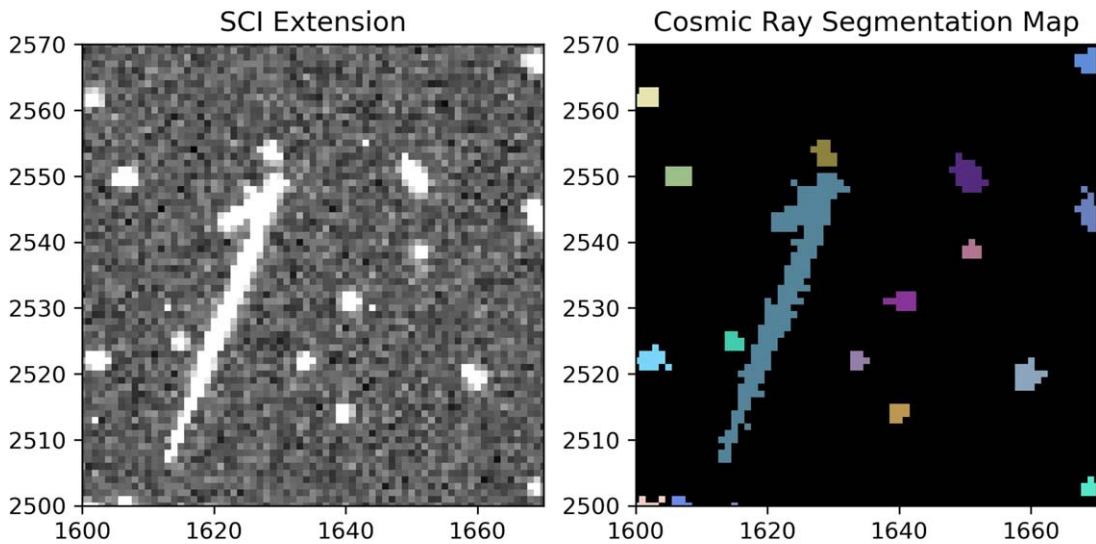


Figure 7. The SCI extension of an arbitrary STIS/CCD dark frame and the cosmic ray segmentation map produced by analyzing the DQ array associated with the SCI extension. As discussed in the beginning of this section, we do not attempt to deblend overlapping cosmic rays.

fluctuations in dark current that can be mistakenly classified as a cosmic ray by noise-based cosmic ray rejection algorithms like ACSREJ.

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & p & 1 \\ 1 & 1 & 1 \end{bmatrix}. \quad (3)$$

Filtering also removes potential single-pixel and two-pixel cosmic rays events. For a single-pixel event to occur on one of the HST CCD detectors, the charged particle (ignoring charge diffusion effects) must physically traverse only a single pixel. In Figure 6, we show a cartoon depiction of the scenario where the path of the particle through the pixel is maximized. The thin, vertical line denotes the normal to the CCD’s surface, and it extends from the boundary between two pixels.

The particle trajectory denoted by the arrow will traverse a single pixel on an HST CCD if the angle incidence is in the interval defined by $[-\theta, 0) \cup (0, +\theta]$, where $\theta = \arctan \Delta x / \Delta z$. For the case shown in Figure 6, as θ approaches 0° , the probability of electron-hole pairs generated in one pixel crossing the pixel boundary via charge diffusion increases (Hopkinson 1987), turning some single-pixel events into multi-pixel events.

The CCDs used by HST instruments are located off the main optical axis, and pick-off mirrors are used to redirect light to each instrument housing. HST’s pointing changes to observe astronomical sources, but during internal observations, e.g., dark frames, the telescope pointing is unconstrained. Hence, the orientation of the CCDs with respect to the average direction of the particle flux is constantly changing. Further, cosmic rays originate outside the magnetosphere, and so their trajectories are significantly altered by the geomagnetic field as they are confined to move along field lines. These factors combine to significantly reduce the probability of single- or two-pixel cosmic ray events in images taken with the HST CCDs.

In Figure 7, we show a cutout of the SCI extension and the corresponding segmentation map generated by labeling the DQ array for an arbitrary STIS/CCD dark. The colors indicate the distinct groups of pixels affected by individual cosmic rays.

3.2.2. Retired Instruments

Because the calibration software for the retired instruments, i.e., WFPC2, is deprecated, a different process is used. Each image is analyzed individually and a hybridization of binary thresholding and connected-component labeling, hereafter referred to as “threshold labeling,” is used to identify cosmic rays. We perform three iterations of sigma clipping to compute the average pixel value, $\langle p \rangle$, and compute a robust measure of spread using the median absolute deviation (MAD). Using the sigma-clipped mean and the MAD, we create a binary image by marking all pixels according to the following condition:

$$\begin{aligned} p(x, y) &\geq \langle p \rangle + 5 * \text{MAD} \\ 1 &\text{ If the condition is True.} \\ 0 &\text{ If the condition is False.} \end{aligned} \quad (4)$$

Next, we run the connected-component labeling analysis to identify groups of pixels with anomalously high values. We reject any object identified that affects 1 or 2 pixels, to remove potential hot pixels, and classify the rest as cosmic rays. We show an example of the resulting segmentation map generated by threshold labeling for WFPC2 in Figure 8.

3.3. Comparing Threshold Labeling and LAcosmic

To compare the results from threshold labeling and LAcosmic (van Dokkum 2001) relative to the “ground truth” of the calibration pipeline’s cosmic ray rejection algorithm, we use dark frames obtained with ACS/WFC.

We utilize the *astropy* (The Astropy Collaboration et al. 2018) implementation of LAcosmic, *astrocrappy*⁸ (McCully et al. 2018). Using data from the ACS/WFC, we run the cosmic ray rejection algorithm, ACSREJ, and use our method of labeling the DQ extension explained in Section 3.2.1 to define the “true” cosmic ray segmentation map. Next, we run LAcosmic with the default parameters to generate a cosmic ray mask and apply connected component labeling to create a second cosmic ray segmentation map. We generate a third cosmic ray segmentation map using the threshold labeling

⁸ <https://astrocrappy.readthedocs.io/en/latest/>

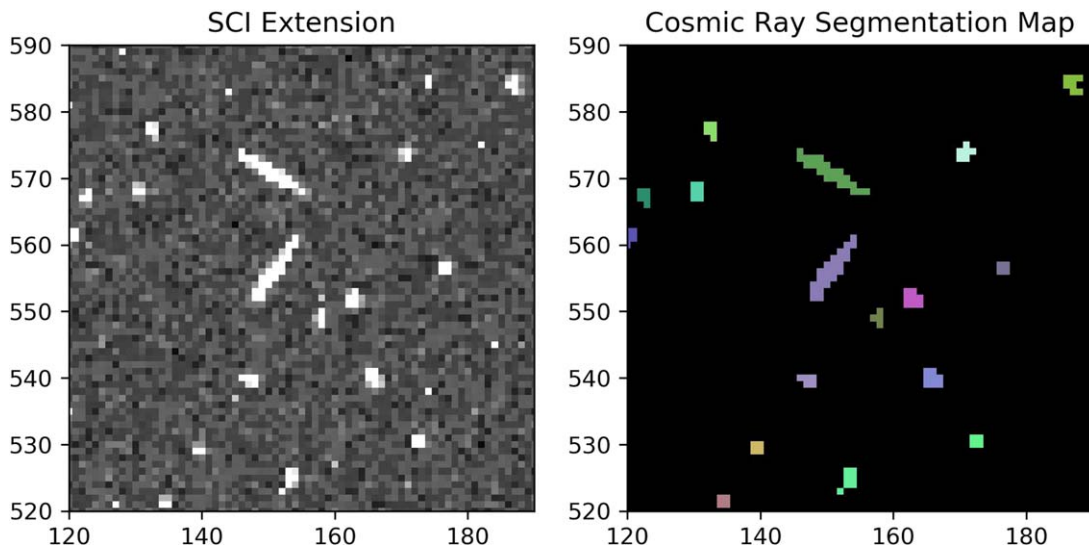


Figure 8. Left: The SCI extension of a WFPC2 dark frame. Right: The corresponding cosmic ray segmentation map produced by analyzing three sigma outliers. As expected, there are no hot pixels identified in the final segmentation map.

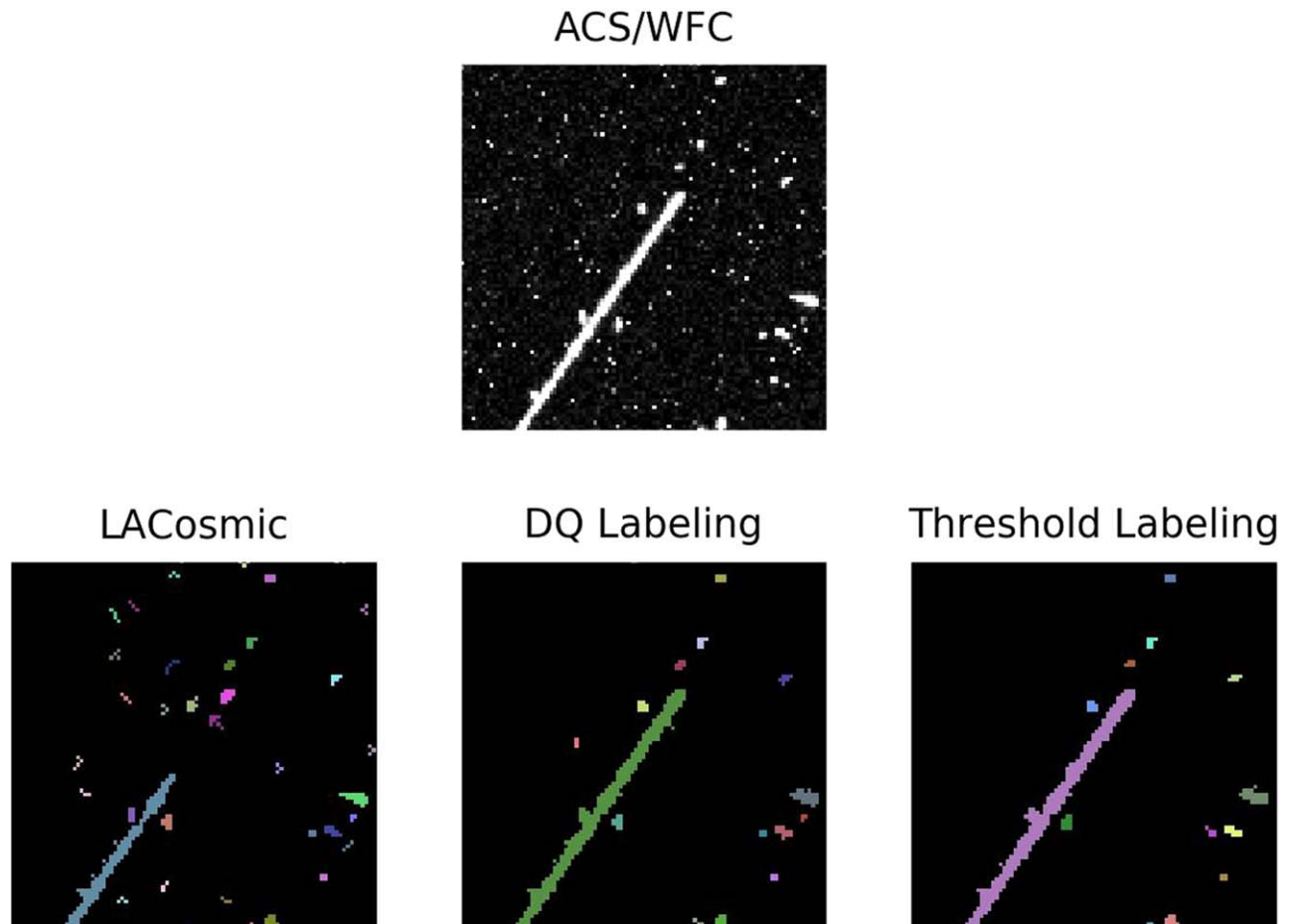


Figure 9. Visual comparison of the segmentation maps generated by the three methods applied to an ACS/WFC dark frame, j8jcnzv1q_ft.fits. Top: A 100 pixel by 100 pixel subsection of the ACS/WFC image. Bottom: The same region in the segmentation map generated by each algorithm.

algorithm. In Figure 9, we show a subsection of an ACS/WFC image and the corresponding segmentation maps generated by the three methods.

We perform a pixel-wise comparison to determine the efficacy of each algorithm in identifying pixels affected by cosmic rays. We define \mathbb{A} to be the set of all pixels identified by

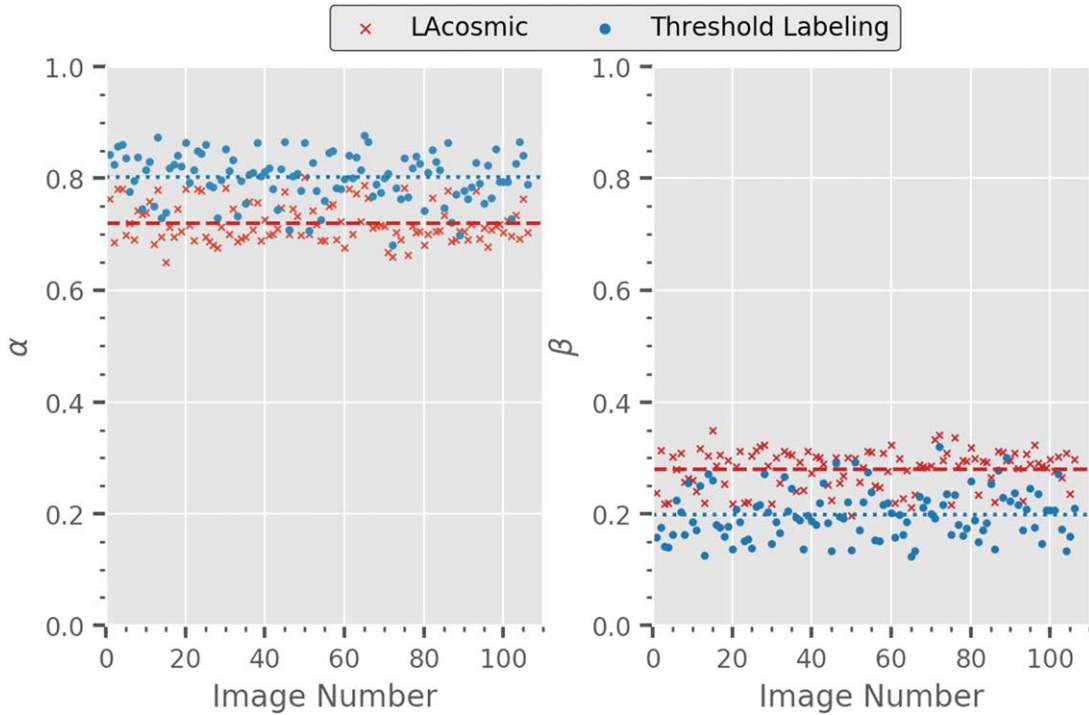


Figure 10. Left: The α value (Equation (5)) for 106 ACS/WFC images. Right: The β value (Equation (6)) for 106 ACS/WFC images.

ACSREJ and \mathbb{C} to be the set of all pixels identified by either LACosmic or threshold labeling. We use the intersection, \cap , and set difference, \setminus , operators in conjunction with the sets defined above to compute the following two parameters:

$$\alpha = \frac{\mathbb{C} \cap \mathbb{A}}{|\mathbb{C}|} \quad (5)$$

$$\beta = \frac{\mathbb{C} \setminus \mathbb{A}}{|\mathbb{C}|}. \quad (6)$$

The first parameter, α , is the fraction of cosmic ray affected pixels that were correctly identified by each algorithm. The second parameter, β , is the fraction of cosmic ray affected pixels that were incorrectly identified by each algorithm. In Figure 10, we show α and β for the 106 images analyzed. LACosmic correctly identified an average of 72% of the cosmic ray affected pixels, and threshold labeling correctly identified an average of 80%.

We find that LACosmic and threshold labeling identified an average of 15% and 22% more pixels, respectively, than ACSREJ. For LACosmic, false positives accounted for nearly 30% of all pixels identified, while for threshold labeling, they accounted for 20%. When analyzing darks with LACosmic, individual hot pixels are routinely misclassified as cosmic rays because they have similarly “sharp” edges. These misclassifications result in a large number of false positives, increasing the total number of detected cosmic rays by 66% across the 106 dark frames.

The aim of our analysis is to study cosmic rays, so we adopt the threshold labeling algorithm for identifying cosmic rays in the WFPC2 darks. Compared to LACosmic, threshold labeling correctly identifies more true cosmic ray affected pixels with fewer false positives. However, because darks do not image external, astrophysical sources, our results should not be interpreted as the ability of LACosmic to correctly distinguish between cosmic rays and external sources with broader profiles (e.g., stars), a regime where it excels.

3.4. Analyzing Cosmic Rays

Once the cosmic ray label has been generated, we apply the label to the SCI extension in order to derive quantities of interest. For each cosmic ray, we compute the following moments of the distribution of the deposited energy, as defined in Riess (2002):

1. $I_0 = \sum_i p_i$.
2. $I_x = \frac{1}{I_0} \sum_i p_i * x_i$.
3. $I_y = \frac{1}{I_0} \sum_i p_i * y_i$.
4. $I_{xx} = \frac{1}{I_0} \sum_i p_i (x_i - I_x)^2$.
5. $I_{yy} = \frac{1}{I_0} \sum_i p_i (y_i - I_y)^2$.
6. $I_{xy} = \frac{1}{I_0} \sum_i p_i (x_i - I_x) * (y_i - I_y)$.

where p_i is the pixel value of the i th pixel in the cosmic ray label and x_i, y_i are the x and y coordinates of the i th pixel, respectively. The first parameter, I_0 , is the total energy deposited by the cosmic ray in units of electrons. The second and third parameters combine to give the centroid of the cosmic ray, (I_x, I_y) . Using the second moments, I_{xx} and I_{yy} , we compute the width or “size” of the distribution of the deposited energy as

$$\text{size} = \sqrt{\frac{I_{xx} + I_{yy}}{2}}. \quad (7)$$

Finally, we use the second moments to assess the symmetry of the distribution of the deposited energy:

$$\text{shape} = \sqrt{\frac{(I_{xx} - I_{yy})^2 + 4I_{xy}^2}{(I_{xx} + I_{yy})^2}}. \quad (8)$$

For comparison, when applying Equations (7) and (8) to a 2D Gaussian with equal variance along both x and y (i.e., $\sigma_x = \sigma_y = \sigma$), Equation (7) yields σ and Equation (8) yields 0,

Table 3
The Number of Detected Cosmic Rays per Instrument

Instrument	CR Count
WFPC2	126,322,987
STIS/CCD	61,717,583
ACS/HRC	24,796,064
ACS/WFC	558,517,641
WFC3/UVIS	526,545,187
Total	1,287,061,978

indicating perfect symmetry (Miles et al. 2020). We compute the total number of pixels affected by each cosmic ray and record this as another metric for the “size” of the cosmic ray. We compute the cosmic ray particle flux as the total number of individual cosmic rays identified divided by the total integration time and the size of the detector from Table 1. The total integration time is defined as the exposure time plus half of the detector readout time, which accounts for cosmic rays that strike the detector during readout. For each cosmic ray identified, we record the coordinates of all the affected pixels.

Finally, for each image analyzed, we use the engineering and telemetry files (*spt.fits) (Desjardins & Lucas 2019) to extract the following metadata:

1. Altitude.
2. Latitude.
3. Longitude.
4. Observation date.
5. Observation start time.
6. Observation end time.
7. Telescope pointing (World Coordinate System (WCS) information).

4. Results

4.1. Cosmic Ray Morphology in HST Images

Cosmic ray morphological properties are of broad interest to the astronomical community because they can be used to discriminate astrophysical transients from cosmic ray events. In Table 3, we report the total number of detected cosmic rays per instrument that we have analyzed thus far.

In general, the morphology of cosmic rays in the two types of CCD detectors analyzed, i.e., thick, frontside-illuminated versus thin, backside-illuminated, is highly consistent. In all the detectors, they appear in a variety of shapes and sizes, from elongated to point-like, and are randomly distributed across the detector. In Figure 11, we show the morphology matrix for cosmic rays identified in a single ACS/WFC dark frame with an exposure time of 1000 s. The cosmic rays shown were randomly sampled from their corresponding distributions defined by the size and shape constraints for the given row and column. The top left corner corresponds to the smaller and more symmetric cosmic rays, while the bottom right corner corresponds to larger and more elongated cosmic rays. Where a given cosmic ray falls in the morphology matrix is almost entirely determined by its angle of incidence with respect to the normal of the plane of the CCD.

At nearly normal incidence, cosmic rays deposit energy in a symmetric manner, leading to a small value for the shape parameter (Equation (8)) and a roundish appearance. While at

oblique incidence, energy is deposited asymmetrically, resulting in large shape and size parameter values and elongated appearance. We postulate that objects in the lower left corner of the matrix are the result of a very high-energy cosmic ray interacting with atoms deep within the silicon substrate, far from the gate that typically traps electrons generated by the photoelectric effect. When electrons are generated far from the gate, regardless of their origin, they are more likely to diffuse into adjacent pixels through a process known as charge diffusion (Hopkinson 1987).

A consequence of multiple angles of incidence, is that the distributions of the computed morphological parameters are highly asymmetric with a positive skew. Table 4 provides the summary statistics for each of the morphological parameters extracted. For each parameter, we report the 50th percentile and the interval bounded by the 25th and 75th percentiles. The typical cosmic ray affects 7 pixels and deposits about 2700 electrons.

4.2. Computing Particle Path Lengths

The number of contiguous pixels affected by a single charged particle event represents the projected path length. For each image, we reconstruct the cosmic ray segmentation map and identify the smallest box that encloses each cosmic ray. The diagonal of the box provides an estimate of the projected path length. Figure 12 demonstrates this technique for one STIS/CCD dark frame.

Once the projected path length has been computed, we calculate the actual path through the detector as shown schematically in Figure 13. The projected path of the particle measured by the CCD pixels is shown on the left. The actual path *through* the detector is shown on the right. We assume the pixel boundaries intersect the top and bottom surfaces of the CCD at 90° angles. Hence, the distance traversed through the detector is $\text{Path Length} = \sqrt{(\text{Projected Path})^2 + (\text{Thickness})^2}$. We use the values listed in Table 5 to convert from units of pixels to units of micrometers.

The distribution of path lengths is most consistent with an isotropic particle flux modulated by strong shielding at large angles of incidence. Using the trigonometric relationships in Figure 13, the probability of a particle having a path length, t , is proportional to

$$A \sin(\theta) \cos^n(\theta) \frac{d\theta}{dt} = A/t^{(n+2)}, \quad (9)$$

where A is the normalization factor and n represents the strength of shielding. The best-fit probability density function for the entire data set is $A/t^{(n+2)}$ with $n = 2.33$, $A = 3.83 \times 10^6$.

Figure 14 shows the best-fit distribution, as well as the distribution of path lengths for each of the five CCD imagers. Although the distributions are similar, they are not identical, and the distribution for WFPC2 notably deviates from that of the other imagers at around 200 μm , corresponding to a path length of about 13 pixels, possibly due to the location of its detectors in the observatory.

4.3. Cosmic Ray Particle Flux

The flux of cosmic rays at 1 AU in the Solar System has been extensively studied by numerous ground- and space-based experiments. In Figure 15, we show the differential intensity of cosmic rays measured by the Payload for Antimatter Matter

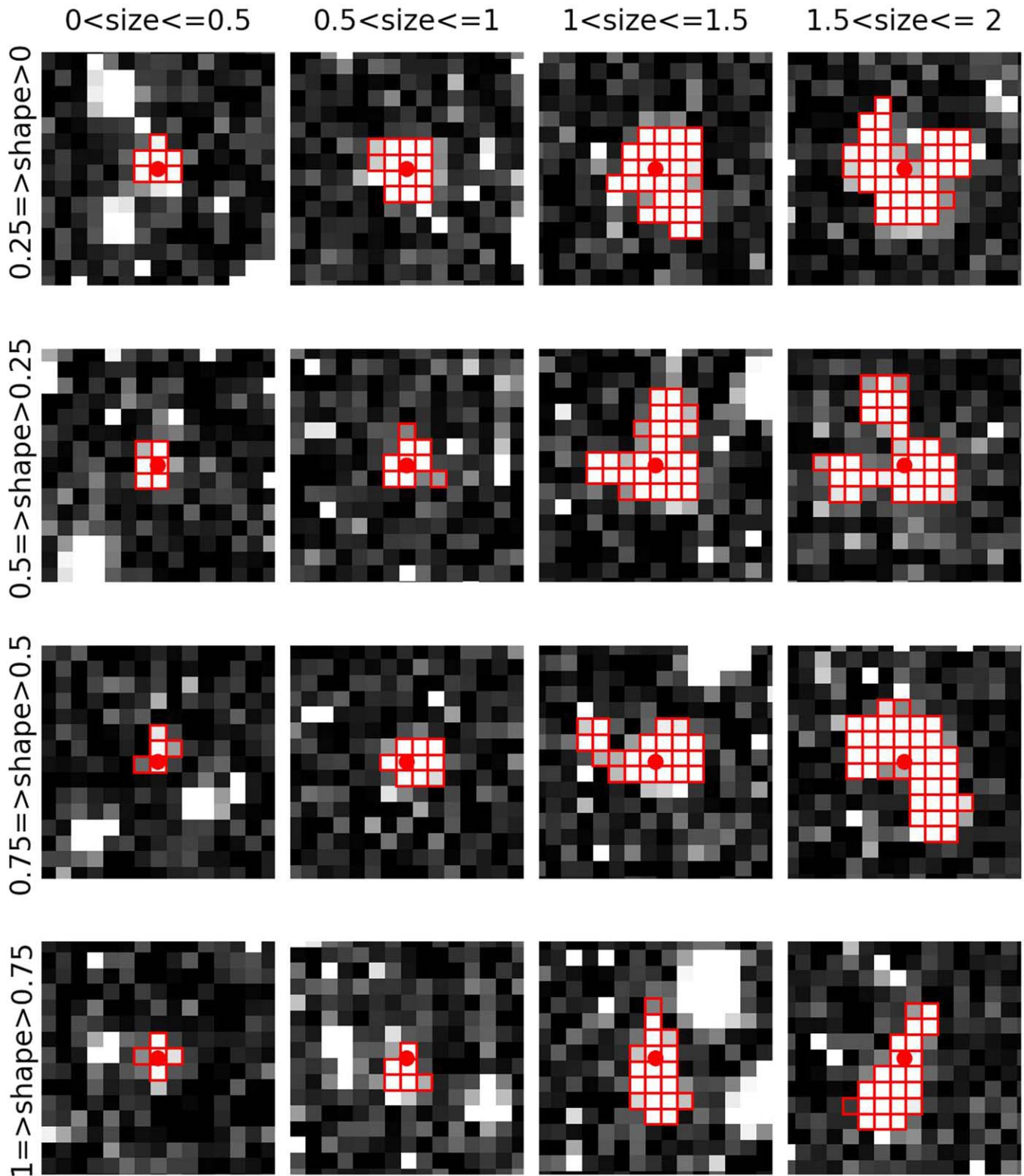


Figure 11. The morphology matrix for cosmic rays identified in a single ACS/WFC dark frame. In each subplot, the cosmic ray and the pixels it affected are marked by red squares. Each row corresponds to the shape bin denoted on the left, while each column corresponds to the size bin denoted on the top. The filled, red circle marks the computed centroid $((I_x, I_y))$ of the cosmic ray.

Exploration and Light-nuclei Astrophysics (PAMELA) experiment. The data were obtained from the Cosmic Ray Database (Di Felice et al. 2017) hosted by the Space Science Data Center. The flux of cosmic rays is primarily dominated by protons and helium at $\sim 90\%$ and $\sim 9\%$, respectively. The

remaining 1% is comprised of heavier nuclei and antimatter. While HST does not have a dedicated detector capable of distinguishing particle types, we may use the CCDs to estimate the bulk cosmic ray particle flux in units of $\text{particles s}^{-1} \text{cm}^{-2}$ at HST's orbital altitude.

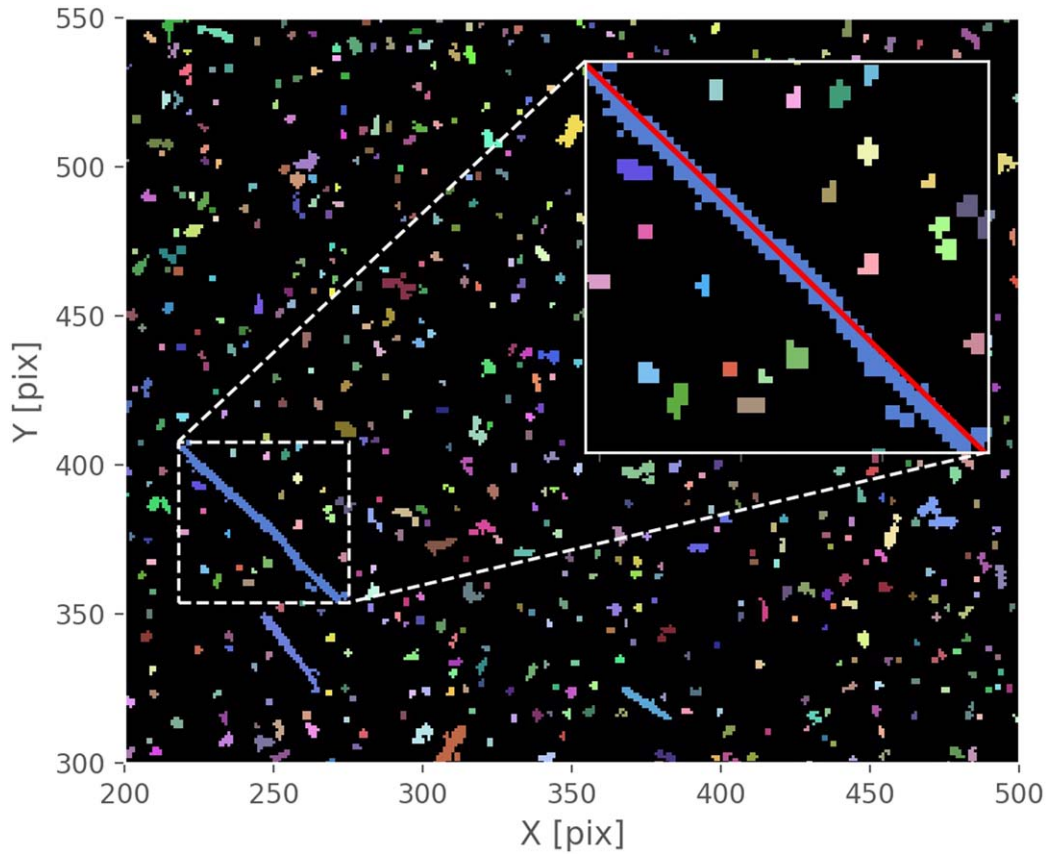


Figure 12. A visual representation of the projected path computed for a single, elongated cosmic ray in a reconstructed STIS/CCD cosmic ray segmentation map. The dashed, white box is the smallest box that fully encloses the elongated cosmic ray. The red line in the zoomed inset (the solid, white box) denotes the projected path length computed as diagonal of the box.

Table 4
Summary Statistics of the Distribution of Cosmic Ray Properties by Detector

	WFPC2	STIS/CCD	ACS/HRC	ACS/WFC	WFC3/UVIS
Energy Deposited (e-)					
25th%	1380.35	1323.74	1469.49	1141.68	1323.78
50th%	2677.9	2621.14	3024.74	1998.51	2239.04
75th%	16134.39	45958.78	115055.12	778211.98	108073.69
Size Parameter ^a (σ)					
25th%	0.438	0.426	0.421	0.481	0.477
50th%	0.536	0.512	0.493	0.591	0.602
75th%	0.928	2.004	1.156	1.674	6.135
Size ^b (pixels)					
25th%	4	5	5	6	6
50th%	6	7	7	9	9
75th%	10	39	13	20	46
Shape Parameter ^c					
25th%	0.423	0.356	0.326	0.339	0.365
50th%	0.646	0.582	0.523	0.538	0.575
75th%	0.915	1.000	4.183	2.126	1.542

Notes.

^a The size parameter σ is derived using Equation (7) and is equivalent to the Gaussian standard deviation.

^b The number of connected pixels.

^c The shape parameter is calculated using Equation (8).

Within the observatory, each instrument is located in a distinct area of the telescope, surrounded by differing amounts of shielding from space radiation. The amount of shielding around each instrument will determine the minimum energy a

cosmic ray can have and still reach the CCD. The effects of this will be twofold. First, the low-energy regime of the power-law spectrum of cosmic ray energies will be truncated as the low-energy particles are completely absorbed by the shielding.

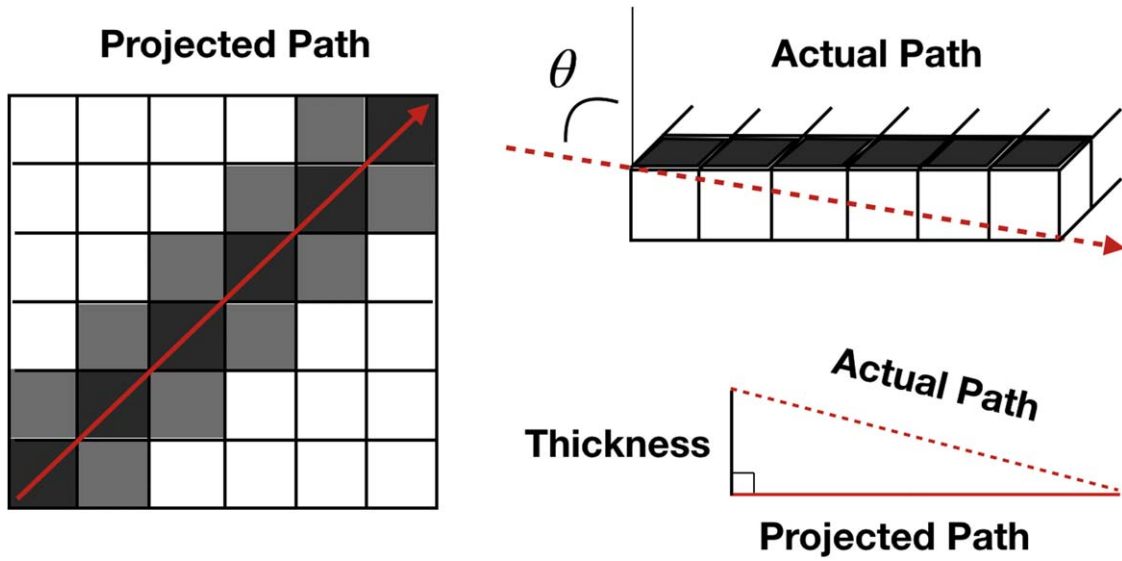


Figure 13. Schematic depiction of the method used to estimate the cosmic ray’s actual path length through the CCD. Dark pixels have higher signal than white pixels; gray pixels are intermediate signal.

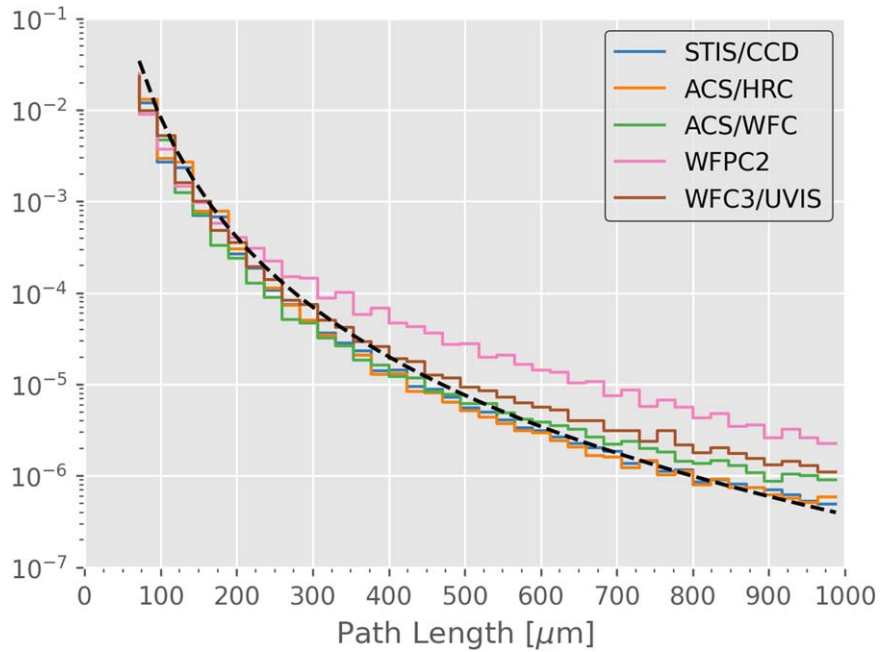


Figure 14. The distribution of path lengths for the five CCD imagers analyzed in this work. The dashed black line is the best fit of the combined distribution of all five imagers.

Table 5

Properties Used to Compute the Path Lengths through the CCD

Instrument	Pixel Size	Average Thickness [μm]
STIS/CCD	21 μm by 21 μm	14.04
ACS/HRC	21 μm by 21 μm	14.26
ACS/WFC	15 μm by 15 μm	14.85
WFPC2	15 μm by 15 μm	10
WFC3/UVIS	15 μm by 15 μm	15.75

Second, the inferred energy of cosmic rays that are detected will be less than their actual energy before encountering the shielding. This underestimation will result in a translation of

the power-law spectrum shown in Figure 15, along the abscissa to lower values of the energy-per-nucleon and along the ordinate to smaller differential intensities.

Additionally, the overall thickness varies from detector to detector, and so a given cosmic ray will deposit more energy in a thicker detector. These differences affect the overall detection rates for cosmic rays. Thus, to perform a direct comparison between instruments, an extensive analysis of the detector characteristics and the shielding for each instrument must be conducted, which is beyond the scope of this paper.

Nevertheless, in Figure 16 we show the distributions of the cosmic ray particle flux for the imagers. We find that the distribution for each instrument has a positive skew and a well-

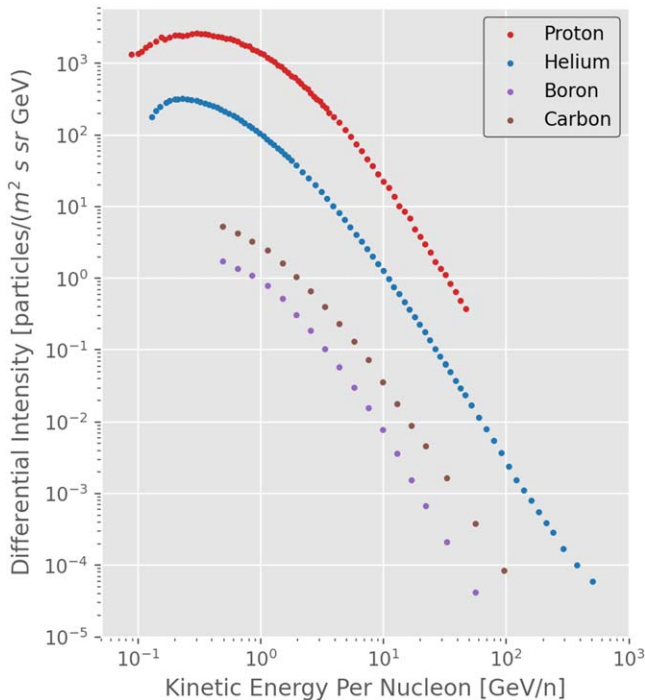


Figure 15. The differential intensity of cosmic rays observed by PAMELA. Data were obtained from the Cosmic Ray Database (Di Felice et al. 2017) hosted by the Space Science Data Center.

defined peak at ~ 1 particle $\text{s}^{-1} \text{cm}^{-2}$. In Table 6, we report summary statistics computed for each distribution.

As a check on our calculated cosmic ray particle fluxes, we compare our results to the flux estimated using the PAMELA data shown in Figure 15. PAMELA was a dedicated cosmic ray detector placed in low Earth orbit, at an average altitude of about 500 km and with an orbital inclination of 70° . While HST and PAMELA share a similar orbital altitude, the higher inclination of PAMELA implies a lower geomagnetic cutoff rigidity (see e.g., Figure 7 in Smart & Shea (2005)). Thus, we should expect the PAMELA cosmic ray particle flux to be slightly larger than that observed by HST. To determine the expected cosmic ray particle flux, in units of particle $\text{s}^{-1} \text{cm}^{-2}$, we make some simplifying assumptions. First, we focus our analysis on the differential intensity of protons, as they comprise the bulk majority of cosmic ray particle types. Second, we assume their distribution to be isotropic. With these assumptions, we estimate the cosmic ray particle flux as

$$F = 2 \int_{E_{\min}}^{E_{\max}} \int_0^{2\pi} \int_0^{\pi/2} I(E) \cos \theta \sin \theta \, dE \, d\phi \, d\theta \quad (10)$$

$$F = 2\pi \int_{E_{\min}}^{E_{\max}} I(E) \, dE. \quad (11)$$

We solve this integral numerically. First, we use cubic spline interpolation to sample the observed distribution at 0.1 GeV resolution. Next, we perform the integration using the sampled data points over the range of kinetic energies in the PAMELA data set, $E_{\min} = 0.1$ GeV to $E_{\max} = 46.6$ GeV.

We find that the flux of cosmic ray protons observed by PAMELA is 2.5 particle $\text{s}^{-1} \text{cm}^{-2}$. If the minimum energy range is restricted to 1 GeV, a more realistic value for HST’s low-inclination orbit, the resulting integral yields 1.3 particle $\text{s}^{-1} \text{cm}^{-2}$. The average of the mean cosmic ray particle fluxes reported in Table 6 is 1.11 particle $\text{s}^{-1} \text{cm}^{-2}$. These values are

Table 6

Summary Statistics Describing the Distribution of Cosmic Ray Particle Fluxes for Each Instrument

	Mean	Error in the Mean	Median	Num. Images
ACS/HRC	1.013	2.55×10^{-3}	0.968	5297
ACS/WFC	1.165	1.92×10^{-3}	1.123	12806
STIS/CCD	0.956	1.21×10^{-3}	0.931	29599
WFC3/UVIS	1.199	1.90×10^{-3}	1.146	12672
WFPC2	1.189	1.77×10^{-3}	1.167	13015

Note. The mean, error in the mean, and median are reported in units of particle $\text{s}^{-1} \text{cm}^{-2}$.

within 17% of one another and demonstrate the capabilities of HST as a particle detector.

4.4. Probing Detector Thickness

When a visible light photon strikes the *Si* detection layer of a CCD, it is readily absorbed within a short distance, due to the extremely high absorption coefficient of *Si*. However, the absorption coefficient is wavelength-dependent, and at longer wavelengths, *Si* becomes increasingly transparent (Ravindra & Narayan 1987). Since CCDs are comprised of layers of different *Si* compounds, each with a different index of refraction, a fraction of the incident light will be transmitted and the remainder reflected at the boundary layer between two compounds. This reflected light can then interfere with any additional light entering, to produce a fringe pattern.

Malumuth et al. (2003), Walsh et al. (2003), and Wong (2010) modeled fringing in STIS, ACS, and WFC3 CCDs, respectively, to correct for the effect of fringing in a given observation. *Si*’s wavelength-dependent absorption distance is well-known, so by modeling the fringing, they derive the *Si* detection layer thickness and produce a corresponding thickness map for each detector (Figure 17, top row). Using the record of pixels affected by cosmic rays, we generate a heat map of cosmic ray strikes for each CCD. This serves as a proxy for the thickness of the detector, as thicker areas of the detector have more potential scattering targets. Comparing thickness maps produced from the fringing analyses to the cosmic ray heat maps, we show in Figure 17 that we reproduce the overall detector structure.

4.5. Solar Modulation of the Cosmic Ray Flux

The periodic nature of solar activity, first observed by counting sunspot numbers, has been extensively monitored since the 1800s and independently confirmed through a variety of other measures of solar activity (e.g., 10.7 cm solar flux, total solar irradiance; see Hathaway 2015). These observations led to the discovery of the 11 yr sunspot and 22 yr magnetic cycles. The 11 yr cycle is defined by the time it takes for the total number of sunspots to progress from a minimum to a maximum and then to the next minimum, with each cycle beginning at a minimum (Russell et al. 2019). The 22 yr cycle, also known as Hale’s Polarity Law, is the time for sunspot pairs to achieve the same magnetic polarity (Hale et al. 1919) with respect to the rotational axis of the Sun. In three consecutive sunspot cycles, the first cycle will have sunspot pairs with a given polarity, the second will have sunspot pairs with the opposite polarity, and the last will have sunspot pairs with the same polarity as the first.

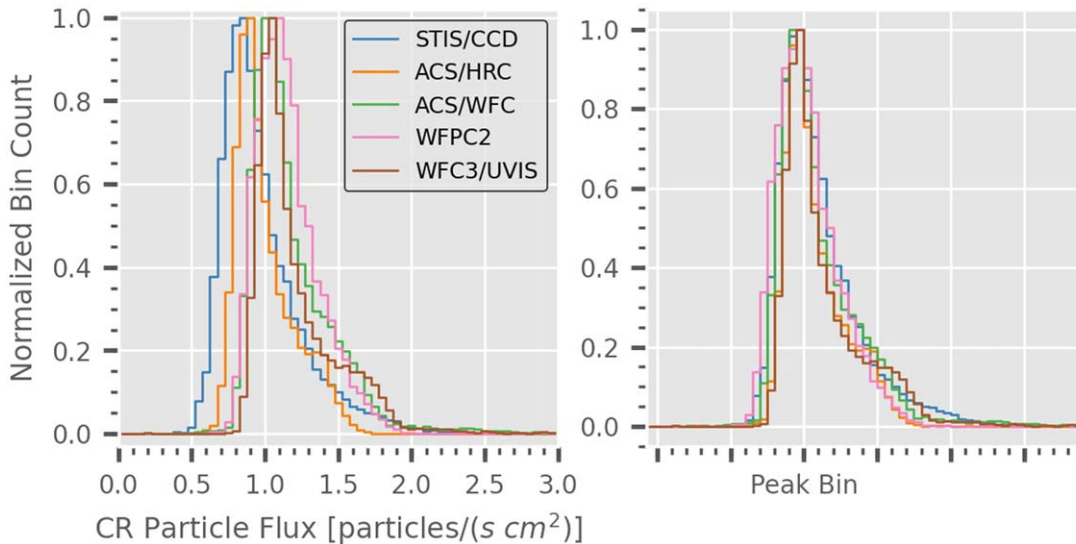


Figure 16. Left: The distribution of observed cosmic ray particle fluxes for each instrument normalized to the peak bin. Right: The same distributions shifted so that their peak bins coincide, to highlight the consistency of the overall distributions.

Table 7
The Extracted Frequencies for Each Instrument in Figure 19

Instrument	Peak 1 (cycles/day)	Period (yr)	Peak 2 (cycles/day)	Period (day)
ACS/HRC	0.001577	1.74	0.02081	48.05
ACS/WFC	0.000231	11.87	0.02089	47.87
STIS/CCD	0.000209	13.10	0.02103	47.55
WFPC2	0.000272	10.09	0.02064	48.46
WFC3/UVIS	0.000237	11.54	0.02098	47.66

It is well known that the observed GCR flux measured at the Earth is anticorrelated with solar activity (Potgieter 2013). In Figure 18, we show the median-normalized cosmic ray particle fluxes measured by HST. We filter out observations in the SAA and smooth the time series with a 30 day rolling median filter.

Qualitatively, the variation of the observed cosmic ray particle flux with the solar cycle is apparent. About one year after each minimum (maximum), the cosmic ray particle flux reaches its minimum (maximum). The delay between the cosmic ray particle flux and solar cycle extrema is expected, as the response of the heliosphere to changes in solar activity is not instantaneous. The observed anticorrelation suggests that the majority of cosmic rays detected by HST are galactic in origin.

Quantitatively, we perform a spectral analysis by computing the Lomb–Scargle periodogram of the observed cosmic ray particle flux as a function of time (see Figure 19). The results are similar for the five instruments, and are given in Table 7, with the exception of the ACS/HRC, which only operated for three years. Its results are therefore unreliable (see Table 1). The first peak occurs at ~ 0.00024 cycles/day, corresponding to the 11 yr solar cycle. The second peak occurs at ~ 0.021 cycles/day, corresponding to 48 days. For each peak, the average false alarm probability is ≈ 0 . In Table 7, we show the peak frequencies and corresponding periods observed by each instrument.

On average, we find that the cosmic ray particle flux varies by about 25% between solar minimum and solar maximum across the detectors, consistent with values from Alpha

Magnetic Spectrometer (AMS). Corti et al. (2019) analyzed seven years of AMS⁹ proton flux data in three rigidity bins: 1.01–1.16 GV, 4.88–5.37 GV, and 33.53–36.12 GV. They found that the lowest rigidity bin has the highest proton flux and the highest variation, $\sim 25\%$ – 30% from the average, with the solar cycle. We thus expect that cosmic rays detected by HST instruments will have kinetic energies consistent with the lowest rigidity of AMS reported in Corti et al. (2019). Using the relationships given in Appendix B, the corresponding kinetic energy range for the lowest rigidity bin is 440 to 553 MeV.

4.6. The South Atlantic Anomaly

Since its discovery, the SAA has been, and will continue to be, an area of great interest as humanity moves to increase its presence in low Earth orbit. The SAA is characterized by an anomalously low value of the geomagnetic field intensity near the Earth’s surface and drifts at a rate of $0^{\circ}.36 \pm 0^{\circ}.06$ W/yr and $0^{\circ}.16 \pm 0^{\circ}.09$ N/yr (Schaefer et al. 2016). As a consequence of the weak intensity, particles trapped in the inner Van Allen radiation belts can more readily penetrate into the Earth’s upper atmosphere (Heynderickx 2002).

HST routinely passes through the SAA (Lupie 2002) whose boundary is defined by avoidance contours. Their coordinates are available via the `costools` Python package¹⁰ maintained by the Cosmic Origins Spectrograph Instrument Branch at the Space Telescope Science Institute. To safeguard against damage to the electronics, astronomical observations are scheduled during SAA-free orbits. Occasionally, an HST calibration program is carried out to map the SAA boundary (Martel et al. 2009; Barker et al. 2010). Here, we show the results of our analysis of HST Proposal 7061 for which dark exposures were made inside and outside the SAA.

Nineteen 60 s darks were taken through the SAA. The position of HST during these observations is shown in Figure 20, and the data set is listed in Appendix A, Table 13. In

⁹ In 2011, AMS was installed on the ISS, which orbits at 450 km above Earth’s surface.

¹⁰ <https://github.com/spacetelescope/costools>

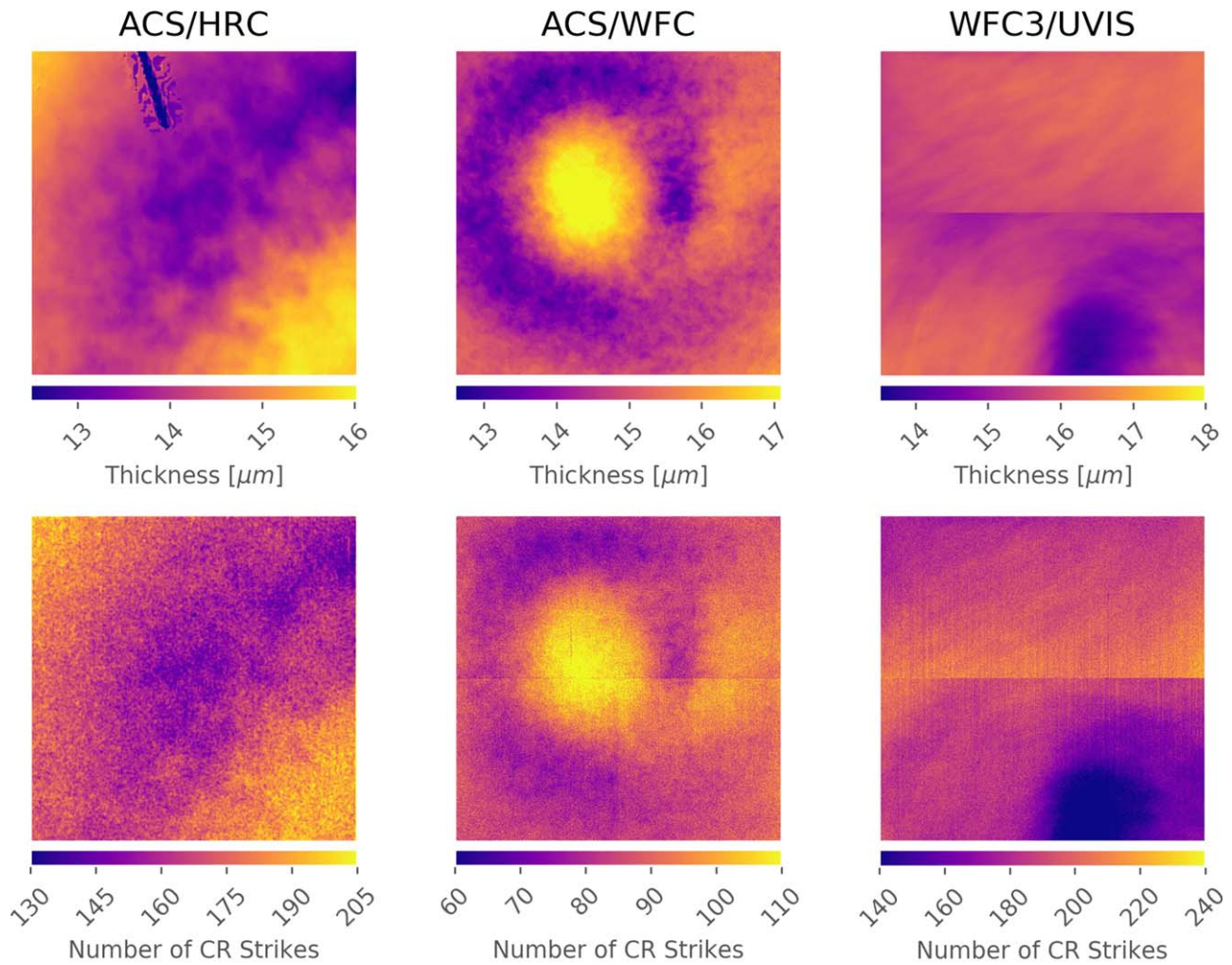


Figure 17. Top row: The thickness maps produced by analyzing fringe patterns in the respective instruments. The ACS data was provided by J.R. Walsh from Walsh et al. (2003), and the WFC3 data was provided by M. Wong from Wong (2010). Bottom row: The corresponding cosmic ray heat maps. Note that, by virtue of the nature of the cosmic ray heat maps, we were able to probe the region of the ACS/HRC detector obscured by the coronagraphic finger in the thickness map generated from the fringing analysis.

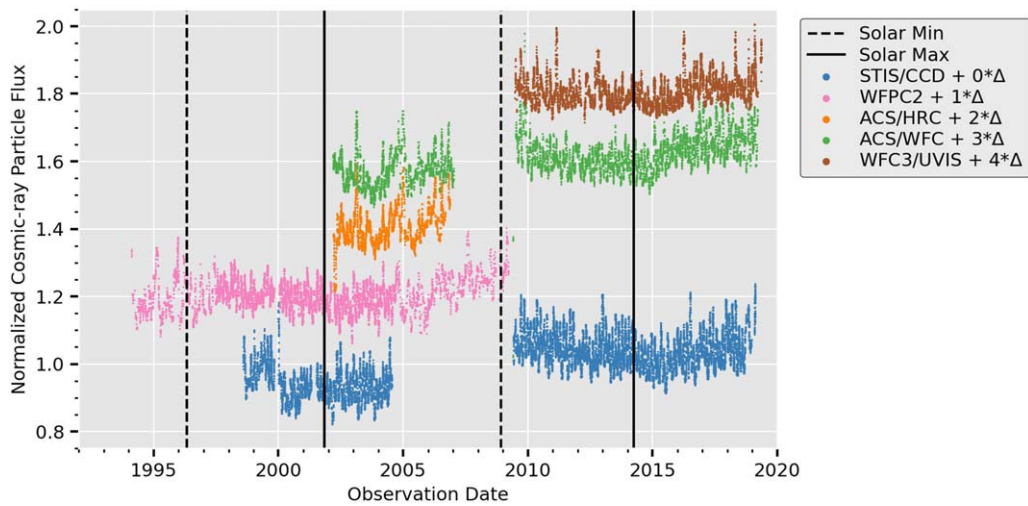


Figure 18. Top: The observed cosmic ray particle flux for each instrument after being smoothed by a one-month (30 day) rolling median and normalized by the median flux observed by the instrument. For clarity, we use a constant offset, $\Delta = 0.2$, to shift the normalized fluxes along the y-axis. The dashed vertical lines correspond to solar minima, whereas the solid lines correspond to solar maxima.

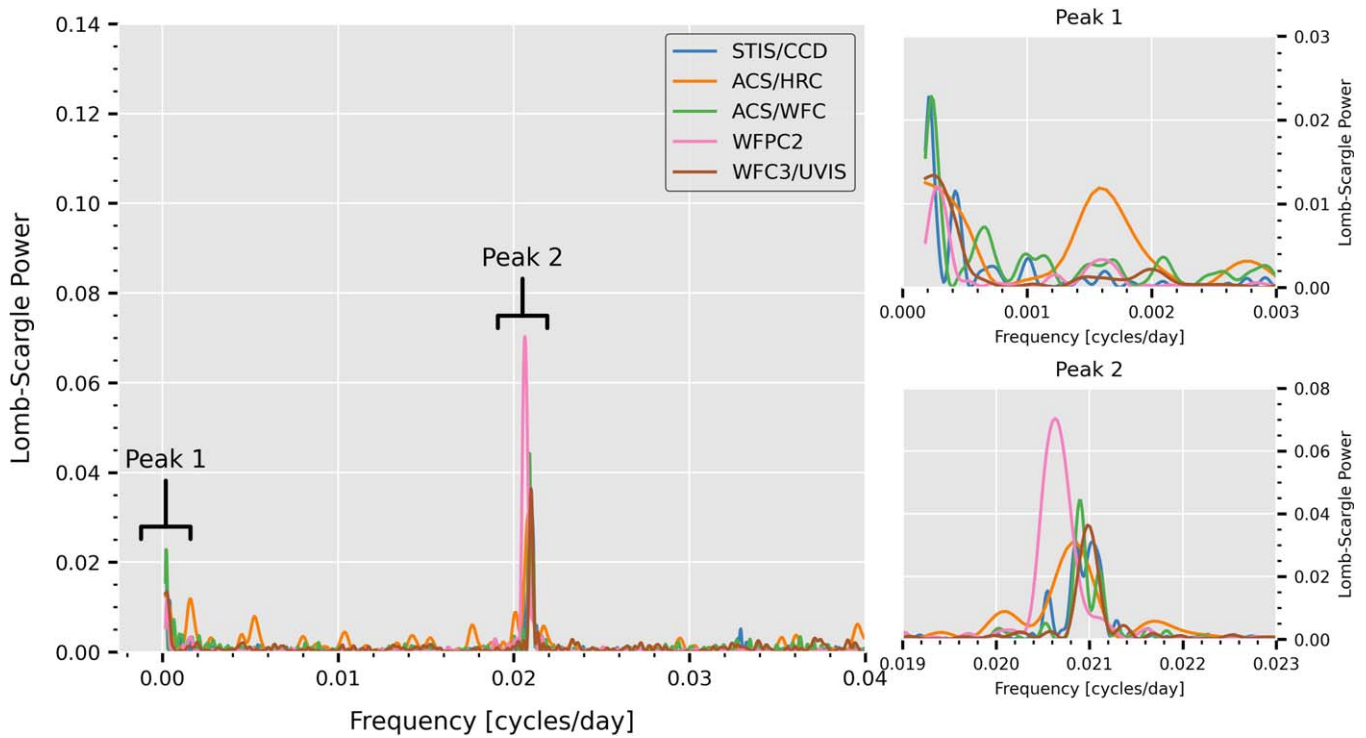


Figure 19. Left: The Lomb–Scargle periodogram for each of the five CCD imagers analyzed. Top Right: Same data, zoomed in to highlight the first peak. Bottom Right: Same data, zoomed in to highlight the second peak.

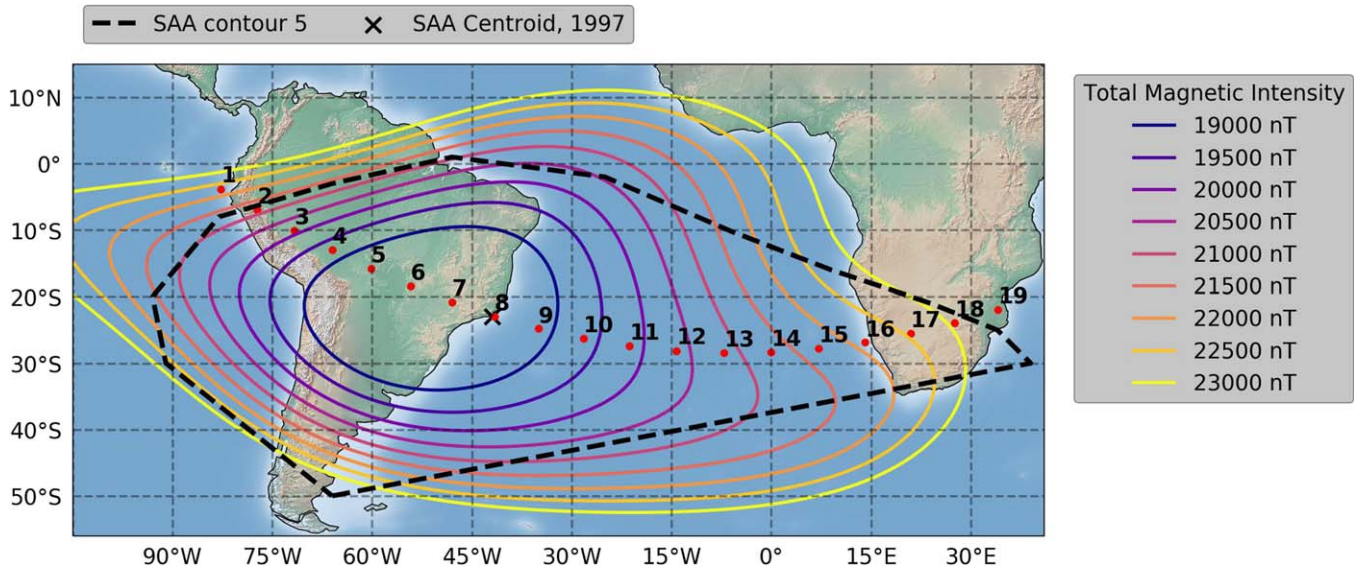


Figure 20. The location of HST during each of the 19 STIS/CCD images taken as part of Proposal 7061. The black, dashed polygon outlines the SAA region defined by SAA avoidance contour 5 (see Lupie 2002). The contour lines correspond to the total magnetic intensity in 1997 at an altitude of 640 km in the vicinity of the SAA computed using the IGRF-13 model using P_{magPy} (Tauxe et al. 2016). Additionally, we show the contemporaneous centroid of the SAA reported by Fürst et al. (2009).

Figure 21, we show 150 by 150 pixel subsections of the darks, to visually demonstrate the level of cosmic ray contamination in the SAA.

For exposures deep in the SAA, e.g., plots 5–10 in Figure 21, it is clear that it is impossible to reliably extract statistics on individual cosmic rays, due to the large number of overlapping cosmic rays. However, the STIS cosmic ray rejection algorithm does a superb job of identifying all the

pixels affected by cosmic rays. Hence, by generating the label from the DQ array, we can compute the total energy deposited by all cosmic rays and estimate the cosmic ray particle flux as

$$\text{Number of Cosmic Rays} = \frac{E_{\text{total}}}{\langle E_{\text{CR}} \rangle * t * A}, \quad (12)$$

where

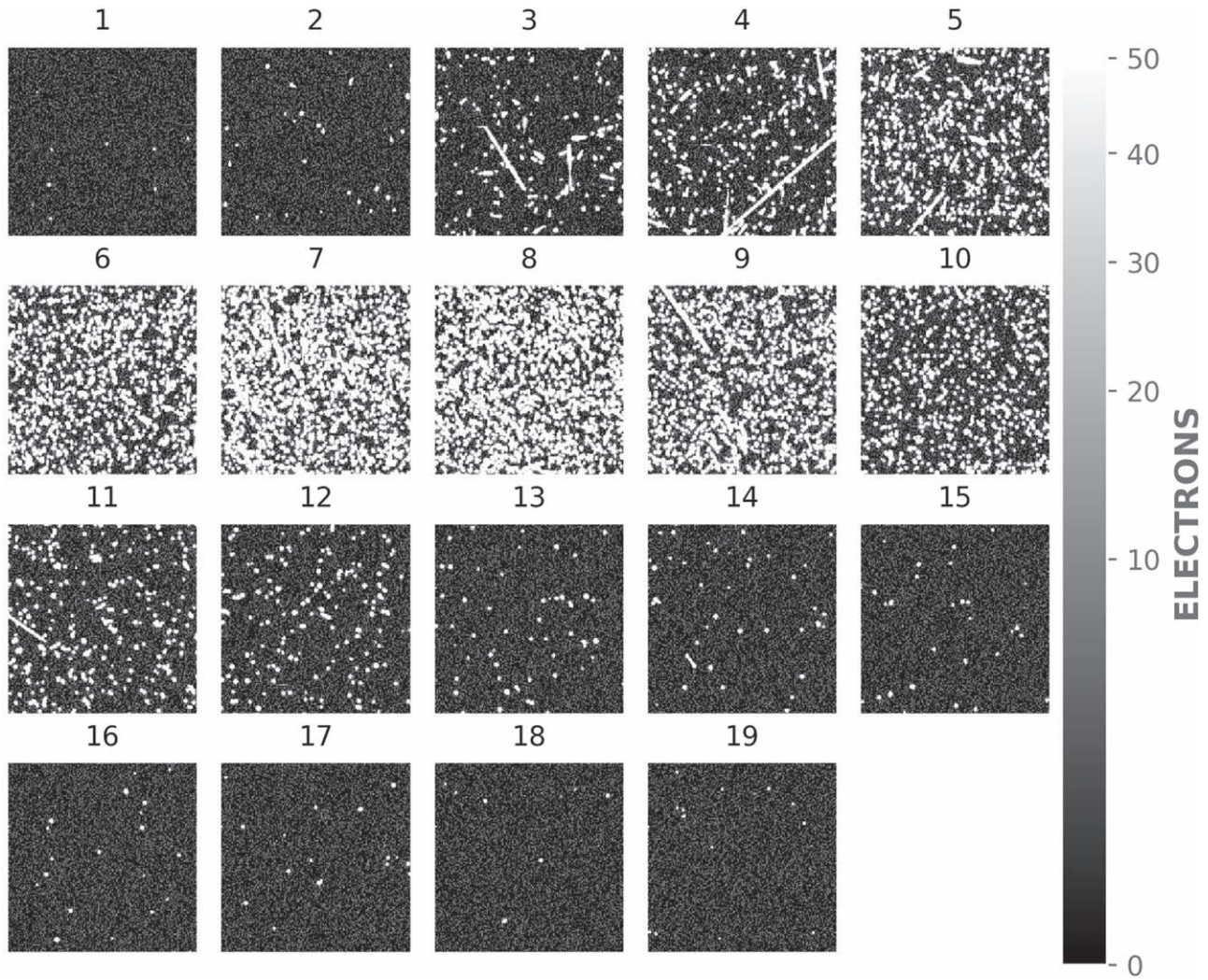


Figure 21. Each subplot shows a 150 by 150 pixel cutout from one of the 19 STIS images. The label at the top of each subplot indicates the image’s number in the observing sequence displayed in Figure 20.

1. E_{total} : = Total energy deposited by all cosmic rays in an image. [e^-].
2. $\langle E_{\text{CR}} \rangle$: = Average energy deposited per image by all cosmic rays. [$e^- \text{ s}^{-1} \text{ cm}^{-2}$].
3. t : = Total integration time for the image. [s].
4. A : = Geometric area of the STIS/CCD (see Table 1). [cm^{-2}].

The value of $\langle E_{\text{CR}} \rangle$ can be determined in two ways, using the data generated by our pipeline. The first method utilizes results presented in Tables 4 and 6. From Table 4, the average total energy deposited by a single cosmic ray is $\sim 2621 e^-$. From Table 6, the average cosmic ray particle flux is $0.96 \text{ particle s}^{-1} \text{ cm}^{-2}$. We compute $\langle E_{\text{CR}} \rangle$ as the product of these two measurements, yielding a value of $2516 e^- \text{ s}^{-1} \text{ cm}^{-2}$.

The second method utilizes the morphological parameters generated by our pipeline. The results for the cosmic rays identified in each image are stored as separate HDF5 data sets within a single HDF5 file. This allows us to derive a value for E_{CR} from each image to generate a distribution. We compute E_{CR} as the sum of the energy deposited by all of the identified cosmic rays, divided by the integration time and detector area. We analyze more than 31000 STIS/CCD observations taken outside the SAA to generate the distribution in Figure 22. The

median value is $3215 e^- \text{ s}^{-1} \text{ cm}^{-2}$, and the most probable bin from the histogram is $2657 e^- \text{ s}^{-1} \text{ cm}^{-2}$. Because of the positive skew, we use the most probable bin for the value of $\langle E_{\text{CR}} \rangle$, as it provides a more accurate estimate than the median.

The values derived using the two different methods agree to within 5%, and we adopt the value generated from the second method as $\langle E_{\text{CR}} \rangle$. We use this value to calculate the cosmic ray particle flux for each SAA image, and the results are listed in Table 13 in the Appendix. In Figure 23, we show the cosmic ray particle flux as a function of the time elapsed since the first image in the observing sequence. The maximum cosmic ray particle flux was $\sim 1100 \text{ particle s}^{-1} \text{ cm}^{-2}$, and it occurred in Observation 8 (o3st20gjq_fit.fits) at a latitude of $22^\circ 96\text{S}$ and a longitude of $41^\circ 61\text{W}$. The minimum cosmic ray particle flux was $\sim 1.4 \text{ particle s}^{-1} \text{ cm}^{-2}$, and it occurred in Observation 1 (o3st20gcq_fit.fits) at a latitude of $3^\circ 81\text{S}$ and a longitude of $82^\circ 76\text{W}$. These observations correspond to the maximum and minimum of the rate of energy deposition by cosmic rays of $3 \times 10^6 e^- \text{ s}^{-1} \text{ cm}^{-2}$ and $3570 e^- \text{ s}^{-1} \text{ cm}^{-2}$, respectively.

We find that the cosmic ray particle flux increases by more than a factor of 800 from the edge to the center of the SAA. When the peak cosmic ray particle flux was observed, HST was nearly collocated with the derived centroid of the SAA in 1997

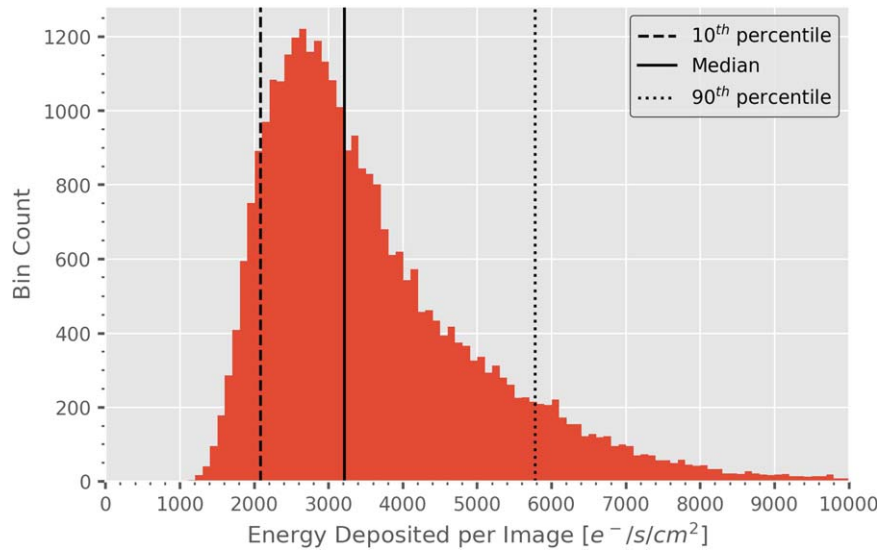


Figure 22. The distribution of the rate of energy deposition by cosmic rays in a single image for the STIS/CCD data set. The dashed line at $2085 \text{ e}^- \text{ s}^{-1} \text{ cm}^{-2}$ marks the 10th percentile, the solid line at $3215 \text{ e}^- \text{ s}^{-1} \text{ cm}^{-2}$ marks the median (50th percentile), and the dotted line at $5778 \text{ e}^- \text{ s}^{-1} \text{ cm}^{-2}$ marks the 90th percentile.

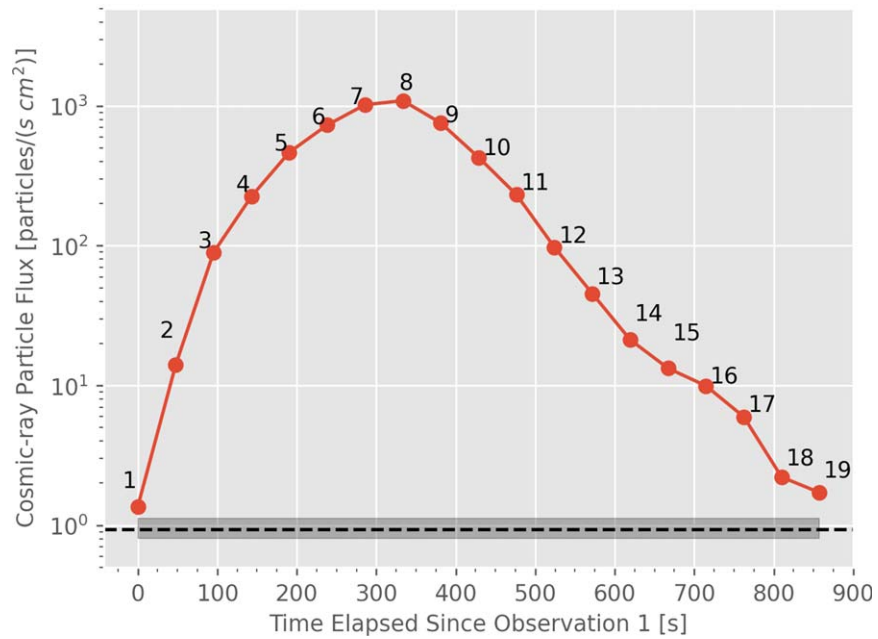


Figure 23. The cosmic ray particle flux as a function of time elapsed since Observation 1 in Figure 20. Each point is labeled to indicate the observation used to make the measurement. The dashed horizontal line is the mean value reported for the STIS/CCD in Table 6. The gray shading denotes the interval bounded by the 25th and 75th percentiles of the cosmic ray particle flux distribution shown in Figure 16.

reported by Fürst et al. (2009) (Figure 20), providing additional constraints on the morphology of the SAA at the time.

4.7. Hot Spots

It is well-known that the trajectories of charged particles are affected by the presence of external electric and magnetic fields. Since cosmic rays are charged particles, we expect the cosmic ray particle flux observed by HST to depend on the geomagnetic field. We compare the cosmic ray particle flux as a function of latitude and longitude to the total magnetic intensity at an altitude of 565 km (HST’s 30 yr average altitude). For this analysis, we use cosmic ray particle flux measurements from images where the integration time is greater than 800 s. In Section 3.4, we defined the integration

time as the exposure time plus half the time required to readout the detector. When the exposure time and readout time are comparable, the fraction of detected cosmic rays that impacted the detector during readout is no longer negligible. This amplifies any error introduced by computing the integration time as the sum of the exposure time and half the readout time. The readout times range from 29 s (STIS/CCD) to 120 s (WFPC2), and are listed in Table 8. Thus, the choice of our 800 s cut is to minimize the error associated with cosmic rays striking during readout. In Figure 24, we show the top 20 most common integration times that are longer than 800 s among the five imagers.

After making this cut, there are 54,215 images distributed between the five CCD imagers, with an average integration time of ~ 1119 s. To compute the total magnetic field intensity,

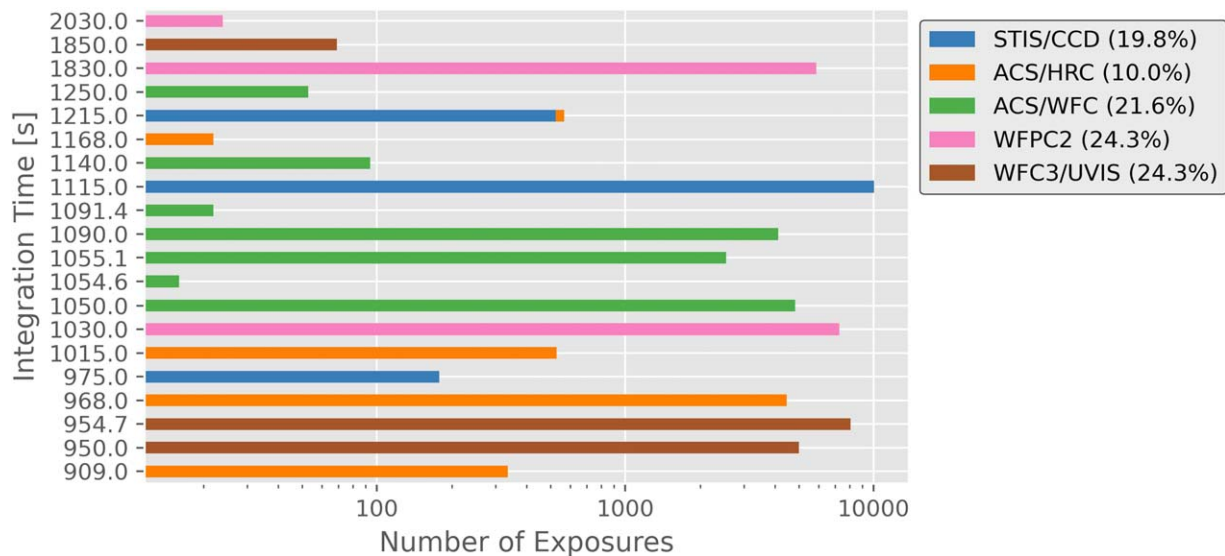


Figure 24. We analyze 54,215 images and compute the top 20 most frequently used integration times among any of the five CCD imagers. Of the total integration time WFC3 and WFC2 each contribute $\sim 24\%$, STIS 19.8%, and ACS (WFC+HRC), $\sim 31\%$.

Table 8
Full Frame CCD Readout Times for Each Instrument

CCD	Readout Time s	Reference
WFC3/UVIS	96	Dressel et al. (2019)
ACS/WFC	100	Mutchler & Sirianni (2005), p. 3
ACS/HRC	26	Mutchler & Sirianni (2005), p. 3
STIS/CCD	29	Riley et al. (2019)
WFC2	$\begin{cases} 60 & \text{If exptime} \leq 180 \text{ s} \\ 120 & \text{If exptime} > 180 \text{ s} \end{cases}$	McMaster (2008)

we use the IGRF-13 model for 2005 (the midpoint of HST’s operational lifetime) and an altitude of 565 km, which approximately corresponds to HST’s average orbital altitude. In Figure 25, we superpose a plot of the distribution of the observed cosmic ray particle flux as a function of HST’s orbital position onto a map of the total magnetic field intensity.

Each point in Figure 25 corresponds to the latitude and longitude at the start of the integration of a single dark frame, and is color-coded (dark purple to bright yellow) according to the observed cosmic ray particle flux. The dashed, black line is an example of HST’s ground track, and the black points mark 250 s intervals along the path. Because of HST’s orbital inclination, the projected latitude coverage is limited to $\pm 28^\circ 5'$. Colored contour lines indicate the magnetic intensity, ranging between 18,000 nT (dark blue) and 51,000 nT (yellow).

Two “hot spots” at the 5σ level are apparent. The northern region extends between 90°W and 150°W and 15° – $28^\circ 5'\text{N}$. The southern area extends from the eastern edge of the SAA to Western Australia, i.e., between 15°E and 120°E , and its north/south range is between 15° and $28^\circ 5'\text{S}$. These regions appear to correspond to locations where the magnetic field intensity is around 36,000 nT. The average integration time of 1119 s corresponds to a track of four black points, so there is some smearing of the actual location and extent of each hot spot, due

the differences in the starting location of each observation. However, the robust sampling of latitude and longitude positions over the 25 yr period minimizes the impacts of this, because there are just as many observations taken while entering each hot spot as there are while exiting.

4.8. Kinetic Energy Estimation

In order to compute an estimate of the average kinetic energy of the observed cosmic rays, we make a series of simplifying assumptions.

First, we assume that the detected cosmic rays are relativistic protons with sufficiently high kinetic energy to penetrate down through the magnetosphere to HST’s orbit: a low altitude of ~ 540 km (see Figure 1) and an orbital inclination of $\pm 28^\circ$. At this orbital inclination, the vertical geomagnetic cutoff rigidity at an altitude of 450 km is $\geq \sim 5$ GV (Figure 7 in Smart & Shea 2005), corresponding to a kinetic energy of approximately 4 GeV or higher (i.e., $\beta \geq 0.98$), for high-energy protons originating outside the magnetosphere.

Second, we ignore shielding losses for arbitrary trajectories through the HST observatory. Hence, the kinetic energy estimated here establishes a *lower limit* for the typical cosmic ray interacting with the CCD imagers. Note that the presence of trapped radiation in low Earth orbit means that this is not a hard limit.

Finally, to determine the energy deposited by cosmic rays, we convert the detected signal (in electrons) to energy (eV). The average operating temperature of the CCDs is -82°C (Dressel et al. 2019; Ryon et al. 2019; Riley et al. 2019). At this temperature, the average energy required to produce an electron-hole pair in silicon is ~ 3.71 eV (Lowe & Sareen 2007). Thus, the energy deposited in eV is $(3.71 \text{ eV/electron}) \times (\text{number of electrons})$. The distribution of energy deposited, equivalent to cosmic ray energy loss, for the five CCD imagers is shown in Figure 26.

4.8.1. Estimating the Energy-loss Probability Density Function

For the subsequent analyses, we use a subset of the data covering the time period between 2001 and 2005 when four

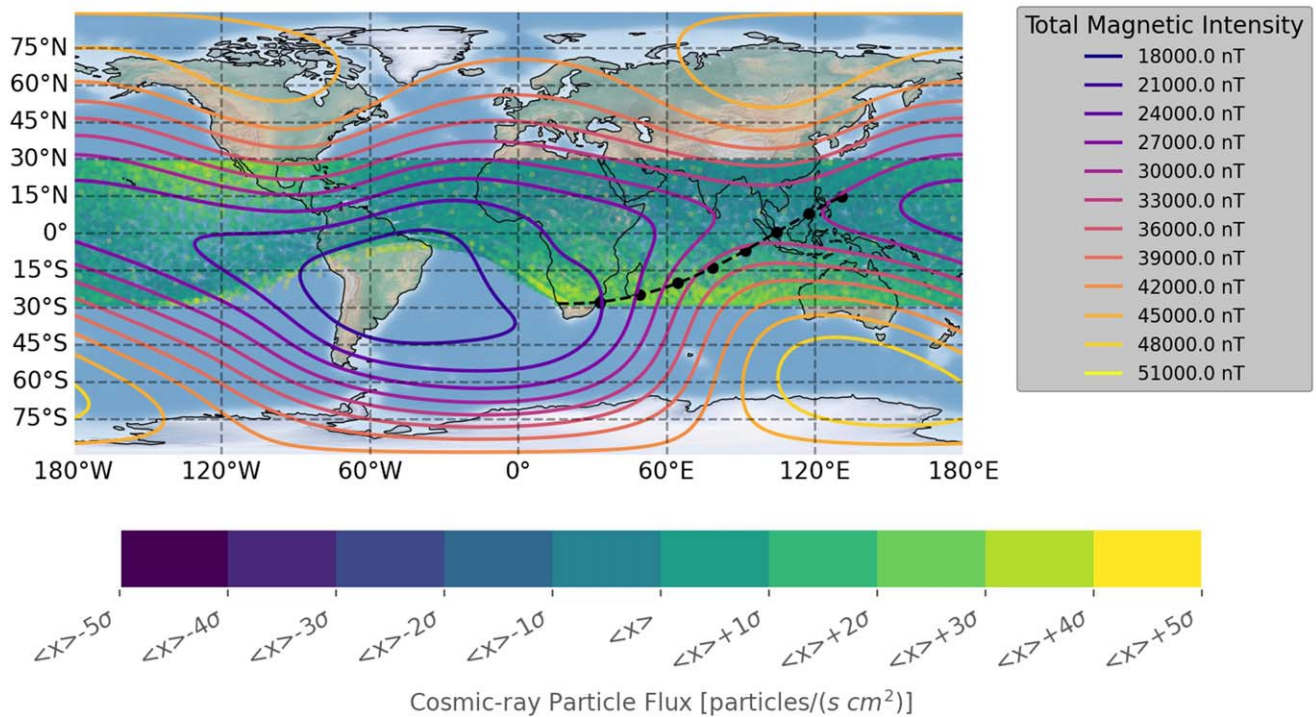


Figure 25. The observed cosmic ray particle flux as a function of orbital position. Each point corresponds to a single observation where the integration was longer than 800 s. The color mapping corresponds to the observed cosmic ray particle flux where the mean and standard deviation, $\langle x \rangle$ and σ , respectively, are sigma-clipped values computed using all 54,215 observations. The dashed, black line is an example of HST’s ground track over a 2000 s exposure. The black points mark 250 s intervals. We assume an average date and altitude of 2005 and 565 km, respectively, for computing the total magnetic intensity using the IGRF-13 model.

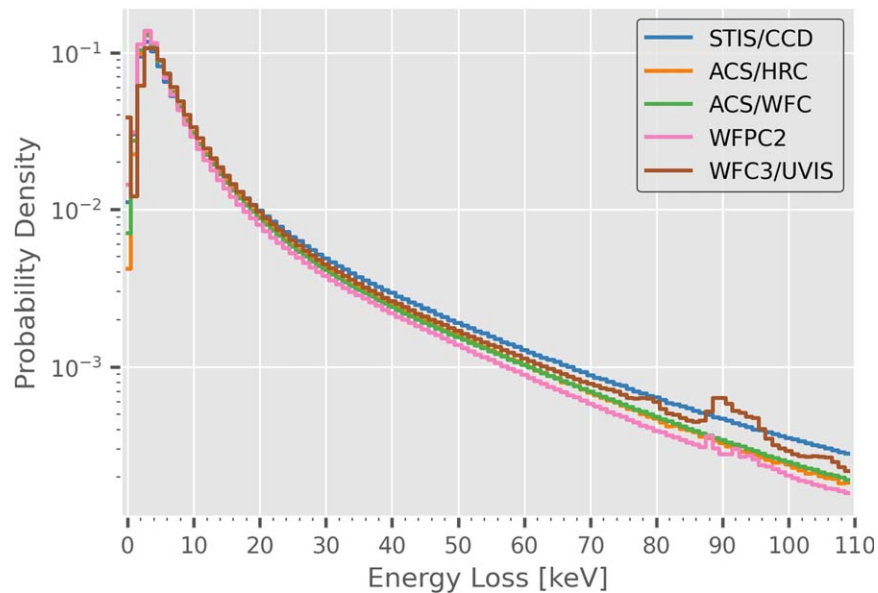


Figure 26. The observed energy loss distributions for each instrument. For energy losses below ~ 20 keV, the distributions are very similar. After ~ 20 keV, the distributions begin to diverge from one another, due to differences in shielding and instrument location within the observatory.

CCD imagers operated simultaneously on HST: ACS/HRC, ACS/WFC, STIS/CCD, and WFPC2 (Table 1). Further, we restrict our analysis to cosmic rays with path lengths between 280 and 300 μm , primarily because this allows us to make a direct comparison to previous laboratory results. Bichsel (1988) report results for protons with relativistic kinetic energies of 1.2 GeV and 7.1 GeV (corresponding to $\beta\gamma$ equal to 2.1 and 8.5, respectively) passing through 290 μm of silicon. The data

presented in Table 9 of Bichsel (1988) are a reproduction of the results obtained by Bak et al. (1987).

We calculate the probability density function of the energy-loss distribution using the statsmodels (Seabold & Perktold 2010) Python package to perform the kernel density estimation, and estimate an optimal bandwidth via a maximum-likelihood cross-validation technique. Figure 27 shows a histogram of the energy-loss distribution with bin size of

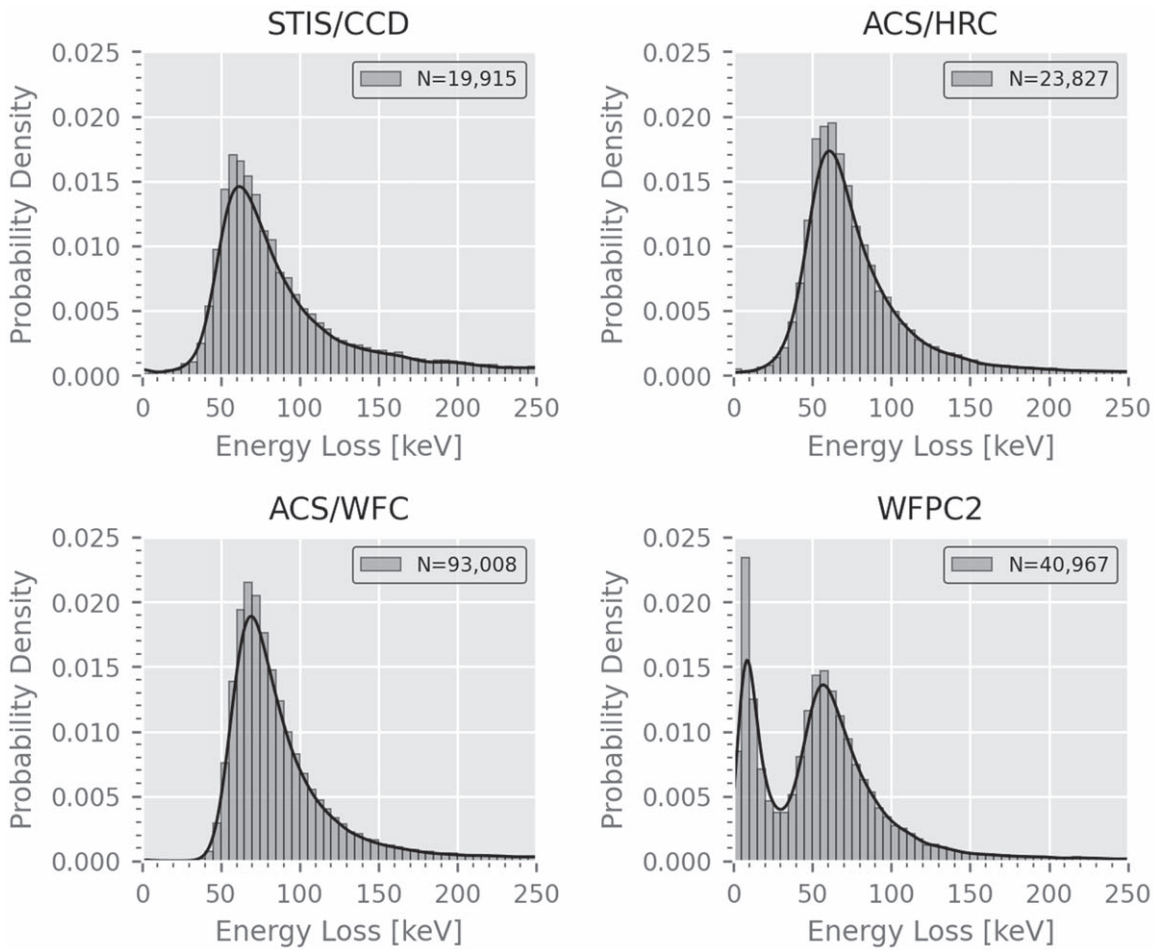


Figure 27. The energy-loss distributions for the four imagers. The solid, black line is the KDE-derived probability density function. The gray shading is a histogram of the data set generated using a bin size of 5 keV. Each histogram has been normalized to yield a probability density function.

5 keV (gray rectangles) for track lengths between 280 and 300 μm , and the corresponding KDE-derived probability density function (solid, black line) for each instrument. The thinned, backside-illuminated CCDs (ACS/HRC, ACS/WFC, and STIS/CCD) all have a single peak at ~ 60 keV. The thick, frontside-illuminated CCD (WFPC2) has two clear peaks. The first is around ~ 5 keV and the second is around ~ 55 keV.

The bimodal energy-loss distribution for WFPC2 is a consequence of its CCD structure. Front-illuminated CCDs that are not thinned can have thick, up to 500 μm , bulk silicon substrates even though the thickness of the photosensitive (epitaxial) layer is comparable to those of thinned, back-illuminated detectors at ~ 10 – 20 μm . Cosmic rays deposit energy throughout their path through the detector. Electrons generated in or near the depletion region are more likely to be collected, whereas those generated in the substrate and field-free regions are more likely to be lost to recombination (partial events). Additionally, split events can occur when charge diffuses from one pixel to another. Thick CCDs have a higher fraction of partial and split events compared to thinned CCDs, resulting in a lower charge collection efficiency (Janesick 2001).

DQ labeling and threshold labeling (Section 3.2) minimize the effects of split events by identifying all pixels directly affected by each cosmic ray, as well as their neighbors. The

calibration of partial events is much more difficult and requires knowledge of the depth of the interaction. Due to its bimodal distribution, we omit WFPC2 from the remainder of the kinetic energy estimation analysis.

4.8.2. The Landau Distribution

Ionization due to inelastic collisions between an incoming particle and the electrons in the silicon substrate is the principal energy-loss mechanism in solid-state detectors. For relativistic charged particles interacting with matter, the maximum possible energy transfer in a single collision between a particle of mass M and an electron is given by Tanabashi et al. (2018):

$$W_{\text{max}} = \frac{2m_e c^2 \beta^2 \gamma^2}{1 + 2(\gamma m_e/M) + (m_e/M)^2}, \quad (13)$$

where m_e is the mass of an electron, $\beta = v/c$, and $\gamma = \sqrt{1/(1 - \beta^2)}$. The energy loss of the incident particle is subject to random fluctuations arising from the probabilistic nature of the collisions.

For the case when the energy loss is small compared to the particle’s kinetic energy (i.e., thin detectors) and the electron binding energy is negligible, the energy-loss distribution is described by the Landau distribution (Landau 1944), where the

Table 9
The Landau Parameters Computed in this Work and Bichsel (1988)

Path Length (μm)	β	ξ (keV)	W_{max} (keV)	ξ/W_{max}	t/β^2 (μm)	Source
290	0.903	6.339	4.496×10^3	1.42×10^{-3}	356	Bichsel (1988)
290	0.982	5.360	2.746×10^4	1.95×10^{-4}	301	This work
290	0.993	5.242	7.316×10^4	7.17×10^{-5}	294	Bichsel (1988)

peak value corresponds to the most probable energy loss:

$$L\Delta_p = \xi \left[\ln \frac{2m_e c^2 \beta^2 \gamma^2}{I} + \ln \frac{\xi}{I} + 0.200 - \beta^2 - \delta(\beta\gamma) \right], \quad (14)$$

where I is the mean excitation energy of the material, and ξ is the Landau parameter describing the typical energy loss for the material. Restating the conditions under which the Landau distribution describes energy loss:

1. $k = \xi/W_{\text{max}} \ll 1$ and
2. $\xi \gg E_{\text{bind}}$, where E_{bind} is the electron binding energy.

When the electronic binding energy is no longer negligible, the distribution is more adequately described by the convolution of the Landau distribution with a Gaussian (Bak et al. 1987; Bichsel 1988; Meroli et al. 2011b)

For silicon, the value of ξ is

$$\xi = 0.017825 * t/\beta^2 \text{ keV}/\mu\text{m}, \quad (15)$$

and t is the thickness of the detector (Bak et al. 1987; Bichsel 1988). The FWHM of the distribution is approximately $w_L = 4.016\xi$. Silicon's maximum electronic binding energy occurs in the K-shell, $E_K = 1.84$ keV. By substituting E_K for E_{bind} for the condition $\xi \gg E_K$, we obtain

$$t/\beta^2 \gg 103 \mu\text{m}. \quad (16)$$

We compute W_{max} (Equation (13)) and the Landau parameter, ξ , (Equation (15)) for the HST data assuming cosmic rays are protons with kinetic energies ≥ 4 GeV, ($\beta = 0.982$); for the Bichsel results, see Table 9. From these quantities, we see that the first condition for the Landau distribution, $k = \xi/W_{\text{max}} \approx 1 \times 10^{-4} \ll 1$, is satisfied, but the second, $t/\beta^2 \gg 103 \mu\text{m}$, is not, indicating that the energy-loss distribution is best described by the convolution of a Gaussian with the Landau distribution.

Using the Python implementation of `Minuit2` (Hatlo et al. 2005), `iminuit` (Dembinski et al. 2020), we fit a Landau convolved with a Gaussian to the kernel density estimation of the underlying probability distribution for the energy losses. In each fit, the free parameters are: the most probable energy loss of the convolved distribution (Δ_p), the width of the Landau distribution (ξ), the width of the Gaussian distribution (σ_{Gauss}), and the height of the convolved distribution (A). We compute the rms error as

$$\sigma_{\text{RMS}} = \sqrt{\langle (y_{\text{fit}} - y_{\text{kde}})^2 \rangle}, \quad (17)$$

and the normalized rms error using the interquartile range (IQR),

$$\sigma_{\text{NRMS}} = \sqrt{\left\langle \left(\frac{y_{\text{fit}} - y_{\text{kde}}}{\text{IQR}_{\text{kde}}} \right)^2 \right\rangle}. \quad (18)$$

Figure 28 shows the best-fit distributions and their fractional residuals. In Table 10, we report the best-fit parameters, the 68% confidence intervals, and the normalized rms error. When the energy losses are small (< 20 keV) the fits are poor. However, the overall shape and peak of the distribution are fit well.

Figure 29 shows the best-fit distributions and their corresponding 3σ error bands generated via MINOS. The peak of the ACS/WFC distribution is offset from the ACS/HRC and STIS/CCD distributions by ~ 8 keV. After losses exceed ~ 150 keV, the ACS/HRC and ACS/WFC distributions converge to within $\sim 8\%$ of one another. In the same regime, the STIS/CCD distribution is systematically higher than both ACS/HRC and ACS/WFC by an average of $\sim 50\%$. We measure the FWHM, w , of each distribution and find widths of 40.27 keV, 37.05 keV, and 44.27 keV for ACS/HRC, ACS/WFC, and STIS/CCD respectively.

The best fit is for ACS/WFC, where the normalized rms error is 3.33%; the worst fit is for STIS/CCD, where the normalized rms error is 11.51%. The poor quality of the STIS/CCD fit is driven by the lack of thermal control of the CCD. In 2001 May, the STIS Side-1 electronics box failed and the instrument began operation using the Side-2 electronics, which cannot precisely control the operating temperature of the CCD. The minimum energy required to produce an electron-hole pair is temperature-dependent; the change in energy per degree is $-0.013\%/K$ (Lowe & Sareen 2007). At warmer temperatures, the same cosmic ray generates more electrons than at colder temperatures.

In Table 11, we list our derived values of Δ_p and w , and for comparison, the values reported by Bichsel (1988) for protons with kinetic energies of 1.2 and 7.1 GeV traversing 290 μm of silicon. The most probable energy losses calculated for the HST CCDs are lower by $\sim 20\%$ and the widths are larger by 36% compared to Bichsel (1988).

The differences are due to the particle populations under investigation. Primary cosmic rays are a mix of 89% protons, 9% helium nuclei, 1% heavy nuclei, and 1% electrons, and have a broad energy distribution with a peak of approximately 0.3 GeV. The values from Bichsel (1988) are for data taken in a laboratory setting for protons with fixed kinetic energies.

We do not expect excellent agreement between our results and those of Bichsel (1988). Rather, the goal of the comparison is to ensure our results are not orders of magnitude off from theoretical models and their high-precision experimental tests.

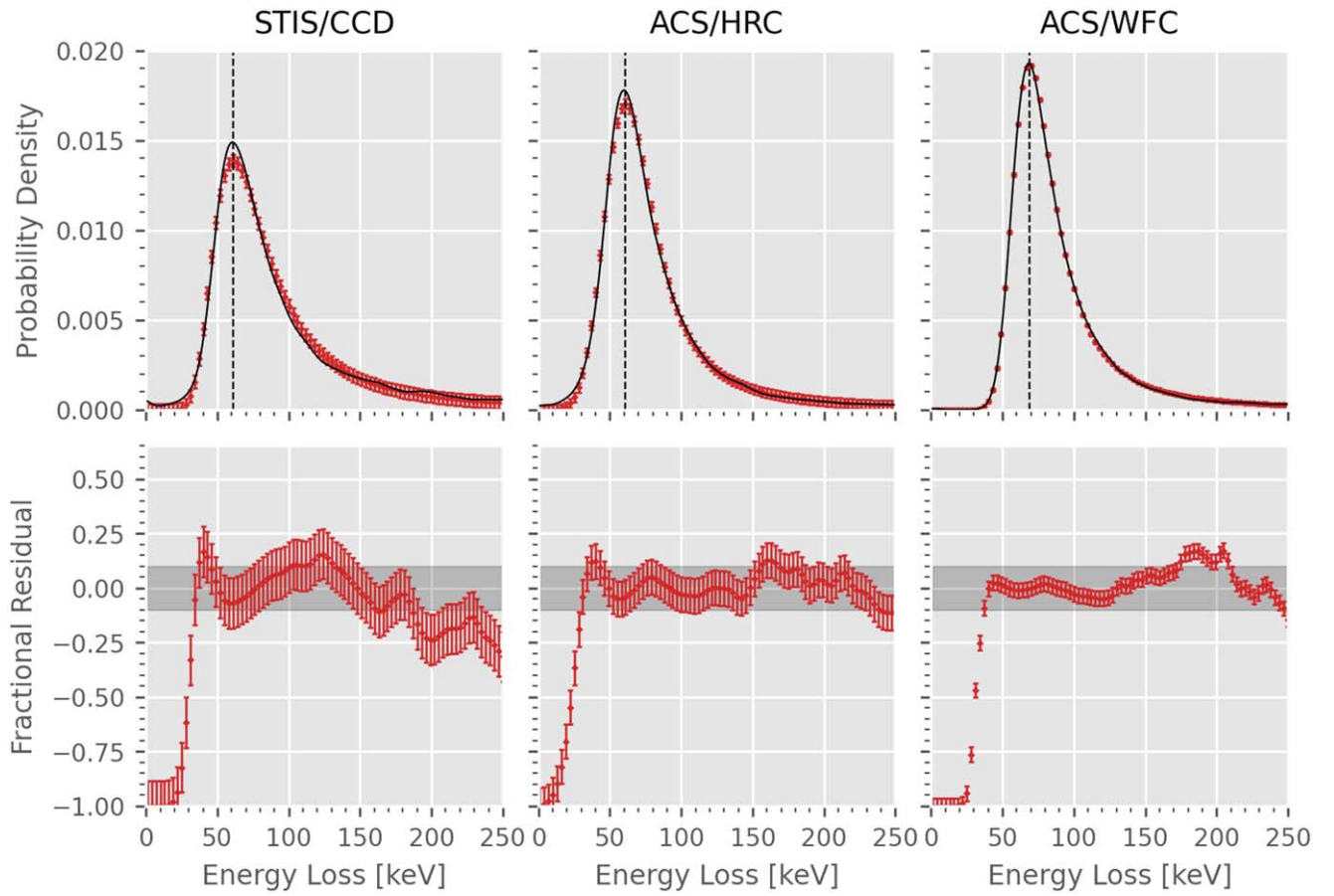


Figure 28. Top Row: The best-fit distribution is shown with red data points, and the error bars correspond to $1 * \sigma_{\text{RMS}}$. The solid, black line is the KDE-derived probability density function. Bottom Row: The fractional residuals for each fit computed as $(y_{\text{fit}} - y_{\text{kde}})/y_{\text{kde}}$. The error bars correspond to $1 * \sigma_{\text{NRMS}}$ and the grading shading indicates $\pm 10\%$.

Table 10
Best-fit Parameters for the Three CCD Imagers Analyzed

	Δ_p (keV)	ξ (keV)	σ_{Gauss} (keV)	$\sigma_{\text{NRMS}}[\%]$
ACS/HRC	60.86, (60.72, 61.00)	7.98, (7.92, 8.04)	10.35, (10.17, 10.52)	7.97
ACS/WFC	68.89, (68.81, 68.98)	7.91, (7.86, 7.95)	6.59, (6.50, 6.68)	3.33
STIS/CCD	61.11, (60.94, 61.28)	12.13, (12.02, 12.23)	4.11, (3.71, 4.48)	11.51

Note. For each parameter, the 68% confidence interval generated using the MINOS algorithm (James & Roos 1975) is given in parentheses.

4.8.3. Computing the Kinetic Energy

Using the relationship between ξ , β , and t defined by Equation (15), we use the value of ξ obtained from the fit and the path length ($280 \mu\text{m} \leq t \leq 300 \mu\text{m}$) to solve for β ,

$$\beta(t) = \sqrt{\frac{0.017825 * t}{\xi_{\text{fit}}}}, \quad (19)$$

and compute the relativistic kinetic energy, $T = (\gamma - 1) * E_o$, where E_o is the rest energy of the proton (Table 12).

The average kinetic energy of particles leaving tracks between 280 and 300 μm is 534 ± 117 MeV. For comparison, the minimum energy for a proton to be stopped in 10 μm of silicon is 5 MeV, indicating that the vast majority of these cosmic rays pass right through the CCDs analyzed.

4.8.4. Discussion

In the preceding section, we determined the relativistic speed, $\beta \sim 0.8$, of particles leaving $\sim 290 \mu\text{m}$ tracks, and thus their average kinetic energy, ~ 0.5 GeV. Particle kinetic energy versus the most probable energy loss is plotted in Figure 30 for a proton traversing 290 μm of silicon. Shown in the figure are the Landau prediction (Equation (14)), the results from Bichsel (1988), and the results from this work. The energy losses for particles with kinetic energy larger than approximately 1.7 GeV are degenerate, i.e., two particles with distinct kinetic energies can have identical values for the most probable energy loss.

At 7.1 GeV, the difference between the Landau prediction and Bichsel (1988) is 5%, but at 1.2 GeV, the prediction is almost 20% larger than the measured value. However, for our estimated average kinetic energy of ~ 0.5 GeV, the most probable energy loss predicted by the Landau distribution is ~ 2

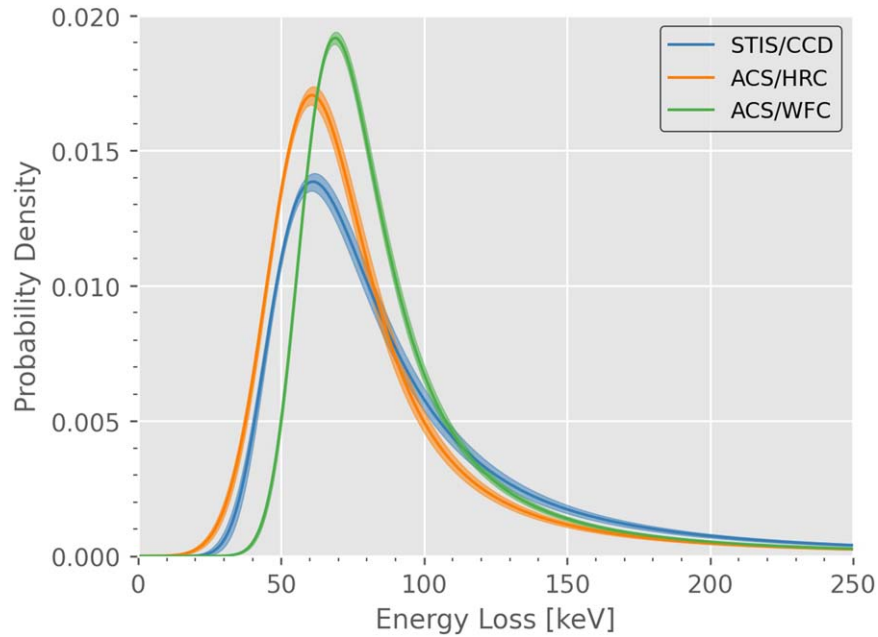


Figure 29. The best-fit straggling distributions with 3σ error bands overlaid.

Table 11

The Most Probable Energy Losses and FWHMs Estimated from this Work (ACS/HRC, ACS/WFC, STIS/CCD) and Those Reported in Bichsel (1988) for 1.2 GeV and 7.1 GeV Protons Traversing $290 \mu\text{m}$ of Silicon

	ACS/HRC	ACS/WFC	STIS/CCD	1.2 GeV	7.1 GeV
Δ_p (keV)	60.86	68.89	61.11	82.46	76.91
w (keV)	40.27	37.05	44.27	31.01	28.53

Table 12

Kinetic Energy of Cosmic Rays with Path Length of $\sim 290 \mu\text{m}$

Instrument	$\beta_{280 \mu\text{m}}$	$T(\beta_{280 \mu\text{m}})$ (MeV)	$\beta_{300 \mu\text{m}}$	$T(\beta_{300 \mu\text{m}})$ (MeV)
ACS/HRC	0.791	594.23	0.818	694.59
ACS/WFC	0.794	606.44	0.822	710.45
STIS/CCD	0.641	284.48	0.664	316.17

times larger. Our measured most probable values are 20% lower than the 1.2 GeV lab measurements. There are two dominant factors that contribute to this apparent discrepancy.

First, HST is exposed to cosmic rays with a broad range of kinetic energies. The population of particles comprising the cosmic rays that impact HST includes low-energy trapped radiation, high-energy solar-energetic particles, and high-energy GCRs. From Equations (14) and (15), we see that the most probable energy is proportional to $1/\beta^2$. Thus, cosmic rays with lower kinetic energies will deposit more energy compared to cosmic rays with higher kinetic energies. The inability of HST to determine the incident kinetic energy of each detected cosmic ray means that each particle is treated with equal weight in our determination of the most probable energy loss. This results in a broadening of the energy-loss distribution (larger ξ), due to contributions from lower-energy particles that lead to a smaller estimation for β through Equation (15).

Second, the CCDs on HST do not have charge collection efficiencies of 100%. This means that some of the electrons generated as a result of the interaction with the cosmic ray are

lost. While this effect is more pronounced in the thick, frontside-illuminated CCDs, it is present in both types (Janesick 2001). This effect is dependent on the depth of the interaction within the pixel; the farther from the gate, the lower the CCE. Janesick (2001) and Meroli et al. (2011a) examined the CCE of CCD and CMOS detectors, respectively, finding typical values of ~ 0.8 . The imperfect CCE manifests as a systematic underestimation of the most probable energy loss. While the effects of an imperfect CCE can be minimized by analyzing specific pixels along the particle track (Meroli et al. 2011a), the inability to determine the incident kinetic energy of each individual cosmic ray cannot be overcome.

The net results of these two factors are the apparent low kinetic energy of the incident cosmic ray and the small value of the most probable energy loss. A full treatment of these effects is beyond the scope of this paper and reserved for future analysis. To further evaluate the validity of our approach, we compare the estimation obtained here with additional works.

Didkovsky et al. (2006) (hereafter D06) estimated the range of kinetic energy of incident particles by modeling the proton energy losses deposited in single-pixel cosmic ray events in the SOHO/EIT CCD. Modeled energies range from ~ 30 MeV to ~ 500 MeV. Energies are converted to differential proton fluxes and compared against data from the Geostationary Operational Environmental Satellite (GOES) Energetic Particle Sensor (EPS). For the two solar energetic particle (SEP) events D06 examined, they find good correlation between their derived differential proton flux and that of GOES. This method works well for the single-pixel events examined by D06 because the constrained angle of incidence (Figure 6) allows isolating distinct energy ranges corresponding to specific CCD signal ranges.

Shen & Qin (2016) (hereafter SQ16) analyzed multi- and single-pixel CR events. They compared the extracted CR count rates to GOES data by analyzing measurements obtained from 23 different solar proton events (SPEs). They found the best correlation to be with the P6 channel on GOES 11, which probes energies between 80 and 165 MeV. They performed a

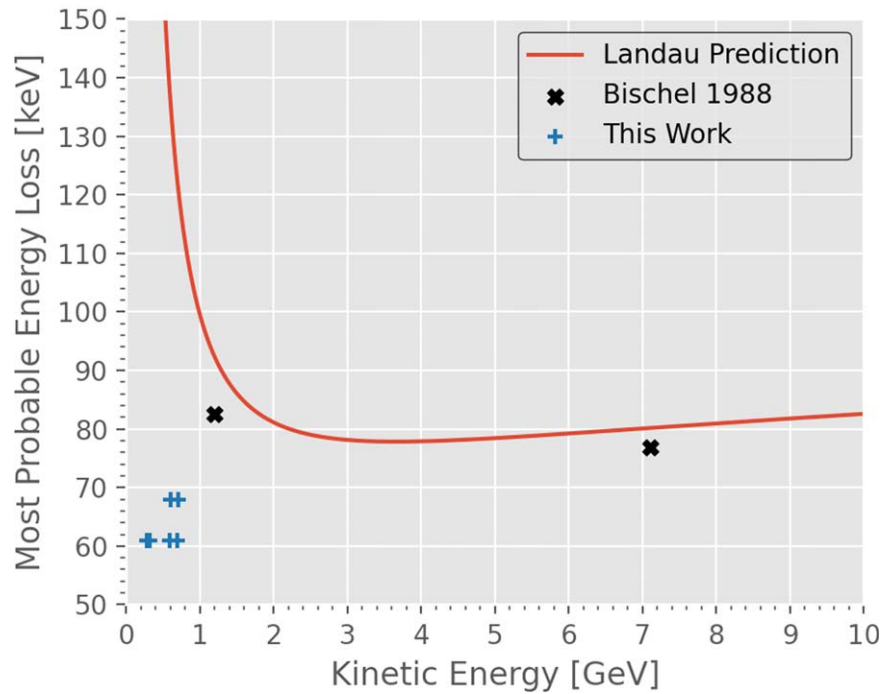


Figure 30. The most probable energy loss for a proton traversing $290 \mu\text{m}$ of silicon. The black “x” denotes the values presented in Table 9 of Bischel (1988). The blue “+” denotes the values obtained in this work. The solid, red line is the value of the peak of the Landau distribution computed using Equation (14), assuming protons and a silicon target with a thickness of $290 \mu\text{m}$.

similar correlation analysis with 39 events that were also observed by the SOHO Energetic and Relativistic Nuclei and Electron (ERNE) experiment. The highest correlation was with SOHO/ERNE channel 3, which probes energies between 118 and 140 MeV. By removing short-term fluctuations from SPEs, they analyzed the slowly varying, low-level background due to GCRs. To convert their GCR count rates to physical units, they utilize similarly smoothed SOHO/ERNE data to derive a scaling factor between the two data sets (Figure 8 in SQ16).

While both these methods provide estimations of the differential proton flux, the differences between the SOHO and HST mission designs reduces their applicability to our work. As a solar observatory, the telescope pointing is always fixed on the Sun. The SOHO/EIT CCD is at the rear end of the spacecraft and it is centered on the optical axis (see Figure 2 in Delaboudinière et al. 1995). The CCD is oriented such that the surface normal is pointing radially toward the Sun. Hence, the CCD surface is perpendicular to the direction of the particle flux during an SEP event, which maximizes the number of single-pixel events during an SEP event. Additionally, HST does not have an in situ particle detector capable of accurately identifying particle types and energies, which complicates the derivation of an accurate scaling factor. Next, HST’s pointing is not fixed on a single target. Instead, it changes constantly as it observes astrophysical sources, and when HST is on Earth’s sunlit side, the pointing is unconstrained. Finally, HST is orbiting in low Earth orbit inside the Earth’s magnetosphere, whereas SOHO is in a halo orbit at L1 outside the Earth’s magnetosphere.

Therefore, we instead compare our results with those derived using observations from the AMS experiment on board the ISS. The ISS has an orbital altitude similar to that of HST, 450 km versus 540 km, though it has a slightly steeper orbital inclination of 51° versus 28° . Both satellites should detect a similar population of cosmic rays. In Section 4.5, we found that

the peak-to-peak variation in the observed cosmic ray particle flux from this work was most consistent with the variation seen in the lowest rigidity bin, 1.01–1.16 GV, analyzed in Corti et al. (2019). Thus, we expect a rigidity derived in this work to be in approximately the same range.

Using the derived estimate of the average kinetic energy for cosmic rays impacting HST of $534 \pm 134 \text{ MeV}$ and the relationship in Appendix B.1, we compute a momentum of $1134.43 \text{ MeV } c^{-1} \approx 1.1 \text{ GeV } c^{-1}$. In these units, the rigidity is 1.1 GV, which agrees with our expectation and confirms the validity of our approach to estimating the kinetic energy of the average cosmic ray.

5. Summary

We developed a Python package, `HSTCosmicrays`, and used it to characterize transient, high-energy particles detected in dark frames that were taken with five CCD imagers on HST; ACS/HRC, ACS/WFC, STIS/CCD, WFPC2, and WFC3/UVIS. Whenever possible, the software will perform a connected-component labeling analysis on the DQ array to identify groups of pixels marked with 8192 (i.e., cosmic ray). When that is not possible, the software will use a technique we refer to as threshold labeling to identify cosmic rays while rejecting hot pixels. For every observation, we record a variety of image metadata and store morphological parameters for each cosmic ray identified.

In total, we have characterized approximately 1.2 billion cosmic rays. We modeled the energy-loss distributions for three of the five CCD imagers and estimated that the typical particle observed by HST has a kinetic energy of $534 \pm 117 \text{ MeV}$. We qualitatively reproduced the overall structure observed in the thickness maps derived from fringing analyses for ACS/HRC, ACS/WFC, and WFC3/UVIS. Next, we observed anticorrelation between solar activity and cosmic ray particle flux.

A spectral analysis of the cosmic ray particle flux over time revealed two signals with peaks at ~ 11 yr (the solar cycle) and ~ 48 days (unknown origin). The modulation of the cosmic ray particle flux by the solar cycle indicates that the majority of the cosmic rays observed by HST are GCRs.

We compiled a variety of useful statistics describing the distributions of the morphological parameters computed for cosmic rays observed in each of the five CCD imagers. These values can be used to quantify the impact of cosmic rays on observations with ACS, STIS, and WFC3 for proposal planning purposes. For example, the number of pixels affected by cosmic rays in a given observation can be computed using the cosmic ray particle flux and the number of pixels a cosmic ray will affect on average.

We analyzed observations made with the STIS/CCD taken during passages through the center of the SAA. We found that the cosmic ray particle flux increased by a factor of 800 over a span of ~ 350 s from the edge of the SAA to the center. Using the rate of energy deposition, we estimated the peak cosmic ray particle flux to be $1092.79 \text{ CR s}^{-1} \text{ cm}^{-2}$ at a location of $41^\circ 61' \text{W}$ and $22^\circ 96' \text{S}$, which is in good agreement with the contemporaneous centroid derived by Fürst et al. (2009). A spatial analysis of 54,215 observations identified two “hot spots,” one over North America and one extending from South Africa to the western coast of Australia, where the cosmic ray particle flux increases to more than 5σ above the nominal value.

The data used in this paper were obtained from the Mikulski Archive for Space Telescopes (MAST) and its cloud-hosted

counterpart, the HST Public Data Set (<https://registry.opendata.aws/hst/>), on AWS. STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. Support for N.D.M. and S.E.D. was provided by NASA through grant number HST AR 14587. Thank you to Chris Long for detailed engineering discussions and information on HST operations, and to Jeremy Walsh and Michael Wong for discussions on fringing models. We also thank the referee, whose comments and suggestions greatly improved this paper.

This research has made use of NASA’s Astrophysics Data System. This work made use of the following open source Python packages; *astropy*, *astroquery*, *cartopy* (Met Office 2010), *dask*, *iminuit*, *matplotlib* (Hunter 2007), *numpy* (Harris et al. 2020), *pandas* (Pandas Development Team 2020), *scipy* (Virtanen et al. 2020), and *statsmodels*.

Appendix A STIS SAA Observations

Observations made during an SAA passage with the STIS CCD in February of 1997 are listed in Table 13. Column 1 is the position (see Figure 20) of HST in the SAA, column 2 is the dataset id of the exposure, OBSID; columns 3 and 4 are Latitude and Longitude of HST, column 5 is the Rate of Energy Deposition by cosmic ray particles during the exposure; column 6 is the Cosmic Ray Flux in particles $\text{s}^{-1} \text{ cm}^{-2}$; and column 7 lists the Number of Cosmic Rays in the exposure. Discussion of the calculations is provided in Section 4.6 of the paper.

Table 13
The Estimated Cosmic Ray Particle Flux and Counts Computed for the STIS/CCD SAA Data Set

Position	OBSID	Latitude (deg)	Longitude (deg)	Rate of Energy Deposition $\text{e}^{-} \text{s}^{-1} \text{ cm}^{-2}$	Cosmic Ray Flux $\text{CR s}^{-1} \text{ cm}^{-2}$	Number of Cosmic Rays
1	o3st20gcq_fit.fits	-3.81	-82.76	3.57×10^3	1.36	472.52
2	o3st20gdq_fit.fits	-6.96	-77.26	3.68×10^4	14.04	4868.80
3	o3st20geq_fit.fits	-10.03	-71.68	2.35×10^5	89.59	3.11×10^4
4	o3st20gfq_fit.fits	-12.99	-66.00	5.91×10^5	225.32	7.81×10^4
5	o3st20ggq_fit.fits	-15.80	-60.17	1.22×10^6	466.61	1.62×10^5
6	o3st20ghq_fit.fits	-18.43	-54.18	1.92×10^6	730.41	2.53×10^5
7	o3st20giq_fit.fits	-20.83	-47.99	2.68×10^6	1021.34	3.54×10^5
8	o3st20gjg_fit.fits	-22.96	-41.61	2.86×10^6	1092.79	3.79×10^5
9	o3st20gkq_fit.fits	-24.79	-35.03	1.98×10^6	755.76	2.62×10^5
10	o3st20glq_fit.fits	-26.28	-28.25	1.12×10^6	428.76	1.49×10^5
11	o3st20gmq_fit.fits	-27.41	-21.32	6.08×10^5	231.84	8.04×10^4
12	o3st20gnq_fit.fits	-28.13	-14.27	2.56×10^5	97.66	3.39×10^4
13	o3st20goq_fit.fits	-28.44	-7.14	1.19×10^5	45.52	1.58×10^4
14	o3st20gpq_fit.fits	-28.33	-0.00	5.57×10^4	21.26	7373.38
15	o3st20gqq_fit.fits	-27.80	7.09	3.48×10^4	13.29	4609.25
16	o3st20grq_fit.fits	-26.87	14.07	2.61×10^4	9.95	3450.12
17	o3st20gsq_fit.fits	-25.55	20.92	1.57×10^4	5.97	2071.02
18	o3st20gtq_fit.fits	-23.88	27.58	5.79×10^3	2.21	766.55
19	o3st20guq_fit.fits	-21.90	34.04	4.49×10^3	1.71	594.29

Note. Exposure time for all images was 60 s.

Appendix B Useful Relativistic Relationships

B.1. Momentum and Kinetic Energy

The relativistic momentum–energy relation defines a relationship between the total relativistic energy, E , the relativistic momentum, p , and the rest mass, m_0 , given by

$$E^2 = (pc)^2 + (m_0c^2)^2. \quad (\text{B1})$$

The total relativistic energy given by $E = \gamma m_0c^2$ and the relativistic momentum is given by $p = \gamma m_0v$. Using Equation (B1), we can express the relativistic momentum of a particle in terms of the relativistic kinetic energy, $T = (\gamma - 1)m_0c^2$, as follows:

$$\begin{aligned} (pc)^2 + (m_0c^2)^2 &= (T + m_0c^2)^2 \\ (pc)^2 &= (T + m_0c^2)^2 - (m_0c^2)^2 \\ p &= \frac{\sqrt{T^2 + 2Tm_0c^2}}{c} \end{aligned} \quad (\text{B2})$$

B.2. Rigidity

For a relativistic particle traveling in a uniform magnetic field, the Lorentz force will cause the particle to undergo uniform circular motion around the field lines. The magnetic rigidity is the product of the radius of the orbit and the magnetic field strength,




$$R = B\rho = \frac{pc}{Ze}, \quad (\text{B3})$$

where p is the momentum, c is the speed of light, Z is the charge number, and e is the elementary charge. For relativistic particles, the momentum is often quoted in units of $\text{GeV } c^{-1}$. When using these units in Equation (B3), the rigidity of a $1 \text{ GeV } c^{-1}$ particle will be

$$\begin{aligned} R &= \frac{pc}{Ze} \\ R &= \frac{(1 \text{ GeV } c^{-1})c}{Ze} \\ R &= \frac{10^9 \text{ eV}}{Ze} \\ R &= \frac{10^9}{Z} \left(\frac{1.602 \times 10^{-19} \text{ J}}{e} \right) \\ R &= \frac{10^9}{Z} \left(\frac{1.602 \times 10^{-19} \text{ J}}{1.602 \times 10^{-19} \text{ C}} \right) \\ R &= \frac{10^9 \text{ V}}{Z} = 1 \text{ GV}/Z. \end{aligned} \quad (\text{B4})$$

Thus, we see that for a proton that has a momentum of $10 \text{ GeV } c^{-1}$, it has a magnetic rigidity of 10 GV.

ORCID iDs

Nathan D. Miles  <https://orcid.org/0000-0001-8936-4545>
Susana E. Deustua  <https://orcid.org/0000-0003-2823-360X>
Gonzalo Tancredi  <https://orcid.org/0000-0002-4943-8623>

References

- Bak, J., Burenkov, A., Petersen, J., et al. 1987, *NuPhB*, **288**, 681
- Barker, E. A., McCullough, P., & Martel, A. R. 2010, WFC3 IR SAA Passage Behavior, Space Telescope Science Institute [WFC3 ISR 2009-40](#)
- Bichsel, H. 1988, *RvMP*, **60**, 663
- Borncamp, D., Grogin, N., Bourque, M., & Ogaz, S. 2017, Pixel History for Advanced Camera for Surveys Wide Field Channel, Space Telescope Science Institute [ACS ISR 2017-05](#)
- Corti, C., Potgieter, M. S., Bindi, V., et al. 2019, *ApJ*, **871**, 253
- Dask Development Team 2016, (Dask: Library for dynamic task scheduling), <https://dask.org>
- Delaboudinière, J. P., Artzner, G. E., Brunaud, J., et al. 1995, *SoPh*, **162**, 291
- Dembinski, H., Ongmongkolkul, P., Deil, C., et al. 2020, scikit-hep/iminuit: v1.4.9, Zenodo, doi:[10.5281/zenodo.3951328](https://doi.org/10.5281/zenodo.3951328)
- Desjardins, T. D., & Lucas, R. 2019, Introduction to the Hubble Space Telescope Data Handbooks, Vol. 9 (Baltimore, MD: Space Telescope Science Institute)
- Didkovsky, L. V., Judge, D. L., Jones, A. R., Rhodes, E. J., Jr., & Gurman, J. B. 2006, *AN*, **327**, 314
- Di Felice, V., Pizzolotto, C., D’Urso, D., et al. 2017, ICRC (Busan), **301**, 1073
- Dressel, L., et al. 2019, Wide Field Camera 3 Instrument Handbook, Version 12.0 (Baltimore, MD: Space Telescope Science Institute)
- Forbush, S. E. 1954, *JGR*, **59**, 525
- Fürst, F., Wilms, J., Rothschild, R. E., et al. 2009, *E&PSL*, **281**, 125
- Ginsburg, A., Sipocz, B., Parikh, M., et al. 2018, astropy/astroquery: v0.3.9 release, vv0.3.9, Zenodo, doi:[10.5281/zenodo.2002585](https://doi.org/10.5281/zenodo.2002585)
- Hale, G. E., Ellerman, F., Nicholson, S. B., & Joy, A. H. 1919, *ApJ*, **49**, 153
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Natur*, **585**, 357
- Hathaway, D. H. 2015, *LRSP*, **12**, 4
- Hatlo, M., James, F., Mato, P., et al. 2005, *ITNS*, **52**, 2818
- Heynderickx, D. 2002, *IJMPA*, **17**, 1675
- Hopkinson, G. R. 1987, *OptEn*, **26**, 766
- Hunter, J. D. 2007, *CSE*, **9**, 90
- James, F., & Roos, M. 1975, *CoPhC*, **10**, 343
- Janesick, J. R. 2001, Scientific Charge-Coupled Devices (Bellingham, WA: SPIE Optical Engineering Press)
- Landau, L. D. 1944, *J. Exp. Phys. (USSR)*, **8**, 201
- Lockwood, J. A. 1958, *PhRv*, **112**, 1750
- Lowe, B., & Sareen, R. 2007, *NIMPA*, **576**, 367
- Lupie, O. L. 2002, Pre-Test SAA Contours for WFC3, Space Telescope Science Institute [WFC3 ISR 2002-01](#)
- Malumuth, E. M., Hill, R. S., Gull, T., et al. 2003, *PASP*, **115**, 218
- Martel, A. R., Borders, T., & Barker, E. A. 2009, WFC3 SMOV Proposal 11448: UVIS SAA Passage Behavior, Space Telescope Science Institute [WFC3 ISR 2009-47](#)
- McCully, C., Crawford, S., Kovacs, G., et al. 2018, astropy/astrocrappy: v1.0.5, Zenodo, doi:[10.5281/zenodo.1482019](https://doi.org/10.5281/zenodo.1482019)
- McMaster, M. 2008, Wide Field and Planetary Camera 2 Instrument Handbook, Version 10.0 (Baltimore, MD: Space Telescope Science Institute)
- Meroli, S., Biagetti, D., Passeri, D., et al. 2011a, *NIMPA*, **650**, 230
- Meroli, S., Passeri, D., & Servoli, L. 2011b, *JInst*, **6**, P06013
- Met Office 2010, Cartopy: a cartographic python library with a matplotlib interface, <http://scitools.org.uk/cartopy>
- Miles, N. D., Deustua, S., & Tancredi, G. 2020, arXiv:[2011.11604](https://arxiv.org/abs/2011.11604)
- Mutchler, M., & Sirianni, M. 2005, Internal monitoring of ACS charge transfer efficiency, Space Telescope Science Institute [ACS ISR 2005-03](#)
- Pandas Development Team 2020, pandas-dev/pandas: Pandas, Zenodo, doi:[10.5281/zenodo.4681666](https://doi.org/10.5281/zenodo.4681666)
- Potgieter, M., Mwiinga, N., Ferreira, S., Manuel, R., & Ndiitwani, D. 2013, *JAdR*, **4**, 259
- Potgieter, M. S. 2013, *LRSP*, **10**, 3
- Ravindra, N. M., & Narayan, J. 1987, *JAP*, **61**, 2017
- Riess, A. 2002, A First Look at Cosmic Rays on ACS, Space Telescope Science Institute [ACS ISR 2002-07](#)
- Riley, A., et al. 2019, STIS Instrument Handbook, Version 19.0 (Baltimore, MD: Space Telescope Science Institute)
- Russell, C., Jian, L., & Luhmann, J. 2019, *RvGeo*, **57**, 1129
- Ryon, J., Grogin, N., Anderson, J., et al. 2019, Advanced Camera for Surveys Instrument Handbook for Cycle 28, Version 19 (Baltimore, MD: Space Telescope Science Institute)

- Schaefer, R. K., Paxton, L. J., Selby, C., et al. 2016, *SpWea*, **14**, 330
- Schnyder, G., Nesmachnow, S., & Tancredi, G. 2017a, in *High Performance Computing*, ed. C. Barrios Hernández, I. Gitler, & J. Klapp, Vol. 697 (Berlin: Springer), 414
- Schnyder, G., Nesmachnow, S., Tancredi, G., & Tcherykh, A. 2017b, in *High Performance Computing*, ed. E. Mocsos & S. Nesmachnow, Vol. 796 (Berlin: Springer), 359
- Seabold, S., & Perktold, J. 2010, in *9th Python in Science Conf., SciPy 2010*
- Shen, Z.-N., & Qin, G. 2016, *JGRA*, **121**, 712
- SILSO World Data Center 1991-2019, *International Sunspot Number Monthly Bulletin and Online Catalogue* (Brussels: Royal Observatory of Belgium)
- Simpson, J. 1958, *Annals of the International Geophysical Year IV, Part VII* (Oxford: Pergamon), 351
- Simpson, J. A. 1994, *JGR*, **99**, 19159
- Smart, D., & Shea, M. 2005, *AdSpR*, **36**, 2012
- Tanabashi, M., Hagiwara, K., Hikasa, K., et al. 2018, *PhRvD*, **98**, 030001
- Tauxe, L., Shaar, R., Jonestrask, L., et al. 2016, *GCG*, **17**, 2450
- The Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, arXiv:1801.02634
- The HDF Group 2010, *Hierarchical data format version 5*, <http://www.hdfgroup.org/HDF5>
- Thébaud, E., Finlay, C. C., Beggan, C. D., et al. 2015, *EP&S*, **67**, 79
- van Dokkum, P. G. 2001, *PASP*, **113**, 1420
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *NatMe*, **17**, 261
- Walsh, J. R., Freudling, W., Pirzkal, N., & Pasquali, A. 2003, *Modelling the fringing of the ACS WFC and HRC chips*, ST-ECF Instrument Science Report, ACS 2003-012, Space Telescope European Coordinating Facility
- Walterscheid, R. L. 1989, *JSpRo*, **26**, 439
- Wong, M. H. 2010, in *2010 Space Telescope Science Institute Calibration Workshop, Hubble after SM4. Preparing JWST*, ed. S. Deustua & C. Oliceira (Baltimore, MD: Space Telescope Science Institute), 22