Toward a Data-driven Model of the Sky from Low Earth Orbit as Observed by the Hubble Space Telescope

Sarah E. Caddy1,2, Lee R. Spitter1,2,3, and Simon C. Ellis2,3

1 School of Mathematical and Physical Sciences, Macquarie University, Sydney, NSW 2109, Australia; sarah.caddy@mq.edu.au
2 Research Centre in Astronomy, Astrophysics & Astrophotonics, Macquarie University, Sydney, NSW 2109, Australia
3 Australian Astronomical Optics, Faculty of Science and Engineering, Macquarie University, Macquarie Park, NSW 2113, Australia

Received 2022 January 18; revised 2022 May 24; accepted 2022 June 6; published 2022 July 15

Abstract

The sky observed by space telescopes in Low Earth Orbit (LEO) can be dominated by stray light from multiple sources including Earth, Sun, and Moon. This stray light presents a significant challenge to missions that aim to make a secure measurement of the extragalactic background light (EBL). In this work, we quantify the impact of stray light on sky observations made by the Hubble Space Telescope (HST) Advanced Camera for Surveys. By selecting on orbital parameters, we successfully isolate images with sky that contain minimal and high levels of earthshine. In addition, we find weather observations from CERES satellites correlate with the observed HST sky surface brightness indicating the value of incorporating such data to characterize the sky. Finally, we present a machine-learning model of the sky trained on the data used in this work to predict the total observed sky surface brightness. We demonstrate that our initial model is able to predict the total sky brightness under a range of conditions to within 3.9% of the true measured sky. Moreover, we find that the model matches the stray-light-free observations better than current physical zodiacal light models.

Unified Astronomy Thesaurus concepts: Cosmic background radiation (317); Diffuse radiation (383); Zodiacal cloud (1845)

1. Introduction

The total sky observed by a low-Earth-orbiting (LEO) space telescope can contain significant levels of “stray light,” which contains complex and time-varying sources like earthshine, sunshine, and moonshine as well as bright stars outside the optical field of view. In this work, we explore whether we can characterize the observing conditions where stray light is most likely to impact observations from LEO and predict the total intensity of the sky in order to support science goals to measure both diffuse and low-surface-brightness signatures that require precise sky photometry.

A number of LEO space telescopes (e.g., SPHEREx, Korngut et al. 2018; SkyHopper,4 MESSIER Lombardo et al. 2019), sounding rocket missions (e.g., CIBER, Korngut et al. 2013; CIBER2, Korngut et al. 2022), and large archival surveys from existing instruments (e.g., SKYSURF, Windhorst et al. 2022; Carleton et al. 2022) must deal with sources of stray light in order to make a secure measurement of the extragalactic background light (EBL). The EBL consists of all optical/infrared sources over all cosmic time and is considered a valuable benchmark to assess competing theories of structure formation of the universe.

Attempts to make a secure measurement of the EBL have resulted in a significant unresolved discrepancy between direct measurements and estimates derived from integrated galaxy counts (see Cooray 2016 for a review). If the EBL is derived from galaxies, these estimates should converge (Driver et al. 2016). However, this is not yet found to be the case (Mattila 1976; Spinrad & Stone 1978; Bougha & Kuhn 1986; Mattila 1990; Bernstein 2007; Matsuoka et al. 2011; Korngut et al. 2013; Driver et al. 2016; Mattila et al. 2017; Mattila & Väisänen 2019; Korngut et al. 2022; Lauer et al. 2022). Current best estimates from direct measurements are up to a factor of 5 times larger than estimates from galaxy counts (Driver et al. 2016). Indirect measurements made by studying the signature of the EBL on the spectra of bright gamma-ray sources also suffer from systematic errors (Korngut et al. 2022). However, it provides another independent detection of a lower limit that leaves little room for the presence of unresolved sources (Desai et al. 2019). In contrast, measurements made by Lauer et al. (2022) with New Horizons LORRI images in the outer solar system reported a potential detection of excess light that was not explained by resolved galaxies (Korngut et al. 2022). This discrepancy may indicate the presence of a truly diffuse component of the EBL undetected by galaxy counts. It could also indicate an incompleteness in the model of sky brightness from LEO that may lead to overestimated sky subtractions in estimates derived from direct EBL observations by space telescopes (see Windhorst et al. 2022 for examples).

Direct measurement of the EBL from all unresolved optical/infrared sources (Matsuoka et al. 2011; Driver et al. 2016; Zemcov et al. 2017; Mattila & Väisänen 2019) has been challenging to isolate from the numerous other components (including stray light) that make up the total observed sky.

The observed sky from the UV to near-IR can be broken into several components. Each component originates from a different source, but once combined they are difficult to separate because the sources do not have strongly defining spectral features that can be used to distinguish one from another. The components of the diffuse sky from LEO include four main contributions:
1. Zodiacal light, composed of both scattered and thermally radiated sunlight from interplanetary dust particles distributed nonuniformly in the solar system (Kelsall et al. 1998).

2. The diffuse galactic light, composed of scattered light from dust and gas in the Milky Way’s interstellar medium (Leinert et al. 1998).

3. The EBL, composed of all radiating and re-emitting sources integrated over all space, time, and wavelengths (Driver et al. 2016). The EBL can be described by two components: the diffuse EBL comprising unresolved sources, and the discrete EBL consisting of resolvable sources.

4. Stray light, which here is defined as any light scattered or emitted from outside the field of view of the instrument. It includes earthshine, moonshine, or sunlight. It also includes other bright astronomical sources like stars (e.g., Borlaff et al. 2021), which we do not address in this work specifically.

The two brightest and most variable components of the sky are the zodiacal light and stray light. Measuring these components poses the most challenging barrier in securing a measurement of the EBL from LEO (e.g., Korngut et al. 2022).

Efforts to constrain zodiacal light have resulted in time-varying models that describe the zodiacal light surface brightness over the entire sky using data from COBE/DIRBE (Kelsall et al. 1998), Spitzer (Krick et al. 2012), and/or SMEI (Buffington et al. 2016). The two most commonly used models are the Wright model (Wright 2001) and the Kelsall model (Korngut et al. 2022). A limitation of these models is their reliance on poorly constrained models of the interplanetary dust distribution within the solar system (Jorgensen et al. 2021). Indeed, the Kelsall and Wright models disagree by as much as the total intensity of the expected EBL itself (Korngut et al. 2022; Driver et al. 2016). This systematic uncertainty makes it challenging to distinguish between the EBL and the zodiacal light sky.

Stray light from earthshine and sunshine can be an important component of the total observed sky for space telescopes in LEO observing at optical wavelengths because the sources are so luminous. At \( \sim 0.3 \)–2 \( \mu \)m, earthshine is dominated by reflected sunlight and is characterized by a scattered solar spectrum with the addition of absorption features from the atmosphere, clouds, forests, oceans, and deserts (Giavalisco et al. 2002; Doelling et al. 2013). As a result, the stochastic nature of the weather combined with time-varying surface features beneath a space telescope makes earthshine stray light difficult to model accurately.

The easiest solution to mitigate the impact of stray light from sunshine and earthshine for observers is to simply avoid it by restricting the range of the telescope’s pointing with respect to potential stray-light sources (Shaw et al. 1998; Giavalisco et al. 2002; Korngut et al. 2018). This comes at a cost to the productivity of the instrument, reducing the visit time for particular targets. For this reason, pointing restrictions are usually not designed to completely avoid earthshine or sunshine stray light, only minimize it. Indeed, using the data set in the present study, we estimated that potentially more than half of the exposures in the Hubble Legacy Archive (HLA) may be contaminated by some sort of stray light based on the Sun altitude at which the exposure was taken.

Various studies have characterized stray light from space telescopes. For example, Murthy (2014) use archival imaging products from the GALEX space telescope to find two components in the UV sky: one that is time-variable symmetrical in intensity about the local midnight of the orbiting observatory and one that is proportional to the angle between the Sun and the observed field. Borlaff et al. (2021) used Hubble Space Telescope (HST) stray-light information in the Hubble Ultra Deep Field (HUDF) to better inform sky estimates for future missions such as Euclid and JWST. Prior work to understand how HST orbital parameters and telescope attitude impact the presence of stray light has led to rough estimates of the intensity of earthshine stray-light contributions (Shaw et al. 1998; Giavalisco et al. 2002; Biretta et al. 2003; Baggett & Anderson 2012; Brammer et al. 2014). Some of this work informs the three options currently available in the HST Exposure Time Calculator for earthshine contribution (average, high, or extremely high) with an important caveat that these often do not reflect true conditions during operations (Giavalisco et al. 2002).

The impact of earthshine on space-based telescopes in LEO is nicely illustrated by the work of Luger et al. (2019). They detect earthshine stray-light variations in Transiting Exoplanet Survey Satellite (TESS; Borucki et al. 2010) data and use it to map the contents on the surface of Earth. In doing so they demonstrate that TESS is an “effective addition to NASA’s fleet of weather satellites.” In this work we explore this idea further and aim to answer the question: Can NASA’s fleet of weather satellites be used to inform the impact of earthshine on LEO space telescopes?

In this work, we create a data-driven model to assess and predict the impact of stray light. The analysis is used to inform when HST data are likely to have minimal to no stray light from Earth, Sun, and Moon. We also illustrate the promising features of having a data-driven model to support EBL science goals. A future version of the model might be used to improve exposure time calculations by using a highly tuned prediction for the sky surface brightness for a given field observed at a particular time with HST or future missions.

In detail, here we extend the work of Caddy (2019) and construct an empirically derived model of the sky brightness observed by HST. We leverage the flexibility and accuracy of the machine-learning algorithm XGBoost (Chen & Guestrin 2016) and the extensive data of the Hubble Legacy Archive (HLA) composed of hundreds of thousands of exposures—spanning decades—in multiple filters to create a useful tool that aims to predict stray light from LEO. We describe the results of the constructed models using calculated and collated orbital parameters of HST, the median-clipped sky in a sample of over 34,000 Advanced Camera for Surveys Advanced Camera for Surveys (ACS; Sirianni et al. 2005) images, and the earthshine below the space telescope derived from simultaneous satellite imagery from the CERES missions. We explore the impact of each pointing parameter on the contribution of local stray light to the total sky and demonstrate the benefits of constructing an empirically generated sky model that incorporates all foreground stray-light sources, as opposed to identifying and modeling each component of the sky separately. Finally, we demonstrate the use and functionality of the model utilizing GOODS North and describe future work made possible by these results.
2. Method

In this section, we describe the observational data and derived quantities used in this work. We also describe the earthshine observations made by orbiting weather satellites and how these data are translated to fluxes beneath HST. The calibration of the HST data, HST data quality control, and the construction of a geometric model describing HST’s pointing relative to Earth are presented, and we describe the XGBoost machine-learning model used in this work to predict the total intensity of the sky.

2.1. Astrodrizzle Sky Estimates

In this study, we use data derived from \( \sim34,300 \) individual exposures taken by the ACS on HST from 2002 to 2020. We use only ACS data because sky estimates are readily available in the (now-deprecated) StarView engineering database (Holl & Deul 1992). Four filters are used for this study (F435W, F606W, F814W, and F850LP) due to the large amount of available data to sample the optical stray light and zodiacal light. The F850LP GOODS North (Dickinson et al. 2006) data are particularly useful because early observations are known to have high levels of stray-light contamination (Kawara et al. 2017). The locations of the fields on sky are shown in galactic coordinates in Figure 1 and, aside from poor sampling near the galactic plane, are relatively well distributed over time and space.

Attitude parameters define the orientation of the telescope’s axes with respect to Earth, Sun, and Moon and are used as indicators of stray-light contamination. These variables are described in Figure 2 and are retrieved for every available ACS exposure in standard FITS headers from the database StarView. The relevant FITS header keywords and StarView keywords are summarized in Table 15. The median sky surface brightness estimated for each exposure is taken from the FITS header keyword MDRIZSKY, which is computed by an automated sky subtraction routine in Astrodrizzle in the STScI Drizpack software (Hack et al. 2019). These FITS header keywords were obtained from the StarView database. The MDRIZSKY is computed by sigma-clipping pixels with outlying values. After each of the five sigma-clipping iterations, the standard deviation of the pixel values is computed, and pixels deviating from the mean value by more than 4\( \sigma \) are rejected. The median value of nonrejected pixels is the adopted estimate of the sky level. The number of pixels used in the calculation of the median sky value is not retained or recorded during the Drizzling process, so we cannot easily estimate uncertainties for individual measurements. We will address this shortcoming in a future work. The sky estimates generated by Astrodrizzle are given in units of uncalibrated instrumental flux e\(^{-}\)/pixel in ACS images. For each observation, the zero-point can be derived for the specific instrument configuration given photometric header keywords. PHOTFLAM (erg cm\(^{-2}\)/\( \text{Å} \)/e\(^{-}\)) for each filter is used as the scaling factor necessary to transform an instrumental flux in units of e\(^{-}\)/pixel to a physical flux density. The subtracted sky estimate can then be represented in terms of physical flux density in units of erg\(^{-1}\) s cm\(^{-2}\) arcsec\(^{-2}\)/\( \text{Å} \) by dividing by the squared pixel scale of the instrument and the exposure time:

\[
F_\lambda = \frac{(\text{MDRIZSKY} \times \text{PHOTFLAM})}{\text{EXPTIME} \times 0.05^2}. \tag{1}
\]

We use the original pixel scale of ACS/WFC of 0.05" pixel\(^{-1}\) because drizzling images to a range of different pixel scales generally occurs after MDRIZSKY is calculated from the physical instrument pixel scale (Hack et al. 2019).

Figure 1. The location of fields used in this study in four selected filters. Field location is plotted in galactic coordinates, with the galactic plane from \( \pm 15^\circ \) shown in shaded gray. The ecliptic plane is plotted in black. ACS F435W, F606W, and F814W are chosen due to the quantity of data available with existing archival sky surface brightness estimates over the entire sky. ACS F850LP is chosen because of data in the GOODS North fields, which are known to have high levels of earthshine contamination.
2.2. Sky Estimates Data Quality Control

Particularly crowded fields such as star clusters, planetary targets, and large foreground targets such as NGC objects take up the entire field of view and can corrupt the automatically generated MDRIZSKY sky estimates. For this reason, fields containing these targets were removed from the data set by screening the StarView field name Target Name, and in some cases confirming this information with the proposal ID available on the HLA. Fields with exposure times of less than 500 s are also removed to reduce the impact of charge-transfer efficiency losses for particularly faint sky observations. Exposure times are limited to less than 2000 s to reduce the impact of longer-duration exposures sampling both day and night sides of the orbit in one exposure. Duplicated data with identical start times and keyword parameters are removed. Calibration fields or unknown targets that are sometimes returned in error from the StarView database were also removed from the analysis.

Unfortunately, due to unknown reasons, the information in the FITS headers provided on StarView is not always correct or complete. As a result, other outliers that may impact the accuracy of the measured sky include grism images, calibration images, and images with failed guide star acquisition. Every effort has been made to pick out these images by manual inspection of the FITS headers and verifying with proposal IDs on the HLA, but some may remain.

From the above quality control, we reduced the initial data set from 280,105 individual exposures to 34,308 exposures, which were used in the analysis. For each filter, we have 2468, 9332, 15,844, and 6664 exposures for the F435W, F606W, F814W, and F850LP filters, respectively.

For each field, we constructed a database consisting of raw FITS header information and derived quantities. See Section 5 for a table of relevant FITS header keywords and StarView field names. This includes orbital information, time, date, exposure times, photometric calibration parameters, start and end date and times for every exposure, filter combinations for wheels 1 and 2, and field coordinates.

2.3. Satellite Weather Data

Historical satellite weather data should be sensitive to the amount of earthshine beneath a space telescope at a given time. In this work we utilize data from the ongoing CERES (Clouds and the Earth’s Radiant Energy System) missions. CERES is a multisatellite project dedicated to observing Earth’s global energy budget. The first satellite was launched in 1997, closely followed by Terra and Aqua in 1999 and 2002, respectively, in Sun-synchronous orbits. CERES still operates to the present time (Doelling et al. 2013), and CERES data are available for 60%–70% of HST’s operational lifetime.

CERES mission data are available online through NASA’s Langley Research Center.5 We choose to focus only on the so-called “shortwave” CERES filter (0.3–5 μm) that most closely measures earthshine levels in the chosen HST filters. An important caveat of interpreting analysis using the CERES data is that it is sensitive to thermal emission, while our HST filters will not be. CERES fluxes are total integrated flux measurements from the top of the atmosphere over a 1° × 1° latitude and longitude grid over a 3 hr period. The CERES shortwave filter, HST filters, and the earthshine and zodiacal light SEDs are shown in Figure 3. CERES shortwave instruments do not operate or are not sensitive to the nightside of Earth. As a result, data products are only strictly relevant to HST when acquired over the daytime portion of Earth.

To compute the relevant CERES irradiance for a given observation, the primary scattering area (see Figure 2) of an exposure is projected onto the CERES spatial image of the entire illuminated surface of Earth (see Figure 4). The average earthshine flux inside the primary scattering area is calculated for the start, middle, and end of each exposure. The calculated

---

5 https://ceres.larc.nasa.gov/data/
earthshine in CERES native units of $W/m^2$ beneath the telescope for any one exposure is then the average of the three primary scattering area averages. For locations that have no CERES data (e.g., nighttime side), the CERES image values are zero and will impact the averages in a way that accurately reflects the amount of earthshine that potentially impacts an observation.

We calculate a primary scattering area beneath the telescope (see Figure 2) for the start, middle, and end of every exposure. For this calculation, the location of the telescope needs to be known while an exposure is being acquired. The ephemeris for HST was retrieved from the North American Aerospace Defense Command Archive. This includes data for HST over the past 20 yr was kindly provided by private request. The ephemeris data were processed by the Python package pyephem to determine the location of HST at the start, middle, and end of an exposure. The altitude of the telescope at the start of the exposure was adopted for the entire exposure.

The radius of the primary scattering area is then given by

$$R_{\text{earthshine}} = \cos^{-1} \left( \frac{R_{\text{Earth}}}{R_{\text{Earth}} + Alt} \right) \times R_{\text{Earth}}$$

For all calculations, the radius of Earth is assumed to be 6370 km and the exposures in this work have an average $R_{\text{earthshine}}$ of 2535 km.

A primary scattering area was computed for the location of HST at the start, middle, and end of an exposure. The final primary scattering area for a given exposure was taken as the region covered by all three primary scattering areas. This tiling approach provides adequate overlapped sampling of the earthshine beneath the telescope over the duration of the exposure.

![Figure 3. Filters used in this work plotted with respect to scaled spectra of components of the total sky. The earthshine spectrum is approximated by an E-490 solar spectrum following Giavalisco et al. (2002) and is shown in black. The zodiacal light spectrum is shown in gray and described in Section 3.1. Both spectra are shown with relatively "high" scaling from Giavalisco et al. (2002) Figure 2. Four HST filters that span the visible wavelength portion of the scattered earthshine and zodiacal light spectra are chosen for this study. A portion of the CERES shortwave filter from weather satellites, which extends to 5 μm, is shown in purple.](image)

2.4. Training the Machine-learning Algorithm

As will be shown in Section 3.2, basic trends between telescope attitude parameters, environmental parameters, and the observed sky surface brightness can be visualized by subselecting the data set manually and utilizing basic regression methods. However, due to the complexity and size of the data, we explored the use of a machine-learning algorithm to capture the underlying trends in the data. XGBoost is an efficient and flexible gradient boosting framework (Chen & Guestrin 2016) that can produce a multidimensional interpolation across the input parameters.

We utilize the XGBoost algorithm to ingest and train on the observed sky surface brightness from HST images and corresponding observational parameters given in Section 5. The model can then be used to predict the sky surface brightness for a desired location on the sky, time of year, and set of telescope attitude constraints (e.g., Sun altitude, limb angle, etc.) As described immediately below, the XGBoost algorithm can be trained in two different ways for each filter.

The single-field trained method: This method focuses on identifying underlying trends in sky surface brightness toward a single HST field, made up of many pointings. It allows us to simplify the parameter space to investigate a single issue. For example, we use this method on GOODS North because the field has minimal zodiacal light, which allows us to better isolate the impact of stray light. Similarly, we use this model to understand the benefit of the satellite weather data. The fields we use for this method are the Legacy HST fields, which provide a large amount of data spanning many years and hence give us access to a fairly well-sampled orbital parameter space.

The all-sky trained method: In contrast to the previous method, this method trains on all available data without any spatial restrictions. It is therefore sensitive to all components of the sky (e.g., stray light, zodiacal light, and Galactic dust) and in principle should produce a model that can be used to predict the sky surface brightness for any field taken at any time of year. We will use this model to generate all-sky maps of the sky as seen by HST.

Despite the data quality control described above, there were a few spurious sky estimates that effectively contributed noise to the models. To mitigate the impact of outliers, for each of the 10 runs of the model, we iteratively removed the upper 0.99 and lower 0.01 quantile in the residuals of the measured sky values and model predictions.

Finally, to optimize performance, accuracy is favored over speed for these tests. To prevent overfitting of the model, early stopping rounds are set relatively high (~20). The learning rate is tuned to a low value of 0.01, and we allow the model a large number of estimators at 10,000. Computation time can be sped up considerably with fewer estimators; however, this comes at the cost of accuracy.

The software developed for this work including worked examples, software versions, calibration data, and example data are available on the Skypy github repository.
motivate how we approach the analysis of the data. The figure illustrates some key physical phenomena relevant to this data set. Data points are colored according to the mean Sun altitude at which the exposure is taken. The sky scatter follows a “Christmas tree” shape defined by ecliptic latitude.

We expect multiple parameters to determine the time-variable sky surface brightness in any field observed from LEO. For example, Figure 5 shows an increase in minimum sky surface brightness for fields close to the ecliptic plane. This is most likely due to the contribution of zodiacal light to the total sky.

The lowest sky for any ecliptic latitude occurs when the Sun altitude is less than 0°, which corresponds to the nightside of an orbit. This happens because exposures should be minimally impacted by earthshine stray light in Earth’s shadow at ~0.8 μm.

We also note there is evidence in the plot for other parameters contributing to the sky surface brightness. For instance, at Sun altitude >0°, the sky surface brightness is typically elevated and can reach values ~4× the lowest observable sky. Furthermore, we note that fields with significantly elevated sky surface brightness but low Sun altitude are likely due to Galactic cirrus contributions or poorly calculated sky measurements in crowded fields.

To conclude, Figure 5 shows that zodiacal light and Sun altitude are both important factors in determining the sky. The reason Sun altitude is important is also reflected in Figure 4. At the center of the illuminated portion of Earth, there is a higher amount of reflection from the Sun due to the more optimal reflection geometry. This is why the Sun altitude is a good first-order indicator of the potential for earthshine stray light to impact the sky. Thus, given that earthshine and zodiacal light appear to dominate the sky, we therefore characterize their impact in the next section before exploring other parameters in Figure 7.

### 3.1. Exploring the Dominant Components of the Sky

The impact of the two most dominant components of the sky—earthshine and zodiacal light—in HST fields can be isolated by considering fields in two different locations. Close to the ecliptic plane, the sky will tend to reflect how the zodiacal light varies in time due to the motion of Earth around the Sun caused by changes in viewing angles through the inhomogeneous zodiacal cloud over the period of a year (Leinert et al. 1998). A field far from the ecliptic plane will have minimal zodiacal light that also does not vary significantly. This enables us to separate the impact of earthshine due to changing viewing angles and variable weather from changes due to zodiacal light.

Figure 4. Earthshine observations by the CERES instruments are provided in 1° × 1° latitude and longitude grids with a temporal resolution of 3 hr on the illuminated side of Earth. This figure illustrates one image from CERES projected onto a map of Earth over Australia. Data used here are from the CERES shortwave filter covering 0.3–5 μm.

Figure 5. Total observed sky is plotted as a function of ecliptic latitude to illustrate the change in minimum sky backgrounds due to the planar distribution of zodiacal dust within the solar system. Scatter upwards from the minimum backgrounds shows some correlation with Sun altitude, indicating a likely increased contribution from earthshine stray light. However, the correlation with Sun altitude is not perfect, indicating other underlying parameters impact the observed sky. Single fields observed over a range of Sun altitudes and/or zodiacal levels are apparent as vertical stripes.
To help isolate time-varying stray-light contributions from time-varying zodiacal light intensities, we use zodiacal models to help us identify when a change is dominated by a changing zodiacal light intensity. For this study, we use the zodiacal model described in Leinert et al. (1998), with amendments by Aldering (2002) and calibrated using fluxes from the GOODS North field (Giavalisco et al. 2002). This method is favored over the Wright and Kelsall models as it can be easily adapted to the shorter wavelengths used in this study. It should be noted that the zodiacal light model is not used in the process of the empirical analysis described in Section 3.2 or the XGBoost models described in Section 3.4. This ensures that the results are unbiased by the assumptions that are typically used to construct zodiacal models.

To implement the Aldering (2002) zodiacal model, we adopted the relative intensity of the zodiacal light with respect to the north ecliptic poles as calculated in Gunagala.10 This follows the prescription of Leinert et al. (1998) by interpolating the results of their Table 17. The absolute intensity of the zodiacal light for each of the HST filters is scaled using the observed sky surface brightness found at the north ecliptic poles given in Giavalisco et al. (2002). Following Aldering (2002), the resulting intensity is multiplied by a correction factor of 0.49 for $\lambda > 0.5 \mu m$ and 0.9 for $\lambda < 0.5 \mu m$ to account for the reddening of the solar spectrum in scattered zodiacal light.

As shown in Figure 6, we use COSMOS F814W observations to probe the ecliptic plane and compare it to GOODS North F850LP observations, which are high off the ecliptic plane. While this comparison uses two different filters, we expect this has a negligible impact on the results because the filters (see Figure 3) cover a similar portion of the zodiacal light and earthshine spectral energy distributions (SEDs).

In Figure 6, the expected zodiacal light contribution to the sky is shown for the same time and location on the sky that each exposure is taken. The sky surface brightness is shown colored by the Sun altitude recorded at the midpoint of the exposure. In the GOODS North field, most exposures have a sky surface brightness close to the zodiacal light model estimate of ~0.12 MJy sr$^{-1}$ over the 5 day period. However, some reach as high as ~0.16 MJy sr$^{-1}$. This difference is almost ~34% of the expected zodiacal light intensity.

In the COSMOS field at the start of the observation period, the lowest sky surface brightness is ~0.75 MJy sr$^{-1}$ with deviations reaching ~0.80 MJy sr$^{-1}$, or ~7% of the mean sky surface brightness at the beginning of this period. At the end of the 10 day period, the lowest sky surface brightness drops to ~0.52 MJy sr$^{-1}$ with a maximum of ~0.65 MJy sr$^{-1}$, or ~23% of the lowest sky surface brightness at the end of the observing period.

These results suggest that there are two main time-variable components of the observed sky from LEO in the ~0.8–0.9 $\mu m$ range. One component corresponds to a gradual change as a function of time on the order of days and is seen to follow the predicted zodiacal light estimate. The second component changes on the timescale of hours and is seen to correlate with the Sun altitude of the exposure. For fields close to the ecliptic pole, the effect of Sun altitude variations is most apparent—with the time-variable sky changing by ~34% of the predicted zodiacal light on the timescale of hours.

In this section, we have shown that the data have two major time-varying components, which we show are likely due to varying zodiacal light levels and a component correlated with the Sun altitudes of a particular exposure. As discussed earlier,

10 https://github.com/AstroHuntsman/gunagala
Sun altitude is likely a measure of the potential of stray light from earthshine, thus we have demonstrated that even within a single orbit, HST backgrounds can vary by as much as ~30% due to stray light from earthshine.

3.2. Isolating the Stray-light Component

In Figure 6 we isolated a time-variable component of the sky that is not due to the change in zodiacal light intensity and is correlated with the mean Sun altitude at which the exposure was taken. This result suggests that the orientation of the telescope with respect to sources of stray light is an important factor in understanding the surface brightness of the sky.

Here we attempt to isolate key parameters available to us in HST FITS headers that correlate most strongly with sky surface brightness and may aid in characterizing the stray-light contribution. This is done by projecting the data into two dimensions to illustrate how each parameter defined in Figure 2 impacts the sky surface brightness. Here we use only use exposures from the GOODS North fields. These fields have high ecliptic latitudes and effectively avoid complications from the time-varying zodiacal light component.

Figure 7 shows the sky surface brightness data against various orientation angles from F850LP observations at a high ecliptic latitude. In the top-left panel of this figure, it is evident that the Sun altitude has a correlation with the sky surface brightness and illustrates the idea that stray light from earthshine is significantly more important when HST is observing over a portion of the daytime Earth (i.e., Sun altitude >0°). At low Sun altitudes, HST will not only be over a nighttime portion of Earth, it will also likely be in Earth’s shadow, further reducing possible stray light from sunshine.

The top-right panel of Figure 7 shows a bifurcation at low limb angles, which reflects the underlying trend with the Sun altitude. For the data taken when HST was over daytime Earth, there is a correlation with limb angle. This likely illustrates that more stray light from earthshine can enter the telescope when the field is close to Earth’s limb from the perspective of HST. This is consistent with preliminary work by Biretta et al. (2003), who report that limb angles <25° can show significantly elevated sky levels. We also note that for exposures of Sun altitude >0°, to protect the instrumentation from potential damage, limb angles are restricted to >20°.

The lower-left panel of Figure 7 shows that the Sun angle, which is determined by the position of Earth with respect to its orbit around the Sun, is discretized into groups. These groups roughly map onto HST “visits” of the GOODS North field over time at varying Sun angles. For Sun angles >80°, the minimum sky surface brightness increases. This could be explained by an increase in sunlight entering directly into HST and/or an increased zodiacal light contribution because HST is looking closer to the Sun. The vertical stripes of data correspond to a single visit and illustrate an increase in sky surface brightness due to the various combinations of viewing angles.

The impact of scattered light from the surface of the Moon is explored in the lower-right panel of Figure 7. To minimize the impact of lunar scattered light, HST is restricted to Moon angles greater than 20°. No evident trend is seen in these data, which is limited to Moon angles >50°. This could reflect the fact that Moon angle restrictions of >50° work well. Or it might reflect the relatively sparse temporal sampling of the current data set, which would limit the range of lunations this analysis can probe.

The data presented in Figure 7 indicate that it is possible to minimize the contribution of stray light in HST data by limiting the range of angles in which the telescope can operate (e.g., Biretta et al. 2003). We note that some data do not follow the trends outlined in the following analysis likely because of limitations in the sample quality control described in Section 2.2. The top-left panel of Figure 7 shows minimal scatter for Sun altitudes ≲ −10°. Employing this Sun altitude selection to the data in the top-right panel of Figure 7 we find that scatter is minimal for limb angles ≥20°. With the same Sun altitude selection, we find that the Sun angle should be limited to >80° to minimize the scatter. And finally, we find no strong evidence for a need to select on Moon angle to minimize scatter in the sky observations.

In summary, the data presented in Figure 7 give some indication of what telescope attitude parameters contribute most significantly to the F850LP sky in the direction of the north ecliptic pole. Maximum sky levels are observed when HST is over the daytime sky and when HST is pointing at a field with a low angular distance to the Earth. We also find the angular distance to the Sun is a secondary factor, at least in the limited data set that was used. Note also that we cannot separate the impact of increased zodiacal light intensity from the possibility of direct sunlight scattering into the telescope aperture. The Moon does not appear to significantly impact the sky surface brightness, though this may only reflect the existing HST Moon angle viewing restrictions and/or a limited range of lunations in the current data set.

3.3. Using Satellite Weather Data to Improve Earthshine Contribution Estimates

In Section 1 we raised the idea of improving stray-light predictions from earthshine levels using satellite weather data of Earth. Here we investigate the possibility of utilizing satellite weather data from CERES to see if it can help identify exposures with increased earthshine levels in HST data.

Figure 8 shows the HST sky surface brightness levels against the CERES irradiance. As detailed in Section 2.3, the irradiance is the average of the flux within the primary scattering area, which is the estimated area below HST during observations (see Figure 2). As in Section 3.2, in this analysis, we use data for the F850LP filter in the GOODS North field to minimize the impact of a time-variable zodiacal light component. As discussed in Section 2.3 the CERES data are only given on the daytime side of Earth. For this reason, this analysis is limited to exposures with a Sun altitude >0°.

The data in Figure 8 show a trend where higher CERES levels generally correspond to higher HST sky surface brightnesses. This confirms that earthshine indeed can dominate HST sky surface brightnesses in certain orbital configurations. To explore the cause of the scatter in the relationship between CERES irradiance and HST sky surface brightness, the data in Figure 8 are divided into three limb angle bins as indicated in the figure. Not surprisingly, higher sky surface brightness levels are found at fixed CERES irradiance due to the fact the telescope is pointing closer to Earth at smaller limb angles.

To conclude, we show that data from the Earth-observi-
3.4. The XGBoost Model of the Sky

In the above analysis, we have isolated the zodiacal light and stray-light time-variable components of the sky by restricting analysis to single fields and breaking the problem down into a series of two-dimensional relationships with key telescope attitude parameters. In doing so we have identified initial constraints on telescope attitude to limit the impact of stray light shown in Section 3.2. We also identify a new parameter—the directly observed earthshine beneath a space telescope—that is useful in identifying exposures contaminated by stray light. However, if this analysis is to be extended to larger data sets in multiple fields, multiple wavelengths, and more extensive telescope attitude parameter spaces, a more efficient method of exploring the data is needed.

As detailed in Section 2.4, we constructed a machine-learning model based on the XGBoost algorithm to predict sky surface brightness levels for any set of telescope attitude parameters and locations on the sky. In the following series of tests, we explore how well the model can predict the brightness of the sky in any field observed with a particular set of telescope attitude parameters. To assess the model’s performance, we define a parameter ($\beta$) in units of percentages that...
describes how well the model can predict the sky surface brightness of a test field. Model performance is the mean percentage difference in flux between the measured and predicted sky as defined below:

$$\beta = \frac{\sum (|S_{\text{test}} - S_{\text{pred}}|)}{N} \times 100\%$$

(3)

where $N$ is the size of the test data set, $S_{\text{test}}$ is the test sky surface brightness measurement, and $S_{\text{pred}}$ is the predicted sky estimate from the model.

First, we explored the impact of sample size on the performance to ensure filters with different numbers of test fields are compared equally. This also allowed us to explore the possibility of wavelength-dependent variation that may exist in the data by training the model with data from four different filters. The model is trained using the all-sky-trained method on a randomly selected subset of data from 200–2000 exposures across the entire sky. The test data used to assess the performance of the model are randomly selected each time from a 70:30 split of testing and training data. The performance per training field is calculated by taking the average of 10 individual training runs, with randomly selected training samples to ensure, e.g., single spurious pointings do not dominate the results and also as a way to estimate uncertainties on the model output. In the following, the average performance and associated standard deviation of the 10 repeated training outputs are reported.

Figure 9 shows how the model performance improves with larger training sample sizes. For these initial tests, neither CERES data nor any zodiacal model input is used. We note the performance of each filter is approximately the same when the sample size is similar, suggesting that there is no apparent wavelength dependence of the model’s performance.

We then utilize all available data for each filter to train and test. The performance of each model is found to be better than 10% for all filters, with F814W performing the best at an average of 3.5% ± 0.1%, followed by F850LP at 3.7% ± 0.2%, F606W at 4.0% ± 0.2%, and F435W at 8.0% ± 0.9%. The improvement in performance appears to reflect the training sample size for each filter. These results demonstrate the ability of the model to predict the sky surface brightness of a particular field given a set of unique telescope attitude parameters to well within the ~34% variation in sky surface brightness found in Section 3.1 due to the presence of stray light.

In Section 3.3 we found that satellite weather data contains some information that might be used to better predict the sky surface brightness of a particular field. To explore how folding in CERES data to the model changes the performance, we train the model on the GOODS North field with the F850LP filter with and without calculated CERES irradiance. We find a change in performance of only 0.1% ± 0.3%, which is within the limit of error of the model. Uncertainties on this change in performance were estimated by running the training 10 times using 10 randomly selected training data sets. While this may indicate that CERES data do not significantly improve the performance of the machine-learning model, it is somewhat surprising given the evident correlation between the observed sky and the CERES data in Section 3.3. A perhaps relevant complication with the CERES data is that they cover 0.3–5 μm, so it will contain thermal emission missed by the F850LP filter.

More CERES fluxes calculated for a greater range of HST fields, wavelengths, and observing conditions may help to understand the contribution of CERES data to the machine-learning model. To evaluate how the model performs in comparison to other existing zodiacal light models, we calculate the expected zodiacal light intensity from the Gunagala implementation of the Leinert model described in Section 3.1. The Leinert model does not contain an estimate for the contribution of stray light, so to ensure a fair comparison, we restrict our sample to nighttime data only with Sun altitudes <−10° and limb angles >20°. We use F814W as it is the largest subset of the data presented in this work and choose to use fields across the entire sky to sample a large range of ecliptic latitudes.

The residual sky is calculated by subtracting the model sky from the observed sky and plotted as a function of ecliptic latitude in Figure 10. The XGBoost model presented in this work achieves a Gaussian-like distribution about a mean residual of 0, a standard deviation of 0.03 MJy sr⁻¹, and a β of 3.04%. We note this is slightly lower than the value determined in Figure 9 as the training and testing data set only contains nighttime data. In comparison, the Leinert model has a mean residual of −0.02 MJy sr⁻¹, a standard deviation of 0.06 MJy sr⁻¹, and a β of 16.50%. The Leinert model residuals contain an offset that varies with ecliptic latitude. This offset is largest when fields are located close to the ecliptic plane and may indicate that the model is not capturing the complex structure of the interplanetary dust cloud in the inner solar system (Korngut et al. 2022; Jorgensen et al. 2021).
In summary, the results of Section 3.4 highlight the benefit of an empirically generated sky model. The model is able to perform as well as, if not better than, existing analytical/physical models of the observed sky. It is flexible and can be trained on single or multiple fields, can be easily modified to incorporate new training parameters to improve model accuracy, and performs equally well across F435W, F606W, F814W, and F850LP for similar-sized training data sets. The model is also able to maintain its high level of accuracy when stray-light-impacted fields are introduced, demonstrating that it may be a valuable tool to predict the observed sky brightness under a range of different observing conditions.

4. Discussion

In this section, we explore how the results presented above can be used for two applications. First, we discuss how the initial set of space telescope pointing constraints can be used to isolate the observed darkest sky from existing data and potentially avoid stray light in future missions. Using these initial constraints we isolate observations that are most likely to contain stray light and observations dominated by zodiacal light. We then demonstrate the use of the model to predict the surface brightness over the entire sky and briefly discuss possible applications.

Most space telescopes have hard attitude constraints that dictate how close a target field can be observed with respect to Earth, Sun, and Moon. As described in Section 1, for applications where the science target is diffuse light, like the EBL, one strategy is to adopt more stringent attitude constraints to minimize sources of stray light. Below we apply our initial recommended set of telescope attitude constraints to reduce the impact of stray light on HST observations and to illustrate the potential value of this type of selection.

In Section 3.2, we outlined our initial recommended set of telescope attitude constraints to avoid stray light. To explore how these constraints might be used, a Sun altitude limit of $\beta < -10^\circ$, a limb angle limit of $> 20^\circ$, and a Sun angle limit of $> 80^\circ$ were applied to the F435W, F606W, and F850LP data in the GOODS North field. The resulting subset of sky measurements was averaged and is shown in Figure 11. This SED should have very minimal levels of stray light and should only reflect the next most important component in the sky, the zodiacal light. When this SED is compared to a scaled zodiacal light template, we find good overall agreement in the shape,
supporting the idea the data have been isolated from stray light and resemble the predicted zodiacal light in this field.

Also in Figure 11, we show the average sky surface brightness from the subset of data that we expect should have the highest earthshine stray-light contribution based on the results of Section 3.2. Observations are selected with a Sun altitude of $>60^\circ$ and a limb angle of $<40^\circ$. These limits are chosen to isolate fields that are most likely to contain stray light, and a larger limb angle range is used to ensure data samples are large enough to be statistically significant. A Sun angle limit of $>80^\circ$ is also applied to ensure the contribution of zodiacal light intensity is approximately consistent between samples of minimal stray light and maximum stray light.
The shape of this SED is markedly different than the one with minimal stray light, indicating there is more than just zodiacal light in these fields. We found a zodiacal light plus solar template based on the results of Giavalisco et al. (2002) did not fit the SED for wavelengths shorter than \( \sim 0.4 \mu m \), so we instead attempted to incorporate an estimate of what earthshine spectra might look like. Rayleigh-scattered sunlight is thought to make up a major component of the stray light from earthshine (Woolf et al. 2002); however, we note this is a simplification of an expected earthshine spectrum that is known to contain many absorption features and other scattering properties from surface features and cloud cover. We fit the data with a Rayleigh-scattered modified solar spectrum (90° scattering angle) to mimic earthshine plus a zodiacal light spectrum using a least-squares fit of

\[
F_\lambda = \alpha F_{\text{Zodi}} + \gamma F_{\text{Earthshine}},
\]

where \( \alpha \) and \( \gamma \) are simple scaling parameters for each component of the sky SED. As shown in Figure 11, this reproduces the same F435W excess seen in images most likely to contain stray light. The depression in F606W may be a result of absorption features in the earthshine spectrum in 0.5–0.6 \( \mu m \) wavelengths (Woolf et al. 2002). The best-fit SED includes an average 30.5% \( \pm \) 2.9% earthshine component across the three filters.

The agreement between the template fits and our SEDs in Figure 11 suggests that the method of applying telescope attitude constraints is indeed effective at reducing or enhancing the impact of stray light. It also suggests that exposures most likely to contain stray light may indeed have observed sky that contains a Rayleigh-scattered earthshine component, possibly originating from scattered sunlight in the atmosphere or reflected from surface clouds reaching the telescope aperture. Future work using multiple filters to more finely sampling the sky SED may be able to identify characteristic traits of Earth’s scattered and reflected SED such as water, oxygen, and ozone absorption features (Woolf et al. 2002) to confirm the origin of the stray light.

While our initial set of constraints appears to be promising, we note they were derived using a single field observed with one filter. Moreover, our data are likely not completely free of bad background estimates (see Section 2.2), which will introduce noise into the data. The promising outcome of this work has motivated a more extensive investigation of these issues using essentially the entire HST imaging archive via the SKYSURF Legacy Archival Survey (Windhorst et al. 2022). In this pending work, we have a significantly larger sample size that will allow us to explore numerous other factors (e.g., instrumental contributions to the sky, Galactic dust, etc.) that we have ignored in the present study.

Ultimately, the aim of our work is to improve models of the sky to enable low-surface-brightness science—including observations of the EBL—from LEO. An important result from our work that will support future efforts is the promising use of machine learning to develop a data-driven tool to predict the sky surface brightness for any HST field, which may be extended to future missions.

In Section 3.4 we showed that the empirically driven model can accurately predict the sky surface brightness of a particular field with a set of telescope attitude parameters to an accuracy of \( \beta \approx 4\% \) for filters with the most training data. Further improvements are likely possible and will be explored in future work. The model is also shown to predict sky surface brightness estimates for the zodiacal-light-dominated sky to an accuracy better than some existing analytical/physical models.

The benefit of this approach is that a large parameter space can be reduced to one flexible model, and subtle trends not captured in the simple analysis of Section 3.2 may be teased out across the entire optical/near-IR spectral range. It may also be used to provide improved telescope attitude constraints for LEO space telescopes that can be used to avoid stray light. The method can be extended to future missions with HST-like optics and inform future HST programs.

To further illustrate the potential of this model, we train the model on our entire data set. This allows us to interpolate data across space and time and fix orbital parameters to see what the entire celestial sphere would look like to HST. This enables us to produce all-sky surface brightness maps that represent the sky when contributions from earthshine, moonshine, and sunshine are minimized (or maximized) using our derived constraints in Section 3.2.

The maps presented in Figure 12 were constructed by fixing the parameters according to our recommended constraints in Section 3.2 to reduce the impact of stray light. The maps themselves were created by random sampling of the model’s celestial sphere to remove any artificial structure that may result from uniformly gridded data. Triangular interpolation is then used to resample the results onto identical projections for each filter. The number of predicted points is set corresponding to the sample size of each filter in order to capture the relative uncertainty for each filter. F606W and F814W with the highest sample sizes are composed of 400 predicted points, resampled onto a 100 by 100 uniform grid. F435W and F850LP have considerably fewer training data points and so are resampled onto a 50 by 50 grid. Finer grid size and larger sampling numbers do not change the results significantly.

COBE/DIRBE 1.25 \( \mu m \)\(^{11}\) is included in Figure 12 as a baseline comparison to our HST sky model. A log scale is used to highlight the faint zodiacal belt in COBE/DIRBE wavelengths. The zodiacal belt in HST data follows similar trends to DIRBE data. The zodiacal light in our model is brighter compared to the Galactic plane, which is marginally detected. This most likely reflects poorer sampling due to the avoidance of crowded fields along the Galactic plane as shown in Figure 1. HEALpix maps of these data will be provided on request, and the results can be reproduced following the examples in the Skyspy repository. This comparison illustrates how data-driven models can recover the structure of the sky off the Galactic plane.

These all-sky maps help illustrate another possible application of this type of model, which is to improve exposure time calculations for future HST observations. This works because we effectively will have a highly tuned prediction for the sky surface brightness for a given field observed at a particular orbital configuration with HST. If the observations are sky-background limited, more optimal scheduling might take place that is appropriate for the exact orbital configuration that has been allocated to a particular program or visit. Future work will be conducted to explore if this will lead to significant efficiency in HST scheduling and perhaps motivate a similar study for future missions.

\(^{11}\) https://lambda.gsfc.nasa.gov/product/cobe/dirbe_products.cfm
In this study, we confirm previous work (e.g., Biretta et al. 2003; Giavalisco et al. 2002; Shaw et al. 1998) that stray-light contamination from earthshine is an important factor that influences the intensity of the sky in HST ACS filters. We find that in worst-case scenarios (i.e., when HST is pointing relatively close to daytime Earth’s limb), the measured sky can increase by as much as $4 \times$ the expected zodiacal light surface brightness. Furthermore, it varies by as much as $\sim 30\%$ between exposures due to changes in the telescope attitude with respect to sources of stray light. Stray light from moonshine is still likely to be a problem but we do not see strong evidence in the limited data set we analyzed. Here we do not attempt to decouple stray light from direct sunlight from zodiacal light. The extent and variability of earthshine stray light in HST ACS images as illustrated by this work underscore the need for space telescopes that are designed and specialized to reduce the significant impact of stray light for low-surface-brightness and EBL observations from LEO.

Fortunately, as in Biretta et al. (2003), we find that there are certain restrictions on orbital parameters that significantly minimize stray-light contributions to the observed sky. We find that the Sun altitude and limb angle are the most important factors in determining the potential for stray-light contamination in HST images and should both be constrained to $< -10^\circ$ and $> -20^\circ$, respectively, in order to minimize the contribution of stray light to HST observations. We also present evidence that satellite weather data may be useful in identifying observing conditions that are most likely to result in high earthshine levels; however, we note that more data are needed to confirm these findings.

We also present a new XGBoost machine-learning model trained on this test data set and demonstrate its ability to accurately predict the sky surface brightness over a range of different observing conditions. We demonstrate that this empirical model is able to achieve an overall accuracy that is better than or comparable to existing sky zodiacal light models. For the data set with the largest training sample F814W, we are able to recover the total nighttime sky surface brightness to an accuracy of $\sim 4\%$, which is comparable to the fraction percentage of the lower limit of the EBL from Driver et al. (2016).

We conclude with a note that we will extend this study to a much larger data set. We will utilize the full HST legacy archive via the SKYSURF Legacy Archival Survey (Carleton et al. 2022; Windhorst et al. 2022) for WFC3 and a more extensive ACS data set. This will enable us to extend our work to a more thoroughly sampled parameter space and work toward decoupling sources of stray light from the spatial dependencies of the zodiacal light. This will enable us to probe the origins and structure of the zodiacal cloud and improve zodiacal light models for future missions such as JWST, SphereX, Euclid, and Roman. In this work, we confirm that stray-light contamination from earthshine impacts the intensity of the sky from LEO, and as a result, this excess diffuse light may help explain some of the observational discrepancies between direct EBL measurements from LEO and estimates derived by galaxy counts. Ultimately, furthering this work with SKYSURF will aid in making progress toward secure measurements of the EBL.

We would like to thank Prof. Rogier Windhorst, Dr. Timothy Carleton, Rosalia O’Brien, and the entire SKYSURF collaboration for their continued support of this work. We would also like to thank the Huntsman Telescope team for their support and guidance. We would like to thank the staff at the HLA for their help sourcing the HST ACS data used in this study.
work and the staff at NORAD for compiling HST ephemeris data for this work on special request. We would also like to thank Dr. Grant Matthews from the CERES collaboration for his guidance using CERES data and for providing the CERES throughput tables for this work. Finally, we would like to thank Arvind Hughes for his helpful comments, which led to the use of XGBoost for this work. This work is made possible by observations made with the NASA/ESA Hubble Space Telescope and obtained from the Hubble Legacy Archive, which is a collaboration between the Space Telescope Science Institute (STScI/NASA), the Space Telescope European Coordinating Facility (ST-ECF/ESAC/ESA), and the Canadian Astronomy Data Centre (CADC/NRC/CSA). S.C. and L. S. acknowledge support from an Australian Research Council Discovery Project grant (DP190102448).

Software: Skippy (Caddy et al. 2022), Gunagala (Horton et al. 2022), XGBoost (Chen et al. 2016), Astropy (The Astropy Collaboration, 2018), PySynphot (The Space Telescope Science Institute, 2015)

Appendix

In this section we present supplementary information in Table 1 describing the StarView and HST FITS Header keywords used in this work.

| Table 1 Description of StarView Keywords and HST FITS Header Keywords Used in This Work |
|-----------------------------------------------|
| Data Start Time = The start time and date of the exposure |
| Data End Time = The end time and date of the exposure |
| Target Name = Proposers target name |
| EXPTIME = Exposure time (s) |
| FILTER1 = Element selected from filter wheel 1 |
| FILTER2 = Element selected from filter wheel 2 |
| GLAT_REF = Galactic latitude of the target (deg) (J2000) |
| GLON_REF = Galactic longitude of the target (deg) (J2000) |
| PHOTMODE = Combination of (INSTRUMENT) + (CAMERA) + (FILTER) |
| PHOTFLAM = Pivot wavelength of the photomode (Å) |
| PHOTFLAM = Inverse sensitivity (erg cm⁻²/Å/DN) |
| PHOTZPT = ST magnitude system zero-point |
| PHOTBW = Root Mean Square bandwidth of the photomode (Å) |
| MDRIZSKY = Astrodizzle median sigma clipped sky surface brightness (DN/pix) |
| SUNANGLE = Angle between Sun and V1 axis (deg) |
| SUN_ALT = Altitude of the Sun above earth’s limb (deg) |
| LOS_MOON = Angle between the Moon and V1 axis (deg) |
| ALTITUDE = The average altitude of the telescope (km) |

References

Aldering, G. 2002, SNAP sky background at the north ecliptic pole, Lawrence Berkeley National Laboratory, Berkeley, CA, doi:10.2172/42543
Baggett, S., & Anderson, J. 2012, WFC3/UVIS Sky Backgrounds, STScI ISR WFC3 2012-12
Bernstein, R. A. 2007, ApJ, 666, 663
Biretta, J., Van Orsow, D., Sparks, W., Reinhart, M., & Vick, A. 2003, ACS Background Light vs. Bright Earth Limb Angle, STScI ISR ACS 2003-05
Borlaff, A. S., Gómez-Alvarez, P., Altieri, B., et al. 2021, A&A, 621, 133
Borucki, W. J., Koch, D., Basri, G., et al. 2010, Sci, 327, 977
Boughn, S. P., & Kuhn, J. R. 1986, ApJ, 309, 33
Brammer, G. Pirzkal, N., McCullough, P., & MacKenty, J. 2014, Time-varying Excess Earth-glowBackgrounds in the WFC3/IR Channel, STScI ISR WFC3 2014-03
Buffington, A., Bisi, M. M., Clover, J. M., et al. 2016, Icar, 272, 88
Caddy, S. E. 2019, Thesis, Macquarie Univ.
Caddy, S. E. 2022, Skippy v1.0.0, Zenodo., doi:10.5281/zenodo.6793035
Carleton, T., Windorsh, R. A., O’Brien, R., et al. 2022, arXiv:2205.06347
Chen, T., & Guestrin, C. 2016, in Proc. of the ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (New York: ACM), 785
Cooory, A. 2016, KOSOS, 3, 150555
Dickinson, M., Giavalisco, M. & GOODS Team 2006, The Mass of Galaxies at Low and High Redshift, ed. R. Bender & E. G. Renzini (Berlin: Springer), 324
Doelling, D. R., Loeb, N. G., Keys, D. F., et al. 2013, JatOT, 30, 1072
Driver, S. P., Andrews, S. K., Davies, L. J., et al. 2016, AJ, 827, 108
Giavalisco, M., Sahu, K., & Bohlin, R. C. 2002, New Estimates of the Sky Background for the HST Exposure Time Calculator, STScI ISR WFC3 2002-12
Hakk, W., Dencheva, N., Sontag, C., et al. 2019, DrizzlePac Documentation (Baltimore, MD: STScI)
Holl, A., & Deul, E. 1992, in Proc. of the ESO Conf. and Workshop 43, Astronomy from Large Databases, II, 43, ed. A. Heck & F. Murtagh (Hagueanu: ESO), 75
Horton, A., et al. 2022, Gunagala v1.0.0, Zenodo, doi:10.5281/zenodo.6796532
Jorgensen, J. L., Benn, M., Connerney, J. E. P., et al. 2021, JGRE, 67, 2017
Kawara, K., Matsuoka, Y., Sano, K., et al. 2017, PASJ, 69, 31
Kelsall, T., Weiland, J., Franz, B., et al. 1998, ApJ, 508, 44
Korngut, B., Bock, J. J., Akeson, R., et al. 2018, Proc. SPIE, 10698, 64
Korngut, P. M., Renbarger, T., Arai, T., et al. 2013, ApJS, 207, 34
Korngut, P. M., Kim, M. G., Ariai, T. et al. 2022, ApJ, 926, 133
Krick, J. E., Glaccum, W. J., Carey, S. J., et al. 2012, ApJ, 754, 53
Lauer, T. R., Postman, M., Spencer, J. R., et al. 2022, ApJS, 927, L8
Leinert, C., Bowyer, S., Haikala, L. K., et al. 1998, A&AS, 127, 1
Lim, P. L., Diaz, R. I., & Laidler, V. 2015, PySynphot User’s Guide, (Baltimore, MD: STScI)
Lombardo, S., Muslimov, E., Valls-Gabaud, D., et al. 2019, SPIE, 11180, 1106
Luger, R., Bedell, M., Vanderspek, R., & Burke, C. J. 2019, arXiv:1903.12182
Matsuoka, Y., Ienaka, N., Kawara, K., & Oyabu, S. 2011, ApJ, 736, 119
Mattila, K. 1976, A&A, 47, 77
Mattila, K. 1990, IAUS, 139, 257
Mattila, K., & Väisänen, P. 2019, ConPh, 60, 23
Mattila, K., Väisänen, P., Lehtinen, K., von Appen-Schnur, G., & Leinert, C. 2017, INRAS, 470, 2152
Murthy, J. 2014, ApJS, 240, 165
Price-Whelan, A. M. 2018, AJ, 156, 123
Shaw, D., Reinhart, M., & Wilson, J. 1998, Scattered Light from the Earth Limb Measured with the STIS CCD, STScI ISR. STIS 98–21
Sirianni, M., Jee, M. J., Benitez, N., et al. 2005, PASP, 117, 409
Spinrad, H., & Stone, R. P. S. 1978, ApJ, 226, 609
Windorsh, R. A., Carleton, T., O’Brien, R., et al. 2022, arXiv:2205.06214
Woolf, N. J., Smith, P. S., Traub, W. A., & Jucks, K. W. 2002, ApJ, 574, 430
Wright, E. L. 2001, ApJ, 553, 538
Zemcov, M., Inmel, P., Nguyen, C., et al. 2017, NatCo, 8, 15003

ORCID iDs

Sarah E. Caddy @ https://orcid.org/0000-0001-6990-7792
Lee R. Spitler @ https://orcid.org/0000-0001-5185-9876
Simon C. Ellis @ https://orcid.org/0000-0002-0742-379X