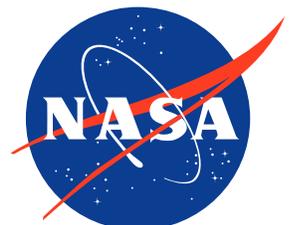**Commercial Satellite Data Acquisition Program**

BlackSky Radiometric & Geometric Quality Assessment Report

Goddard Space Flight Center
Greenbelt, MD

Commercial Satellite Data Acquisition Program BlackSky Radiometric & Geometric Quality Assessment Report

Signature/Approval Page

Approval by:

_____   _____
Melissa Yang Martin   Date
Commercial Satellite Data Acquisition Program Manager
Earth Science Division
Headquarters/NASA

Accepted by:

_____   _____
Dana Ostrenga   Date
Commercial Satellite Data Acquisition Project Manager
Earth Science Division
GSFC/NASA

Preface

This document is under CSDA Project configuration control. Once this document is approved, CSDA approved changes are handled in accordance with Class I and Class II change control requirements described in the CSDA Configuration Management Procedures based on NASA standard configuration practices, and changes to this document shall be made by document change notice (DCN), documented in the Change History Log or by complete revision.

Abstract

The evaluation summarized in this report was conducted by subject matter experts (SMEs) funded by NASA's Commercial Satellite Data Acquisition (CSDA) Program. The SMEs evaluated the radiometric and geometric quality of BlackSky data for the NASA Earth science research and applications community. The results of the evaluation help to inform NASA program management on the quality of the data for NASA science.

Cover Art: Cover art is AI generated graphic using Microsoft Copilot Designer using term "commercial satellite constellation Earth observation across Atlantic AND Northern Hemisphere AND digital downlink"

Authored and prepared by

Frederick Policelli
CSDA Project Scientist
National Aeronautics and Space Administration

Jaime Nickeson
CSDA Technical Science Coordinator
Science Systems and Applications Inc
National Aeronautics and Space Administration

Guoqing (Gary) Lin
Geolocation Subject Matter Expert
National Aeronautics and Space Administration

Alana G. Semple
Geolocation Subject Matter Expert
Science Systems and Applications, Inc.
National Aeronautics and Space Administration

Chuanmin Hu
BlackSky Evaluation Team Member
Texas A&M University

Brian Barnes
BlackSky Evaluation Team Member
University of South Florida

Christopher Neigh
BlackSky Evaluation Team Lead and
Electro/Optical Subject Matter Expert
National Aeronautics and Space Administration

Alexei Lyapustin
Radiometric Calibration Subject Matter Expert
National Aeronautics and Space Administration

Eric Vermote
Radiometric Calibration Subject Matter Expert
National Aeronautics and Space Administration

Bin Tan
Geolocation Subject Matter Expert
Science Systems and Applications, Inc.
National Aeronautics and Space Administration

Myungje Choi
Radiometric Calibration Subject Matter Expert
National Aeronautics and Space Administration

Nima Pahlevan
BlackSky Evaluation Team Member
Science Systems and Applications Inc
National Aeronautics and Space Administration

Change History Log

Revision	Effective Date	Description of Changes
1.0	03/22/2024	First Draft Completed

Table of Contents

List of Figures

List of Tables

Acronyms & Abbreviations

AOD	Aerosol Optical Depth
APA	Absolute Positional Accuracy
ARD	Analysis Ready Data
ATBD	Algorithm Theoretical Basis Document
BBR	Band-to-Band Registration
BRDF	Bi-directional Reflectance Distribution Factor
CE90	Circular Error at the 90th percentile
CF	Climate & Forecast (Metadata Convention)
CWV	Column Water Vapor
CEOS	Committee on Earth Observation Satellites
DEM	Digital Elevation Model
DESIS	DLR Earth Sensing Imaging Spectroradiometer
DN	Digital Number
DLR	Deutsches Zentrum für Luft- und Raumfahrt (German Aerospace Center)
DOI	Digital Object Identifier
EDAP	Earthnet Data Assessment Pilot
EO	Earth Observation
ESA	European Space Agency
ESF	Edge Spread Function
FAIR	Findable, Accessible, Interoperable and Reusable
FRM	Fiducial Reference Measurement
FWHM	Full-width Half-maximum
GRD	Ground Resolved Distance
GUM	Guide to the Expression of Uncertainty in Measurement
INSPIRE	Infrastructure for Spatial Information in Europe
ISS	International Space Station
L1	Level 1
L2	Level 2
LSF	Line Spread Function
MAIAC	Multi-Angle Implementation of Atmospheric Correction
MERRA	Modern Era Retrospective-analysis for Research and Applications
MODIS	MODerate-resolution Imaging Spectroradiometer
MSI	Multispectral Instrument (on Sentinel-2 platform)
MTF	Modulation Transfer Function
NAIP	National Agricultural Imagery Program
NASA	National Aeronautics and Space Agency
NOAA	National Oceanic and Atmospheric Administration
NPL	National Physical Laboratory, UK
OLI	Operational Land Imager (instrument on Landsat 8)

PCC	Pearson Cross Correlation
PSF	Point Spread Function
PUG	Product User Guide
PUM	Product User Manual
QA	Quality Assessment
QA4EO	Quality Assurance Framework for Earth Observation
QA4ECV	Quality Assurance Framework for Essential Climate Variables
RMSE	Root Mean Squared Error
RSR	Radiometric Spectral Response
S2	Sentinel-2
SBAF	Spectral Band Adjustment Factors
SI	Système International (International System of Units)
SNPP	Suomi-National Polar-orbiting Partnership
SSR	Sensor Spatial Response
SZA	Solar Zenith Angle
SR	Surface Reflectance
SRTM	Shuttle Radar Topography Mission
TOA	Top of Atmosphere
URL	Universal Resource Locator
VIIRS	Visible-Infrared Imaging Radiometer Suite
VZA	View Zenith Angle
WLEF	Telecommunications tower in northern Wisconsin

Executive Summary

The CSDA Program was established to identify, evaluate, and acquire data from commercial sources that support NASA's Earth science research and application goals. NASA's Earth Science Division (ESD) recognizes the potential impact commercial satellite constellations may have in encouraging/enabling efficient approaches to advancing Earth System Science and applications development for societal benefit. Commercially acquired data may also provide a cost-effective means to augment and/or complement the suite of Earth observations acquired by NASA and other U.S. government agencies and those by international partners and agencies.

In this report, CSDA provides an evaluation of the quality of data provided by the BlackSky satellite constellation for advancing NASA's Earth system science research and applications. This evaluation of BlackSky radiometric and geometric performance was carried out by NASA subject matter experts (SMEs) that were enlisted to evaluate the fundamental quality of the BlackSky data in these domains. BlackSky provides very high-resolution commercial images in three spectral bands (Blue, Green and Red) and a panchromatic band from a satellite constellation. The provided imagery consists of uncalibrated digital numbers for each band. Lack of calibration would limit science use of the BlackSky data, including retrieval of geophysical properties such as land surface reflectance or atmospheric aerosol properties. Details about the assessment performed by a group of selected principal investigators on the utility of BlackSky data for NASA science is available in a separate report, *Commercial Satellite Data Acquisition Program BlackSky Principal Investigator Evaluation Report*.

To assess the radiometric properties, the BlackSky "Proto" product was analyzed for scenes acquired over the Libya-4 desert calibration site and the WLEF telecommunication tower located in a forested region in north-central Wisconsin near Park Falls. This product precedes the dynamic range adjustment (DRA) product, referred to by BlackSky as "Anthro", which is the default processing applied to their delivered products.

Analysis performed by one of the calibration teams found that the Proto data shows many artifacts in the radiometric performance (negative and variable dark current), a large variability of the calibration coefficient between sensors and instability of calibration for the same sensor. This may be related to the demosaicing of BlackSky's extended Bayer color mosaic filter output from other nearby bands in an image sharpening process, and would be consistent with the fact that both the calibration slope and the dark current in green and red bands are correlated.

Another calibration team addressed the lack of calibrated data by developing and applying a vicarious calibration approach based on multi-angle implementation of atmospheric correction (MAIAC) processing. Their method utilizes the ancillary MAIAC atmospheric and surface reflectance properties from the well-calibrated MODerate-resolution Imaging Spectroradiometer (MODIS) sensor and the German Aerospace Center's (DLR's) Earth Sensing Imaging Spectroradiometer (DESIS) hyperspectral sensor measurements to compute spectral band adjustment factors (SBAFs) as well as radiative transfer to simulate top-of-atmosphere (TOA) radiance for the BlackSky imagery. The comparison of simulated TOA radiance and BlackSky-measured DN values provides the calibration gain and offset.

There is no signal-to-noise ratio (SNR) specification provided by BlackSky. The SNR analysis conducted over the Libya-4 desert site aligns closely with SNR values observed in other VHR satellites, such as Maxar's WorldView-2 and Earth Observing One (EO-1) Hyperion.

A geolocation team evaluated the BlackSky image spatial performance and absolute geolocation accuracy over 26 globally distributed locations, as well as band-to-band registration and temporal stability of BlackSky geolocation accuracy at 3 USA cities.

BlackSky does not make claims for spatial response beyond a 0.8 – 1.1 m pixel size. The geolocation analysis team constructed the BlackSky sensor spatial response for 12 of their 16 sensors. The average spatial response for all sensors and all bands is 2.48 pixels in the row direction and 2.63 pixels in the column direction in terms of full width half maximum of the constructed line spread functions. Modulation transfer function (MTF) at Nyquist frequency is 0.021 for all bands in the row direction, and 0.012 in the column direction. We make a cautionary note about the radiometric performance since they demosaic their extended Bayer's color mosaic filter output not only from other pixels but also from other bands nearby.

The absolute positional accuracy could not be assessed due to lack of available absolute references. Instead, relative positional accuracy was used as a proxy. We found the positional accuracy overall does not meet BlackSky's claimed accuracy of 10 m circular error at the 90th percentile (CE90) in the USA and Australia, or the 20 m CE90 elsewhere. However, their self-consistencies are within their specifications in terms of CE90-demean (CE90 relative to the mean error). This accounts for possible biases in using different reference points from those that BlackSky uses for georeferencing and the Maxar WorldView-2 images that we used for assessment. Geolocation accuracy by city varies by location, with some performing better than specification and some performing much worse. Belo Horizonte, Brazil and Sapporo, Japan are of special note. We found the BlackSky images in these areas to have offsets of 40 - 60 m.

Overall accuracy for USA and Australia locations compared to other locations:

Regions	# Images	CE90	CE90-demean
USA & Australia	53	23.6	9.0
Other global locations	120	23.2	16.4

Temporal stability of geolocation accuracy for one year was assessed for three USA cities. BlackSky does not make separate claims for temporal stability beyond the geolocation accuracy claim of 10 m CE90 in the USA. One location, in Kansas City, MO, exceeded the claimed CE90 threshold. All locations evaluated would not be suitable for time series analyses as delivered, based on the average image offsets shown in the table below. Users will have to either aggregate pixels or register their imagery to a selected "golden" reference image to perform a time series analysis with BlackSky data for a specific site.

City	# Images	Range of E-W offsets (m)	Range of N-S offsets (m)
Boston, MA	10	3.1 - 11.3	0.5 - 14.1
Kansas City, MO	12	2.3 - 26.6	0.1 - 11.0
San Diego, CA	13	0.1 - 7.4	0.1 - 7.4

1 Cal/Val Maturity Matrices

1.1 Summary Cal/Val Maturity Matrix

Data Provider Documentation Review		
Product Information	**Metrology**	**Product Generation**
Product Details	Radiometric Calibration & Characterization 🔒	Radiometric Calibration Algorithm 🔒
Availability & Accessibility	Geometric Calibration & Characterization 🔒	Geometric Processing 🔒
Product Format, Flags & Metadata	Metrological Traceability Documentation 🔒	Mission Specific Processing 🔒
User Documentation	Uncertainty Characterization 🔒	
	Ancillary Data 🔒	

Validation Summary
Radiometric Validation Method 🔒
Radiometric Validation Results Compliance 🔒
Geometric Validation Method 🔒
Geometric Validation Results Compliance

Key
Not Assessed
Not Assessable
Basic
Good
Excellent
Ideal

🔒 Not Public

Figure 1. Summary Cal/Val Maturity Matrix

1.2 Detailed Validation Maturity Matrix

Figure 2. Detailed Validation Maturity Matrix for the optical domain, showing the Validation Summary column from the Summary Cal/Val Maturity Matrix

2 Data Provider Documentation Review

2.1 Product Information

Product Details	
Grade: Basic	
Justification	Only a minimum amount of information is provided. The information provided lacks details on radiometric calibration aspects such as dark current, conversion factors, etc. Furthermore, the Anthro product involves adjusting the image's color balance and contrasts through a dynamic range adjustment, but this process is not thoroughly explained.
Product Name	BlackSky Anthro
Sensor Name	Global-1, 2, 4, 8, 9, 12, 13, 14, 15, 16, 17, 20 Two generations of sensors, with Global 1 and 2 being from generation 1.

Sensor Type	Framing Camera RGBW Color Filter Array
Mission Type	Optical Constellation
Mission Orbit	Mid-inclination orbits
Product Version Number	
Product ID	<SensorName>-<SatelliteID>-<AcquisitionDate>-<AcquisitionTime>-< ImageID> Example: BSG-104-20191025-094224-528654
Processing level of product	Anthro is their default processing level. Proto level product was evaluated. Both can be ordered as either Orthorectified or Georeferenced.
Measured Quantity Name	Digital Number
Measured Quantity Units	No units
Stated Measurement Quality	Information on the radiometric measurement quality is unavailable.
Spatial Resolution	(eNIIRS 4.0 for Global 07-09, Global 12-17) 0.83 m – 0.94 m, (eNIIRS 3.5 for Global 01, 02, 04) 1.1 m – 1.3 m
Spatial Coverage	The spatial coverage depends on the mission. (eNIIRS 4.0) 21-25 km^2 at nadir; 26-33 km^2 at off-nadir; +53.5°N to -53.5°S (eNIIRS 3.5) 24-36 km^2 at nadir; 40-59 km^2 at off-nadir; +85.3°N to -85.3°S
Temporal Resolution	On demand image (daily revisit with the constellation) depending on the location.
Temporal Coverage	On demand image (daily revisit with the constellation) depending on the location.
Point of Contact	customersuccess@blacksky.com
Product locator (DOI/URL)	N/A
Conditions for access and use	All data used in this evaluation were purchased by CSDA under U.S. Government-wide license.
Limitations on public access	License can be found at the CSDA website.
Product Abstract	N/A

Availability & Accessibility

Grade: Good	
Justification	Good access to the validation site through tasking. The "Anthro" dataset is available through the official BlackSky data Platform. However, the "Proto" dataset is unavailable from the site because "Proto" is not their official product. It is only available by individual request by contacting the provider.
Compliant with FAIR principles	No
Data Management Plan	Unknown.
Availability Status	Only available by individual request.

Product Format, Flags and Metadata

Grade: Good	
Justification	Data are provided by formats of GeoTiff and NITF, not HDF or netCDF. GeoTiff is sufficient for analyzing the imagery.
Product File Format	GeoTiff, NITF

Metadata Conventions	(Example of info in metadata for "Proto")
	"collection_conditions": {
	    "cats_angle_degrees": 22.66637980721512,
	    "cloud_cover": 0,
	    "gmt_hour": 11.288844993888889,
	    "ground_range_m": 130944.6341207759,
	    "ground_sampling_distance_m": 1.0374945272355385,
	    "local_hour": 12.848178327222223,
	    "off_nadir_angle_deg": 15.640897319871081,
	    "off_nadir_angle_degrees": 15.640897319871081,
	    "off_velocity_azimuth_deg": 80.6681859723497,
	    "off_velocity_azimuth_degrees": 80.6681859723497,
	    "predicted_niirs": 3.821949055857863,
	    "satellite_azimuth_degrees": 322.491,
	    "satellite_elevation_degrees": 71.73,
	    "slant_range_m": 485282.9821065598,
	    "sun_azimuth_degrees": 237.2,
	    "sun_elevation_degrees": 74.71,
	    "time": "2022-05-18T11:17:19.841978+00:00"
	    },
	    "target_info": {
	        "auto_assign_elevation": true,

"baseline_priority": null, "elevation_meters": 128.0, "ground_reflectance": 0.41141027, "latitude_degrees": 28.55, "longitude_degrees": 23.39, "name": null, "target_id": 1650987, "target_status": "active", "target_type": "platform_orders" }
Analysis Ready Data?	No.

User Documentation		
Grade: Basic		
Justification	The "Anthro" processing enhances the image's contrast (scaled DN), but the description of enhancement is not available.	
Document	*Reference*	*QA4ECV Compliant*
Product User Guide	BlackSky Imagery Product Guide (v3.7)	No
ATBD	Not available to users	No

2.2 Metrology

Sensor Calibration & Characterization	
Grade: Not Assessable	
Justification	BlackSky data are not calibrated. The product is "Digital Number (DN)" only.
References	• BlackSky Imagery Product Guide (v3.7)

Geometric Calibration & Characterization	
Grade: Basic	
Justification	Less than half of the evaluated locations meet BlackSky's claimed geolocation accuracy. Over half, but not all, of these locations have good Temporal Stability. Their image footprint size is over 2 pixels. BBR is satisfactory.
References	• *See sections 4.1.2, 4.2.2, and 4.4.2*

Metrological Traceability Documentation	
Grade: Not Assessable	
Justification	Not available.
References	•

Uncertainty Characterization	
Grade: Not Assessable	
Justification	Not available.
References	•

2.3 Product Generation

Radiometric Calibration Algorithm	
Grade: Not Assessable	
Justification	Not available.
References	•

Geometric Processing	
Grade: Good	
Justification	Geolocation processing follows the standard reference image method. Reference images and orthorectification methods are clearly stated. SSR and BBR processing is not public.
References	• BlackSky Imagery Product Guide (v3.7) • *ESA-NASA, 2021*

Mission-Specific Processing	
Grade: Not Assessable.	
Cloud and Water Masks	
Justification	Cloud percentage and water masks that BlackSky provided were found to be inaccurate. Opaque clouds and cloud shadows were identified in the cloud mask, but transparent/hazy clouds were not. The processing steps for the creation of the masks were not documented.
References	

3   Detailed Validation – Radiometric

3.1   Absolute Calibration

3.1.1   Method

The Libya-4 desert calibration site, centered at 28.55°N, 23.39°E, was selected for the radiometric calibration. The Libya-4 site is a Committee on Earth Observation Satellites (CEOS)-recommended pseudo-invariant desert calibration site due to its high radiometric stability, low aerosol loading and spatial extent (https://calval.cr.usgs.gov/apps/libya-4). The Multi-angle Implementation of Atmospheric Correction (MAIAC) [Lyapustin et al., 2018] group at GSFC has been using this site successfully for the vicarious calibration of the MODerate Resolution Imaging Spectroradiometer (MODIS) [Lyapustin et al., 2014], Visible Infrared Imaging Radiometer Suite (VIIRS) [Lyapustin et al., 2023], and the Maxar sensors [Choi et al., 2024]. The MAIAC-based technique enables evaluation of the long-term calibration trend and cross-calibration of sensors to the MODIS Aqua sensor which is presently considered to be a calibration standard. Here, we analyze the absolute calibration of BlackSky sensors as a long-term record of data is not available.

BlackSky processes their data through up to three processing levels that include Proto, Anthro and Ortho. Proto is the most basic level and all BlackSky images are processed to at least this level. Proto processing reduces sensor-to-sensor variation by correcting for defects or artifacts, improves the geolocation and assesses the cloud score of the imagery. However, Proto data products are not the standard commercial product that BlackSky distributes, and these data were not available to CSDA investigators at the outset of the evaluation. The BlackSky production team worked with the CSDA team to make their Proto data product available for the evaluations. The further BlackSky Anthro processing enhances the image contrast with the aforementioned dynamic range adjustment steps. The Anthro processing is the default imagery product that BlackSky produces. The Ortho processing can be applied to both Proto and Anthro products.

The Proto images are the only ones that can be used for the calibration analysis. To accommodate our request, the BlackSky team collected a total of 51 images over the Libya-4 site from May to

July 2022 and provided them as Proto data (datatype "preDRAOrthorectified") along with the view geometry and cloud cover. Figure 3 shows the acquisitions from different Global satellites in the BlackSky constellation.
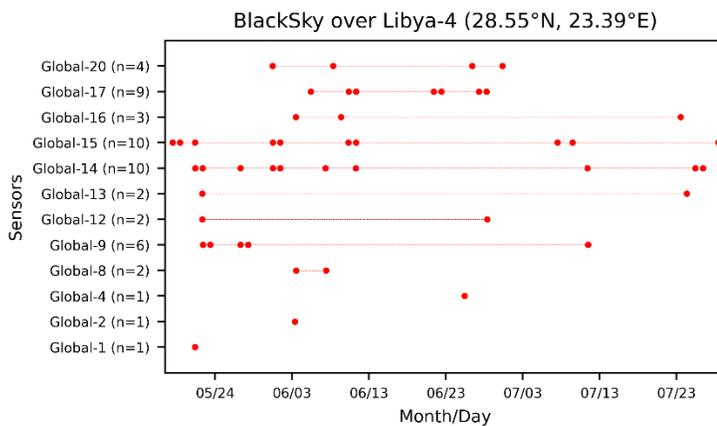
Figure 3. BlackSky acquisitions over Libya-4.

The MAIAC calibration approach (Fig. 4) assumes the standard linear relationship between the TOA radiance and the measured BlackSky DN:

$$TOA\ radiance = gain * measured\ DN + offset$$

The expected TOA radiance was computed for the geometry of BlackSky acquisitions using MERRA-2 ozone and MODIS ancillary aerosol optical depth (AOD), column water vapor (CWV), and spectral surface bidirectional reflectance distribution function (BRDF) information from MAIAC. To account for the MODIS/BlackSky spectral differences in the RGB channels, we generated the Spectral Band Adjustment Factors (SBAFs) for the surface reflectance (SR) using the hyperspectral SR data from DESIS.
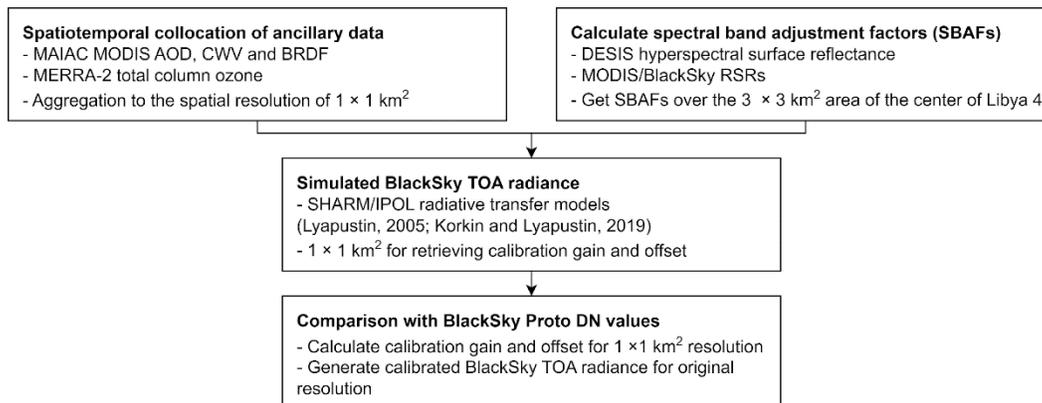
Figure 4. Diagram of BlackSky radiometric calibration analysis using MAIAC

The BlackSky relative spectral response (RSR) functions are shown in Figure 5. Here, "Gen 2.0" corresponds to Global-01, 02, and 04 satellites, and "Gen 2.1" indicates the remaining constellation sensors (i.e., Global 07-09, 12-18, and 20).
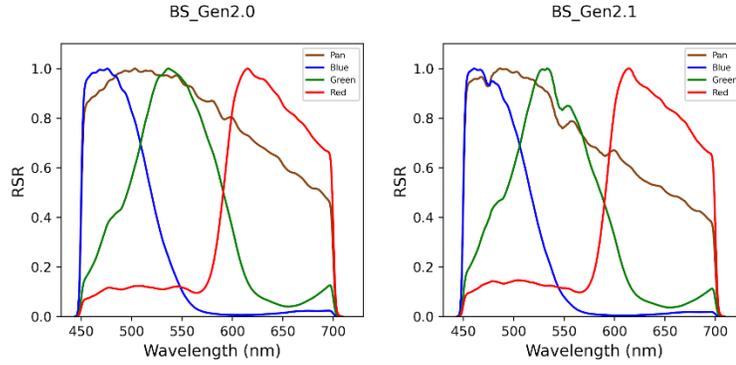
Figure 5. Relative spectral response functions of BlackSky sensors.

As mentioned above, to account for different RSRs between MODIS and BlackSky sensors, we collected a hyperspectral measurements dataset of DESIS. DESIS is the DLR (German Aerospace Center) hyperspectral sensor operating onboard the International Space Station (ISS) since 2018. DESIS measures spectral radiance reflected by Earth from 400 nm to 1000 nm with a spectral sampling of 2.55 nm and a full width of half maximum (FWHM) of about 3.5 nm, a total of 235 channels (Krutz et al., 2019; Alonso et al., 2019). The spatial resolution is 30 m with 1024 across-track pixels with swath of about 30 km. Importantly, our analysis shows a very good agreement of calibration between DESIS and MODIS Aqua [Lyapustin et al., 2023].

Hyperspectral DESIS surface reflectance was convolved with the RSRs of MODIS and BlackSky in the RGB bands as follows $\rho_{simulated} = \frac{\sum \rho_\lambda E_\lambda RSR_\lambda d\lambda}{\sum E_\lambda RSR_\lambda d\lambda}$, where ρ_λ is DESIS SR, E_λ is solar irradiance, and RSR_λ is spectral response function. The ratio of convolved MODIS and BlackSky SR values (over Libya-4) gives the SBAFs required for spectral adjustment of MODIS BRDF to simulate the BlackSky TOA radiance/reflectance. Figure 6 illustrates the idea of spectral convolution based on DESIS and computation of SBAFs, and Figure 7 shows convolved SR values in all three bands for MODIS and BlackSky.
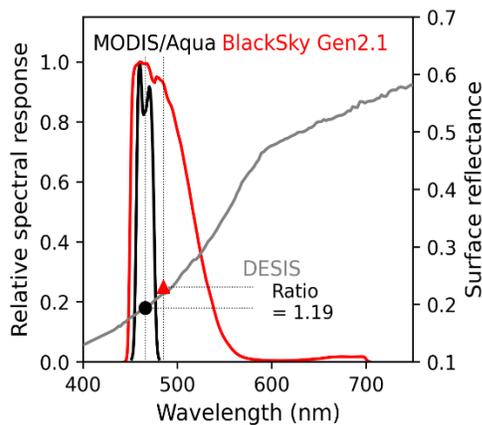
Figure 6. Relative spectral responses of MODIS B3 and BlackSky Blue bands, DESIS hyperspectral surface reflectance, and derived SBAF in the Blue band.
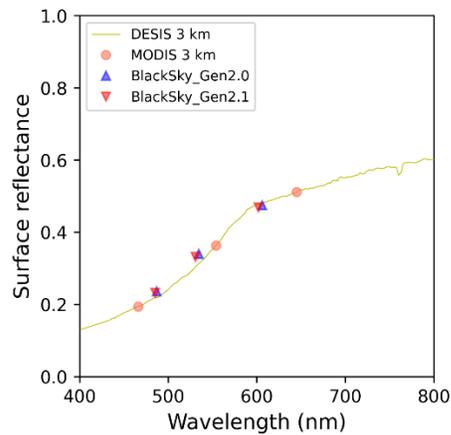
Figure 7. Surface reflectance of DESIS (yellow line) and calculated for MODIS (red circles), BlackSky Gen 2.0 (blue triangles) and Gen 2.1 (red triangles).

Similar to BlackSky, DESIS observations of Libya-4 site were collected in varying sun-view geometries. Therefore, the BRDF normalization is also conducted to normalize different geometries of measurements to a standard geometry (VZA of 0° and SZA of 20°), which can significantly reduce surface reflectance variability [Lyapustin et al., 2014]. BRDF normalization factor $(c(\lambda) = \frac{BRF\ from\ fixed\ view\ geometry\ (SZA\ 20°,\ VZA\ 0°,\ RAA\ 0°)}{BRF\ from\ various\ DESIS\ view\ geometries})$ converts surface reflectance at all different geometries $(\rho_{simulated}(\lambda))$ to the normalized geometry $(\rho^n_{simulated}(\lambda) = \rho_{simulated}(\lambda) * c(\lambda))$. The final SBAFs obtained from multiple measurements at the normalized geometry are given in Table 1.

**Table 1. Spectral Band Adjustment Factors (SBAFs) for the surface reflectance from MODIS to BlackSky over Libya-4 site.**

SBAFs	Blue / MODIS B3	Green / MODIS B4	Red / MODIS B1
BlackSky Gen2.0	1.164	0.934	0.920
BlackSky Gen2.1	1.150	0.914	0.910

As a verification of our calibration procedure, Figure 8 compares the "measured" and "simulated" BlackSky-like TOA radiance using 44 DESIS scenes over Libya-4 at aggregated 1 km resolution. The "measured" indicates reflectance values from DESIS scenes spectrally integrated to the BlackSky channels. The "simulated" data represent radiative transfer simulation of the spectral TOA reflectance using the MAIAC MODIS ancillary surface-atmosphere information in the specific view geometry of the DESIS scenes. Effectively, Figure 8 shows that DESIS, aggregated to the BlackSky bands, is properly calibrated, and the calibration technique works as expected. The different colors indicate different scenes. The X-axis of Figure 8 shows the DESIS measurements spectrally aggregated to the BlackSky channels. The Y-axis of Figure 8 gives the calculated TOA radiance in the DESIS view geometry using MAIAC MODIS ancillary AOD, CWV, and BRDF spectrally adjusted to BlackSky channels using SBAFs. The strong linear relationships with R^2 above 0.92 and close to the one-to-one line confirms the validity of our vicarious radiometric calibration method for the BlackSky.
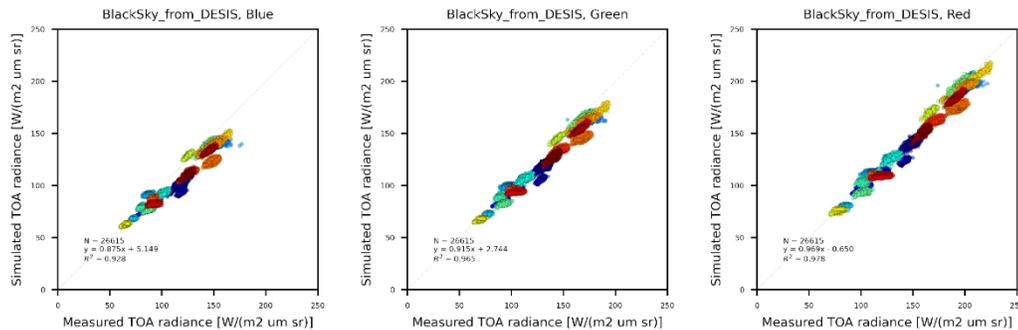
Figure 8. Comparison of "measured" (based on DESIS) and "simulated" (based on ancillary MODIS MAIAC products) BlackSky-like TOA radiance for the Blue, Green, and Red bands.

3.1.2 Results Compliance

A direct comparison between measured BlackSky DN and simulated TOA radiance, as described above, shows highly scattered features, as in Figure 9 when all scenes are composited. However, the linear relationship can be found from pixels in a single scene, with a correlation coefficient R up to 0.77 (Fig. 10). Here, the simulated TOA radiance is normalized, assuming that incoming TOA solar irradiance is 1.
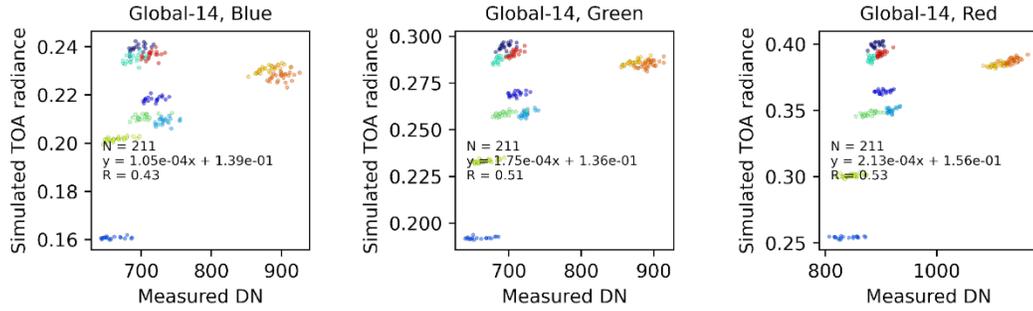
Figure 9. Comparison of BlackSky measured DN and simulated TOA radiance (using MAIAC MODIS ancillary dataset) in RGB bands. Colors indicate ten different Global-14 scenes.
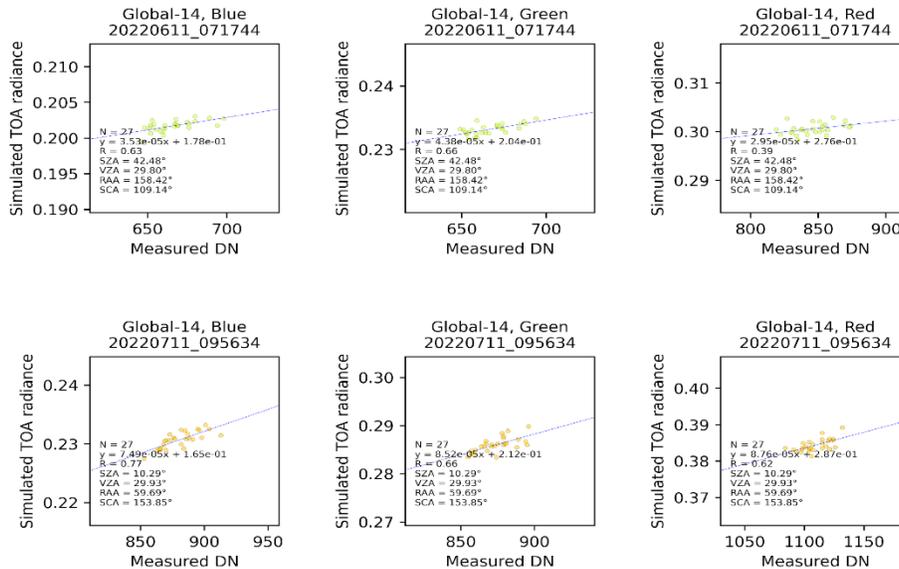
Figure 10. Comparison of measured DN and simulated TOA radiance for BlackSky bands from two scenes: June 11, 2022 (top panels) and July 11, 2022 (bottom panels).

Based on these results and our linear calibration model ($Simulated\ TOA\ radiance = gain * measured\ DN + offset$), we assume that the common gain exists per sensor and channel, and the offset is scene-dependent. Aggregating gains of ten Global-14 scenes provides a common gain and its uncertainty shown in Figure 11. The blue band gain of 4.5e-05 indicates that 100 DN units give 0.0045 in the normalized TOA radiance. The average, standard deviation and relative uncertainty of the gain across collected scenes are summarized for each satellite and sensor in Figure 12. The scales in the Green and Red bands are similar but greater than those in the Blue band for most satellites.
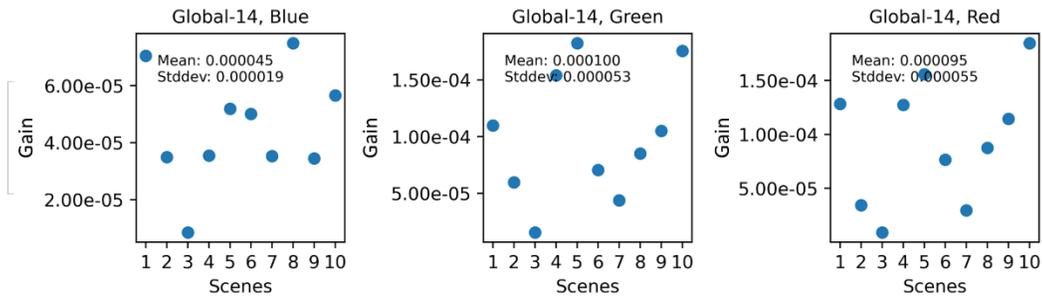
Figure 11. Gain for 10 Global-14 scenes assuming linear regression between measured DN and TOA radiance.
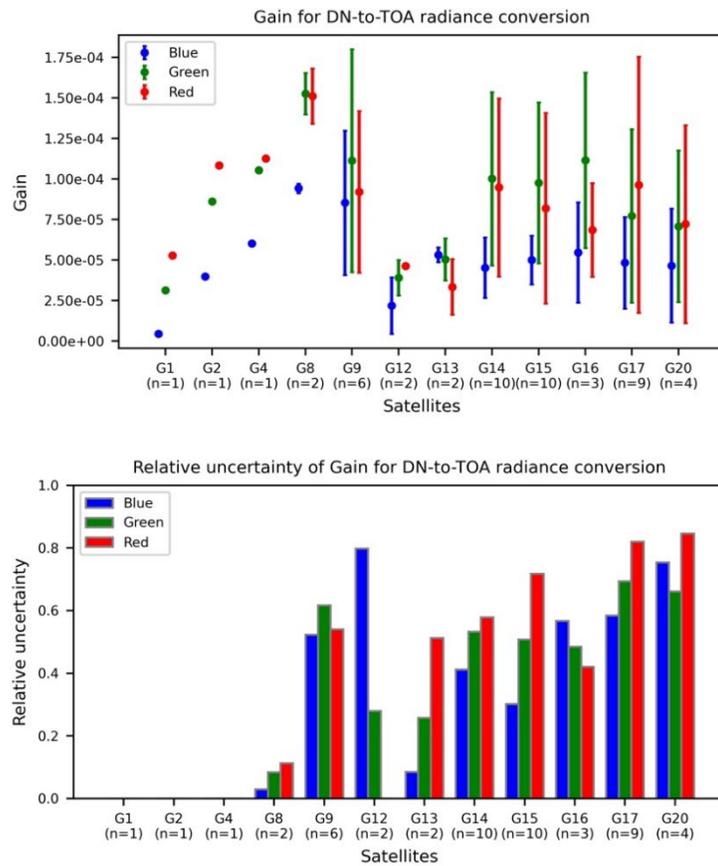
Figure 12. Average, standard deviation, and the relative uncertainty of gain for radiometric conversion for selected BlackSky satellites and sensors. Here, the results are for the normalized TOA radiance assuming the incoming solar irradiance of 1.

The relative uncertainty of gains, defined as the ratio of standard deviation and mean, varies from 3% to 85%. The gain factors converting DN to TOA radiance with the unit of $W/m^2/\mu m/sr$ are presented in Figure 13.
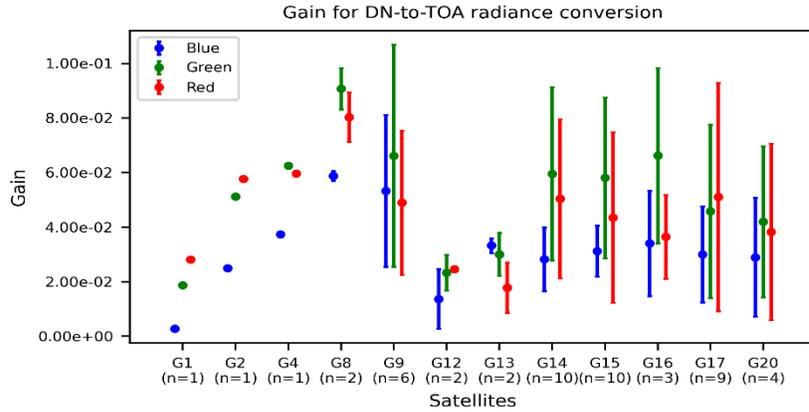
Figure 13. Same as the top panel of Figure 12 except for TOA radiance in W/m²/µm/sr units.

Applying the gains to the Global-14 measured DNs with scene-specific offsets results in a calibrated TOA radiance that agrees well with the simulated radiance as shown in Figure 14 and Figure 15. Such comparison confirms the MAIAC-based vicarious radiometric calibration of the BlackSky data for each satellite/band over the Libya-4 site. However, the short time record and the scene-dependent offsets do not allow us to assess the temporal stability of the radiometric calibration.
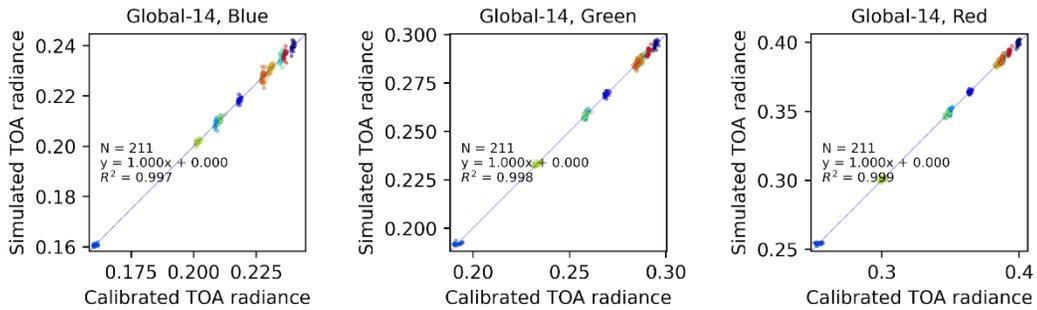
Figure 14. Comparison of calibrated BlackSky TOA radiance and simulated TOA radiance for ten Global-14 scenes.
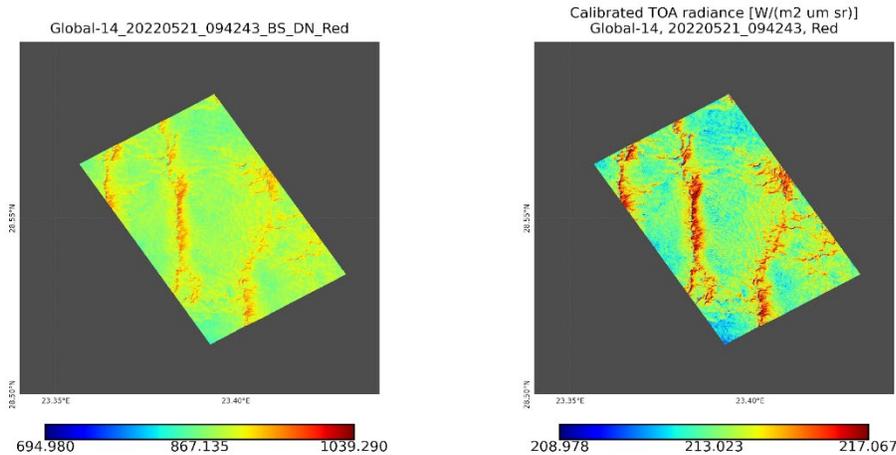
Figure 15. Measured Proto DN values (left) and calibrated TOA radiance (right) of Global-14 5/21/2023 measurement over Libya-4.

The above calibration analysis was based on the assumption that the "common gain exists per sensor and channel, and the offset is scene-dependent". Our analysis and high variability of the gain factor between different acquisitions, up to 50-100%, clearly show that this is not the case (e.g., Fig. 13). However, using the described MODIS-MAIAC technique, based on ancillary surface-atmosphere data, the BlackSky imagery could be calibrated on the scene-by-scene basis for images with spatial contrast.

Analysis performed leads to the following conclusions:

1. There is a significant variability of gain and offset values in different acquisitions of BlackSky data for all sensors. During a dedicated meeting, the BlackSky team was not able to provide the answer as to the cause of such variability.
2. The MAIAC-based technique can be used to calibrate the individual BlackSky scenes provided that the hyperspectral surface reflectance data (e.g., from DESIS), required to derive SBAFs, are available.
3. Without reliable calibration, potential science analysis of BlackSky data would be limited.

A separate study of the BlackSky radiometry was conducted by a second calibration team. This study compared BlackSky Proto products to camera measurements collected from a telecommunications (WLEF) tower located in a heavily forested region in northern Wisconsin. The camera system, called CAMSIS, collected surface reflectance measurements in the red, green and blue that were used to simulate top of the atmosphere reflectance using MODIS ancillary data (aerosol optical thickness, water vapor, ozone).

The same method was used for Sentinel-2 (S2) calibration, Figure 16 shows the result of the comparison (at bottom of the atmosphere). The agreement is excellent given the uncertainties on both sides.
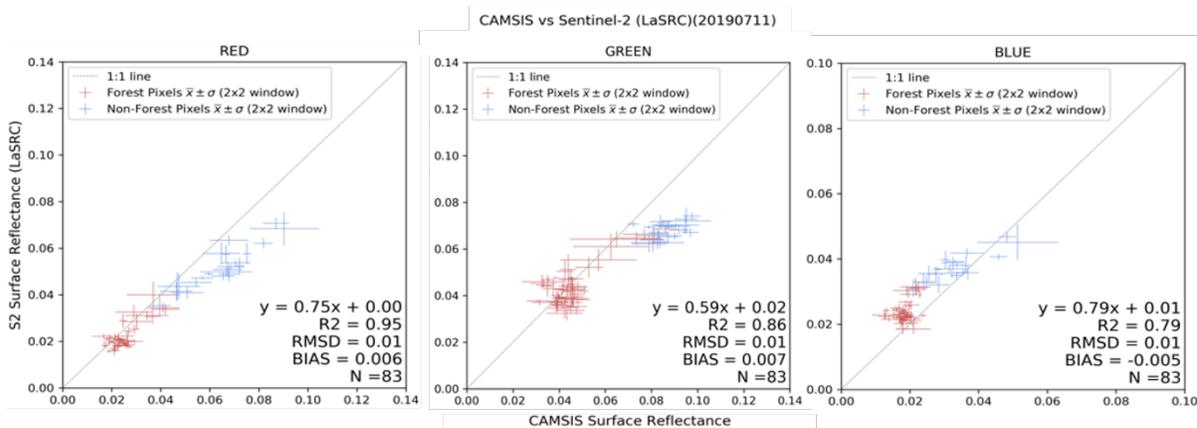
Figure 16. Comparison of the CAMSIS reflectance in Red, Green, Blue with the Sentinel 2 Land Surface Reflectance (LaSRC) derived reflectance.

The results with BlackSky shows poor correlation between the BlackSky Anthro product and CAMSIS (Fig. 17) and better correlation (almost a factor 2) with the BlackSky Proto product (Fig.18).
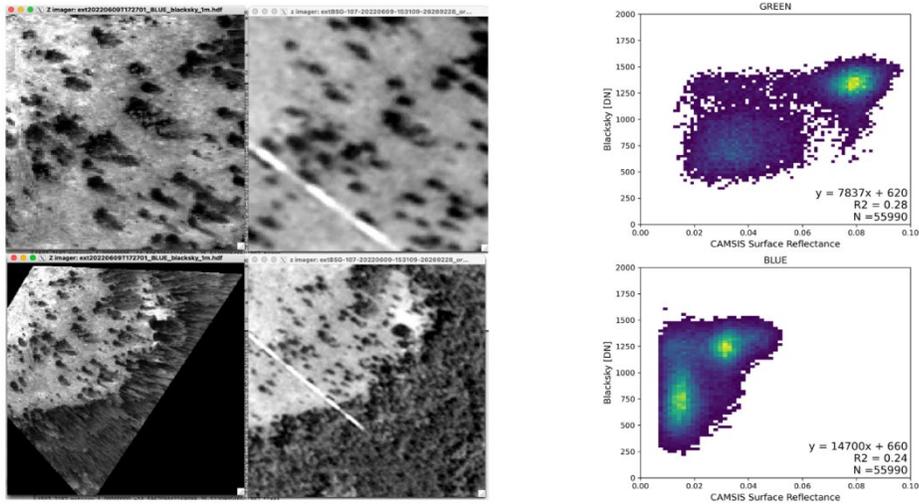
Figure 17. Comparison of BlackSky Anthro product with CAMSIS surface reflectance at WLEF site.
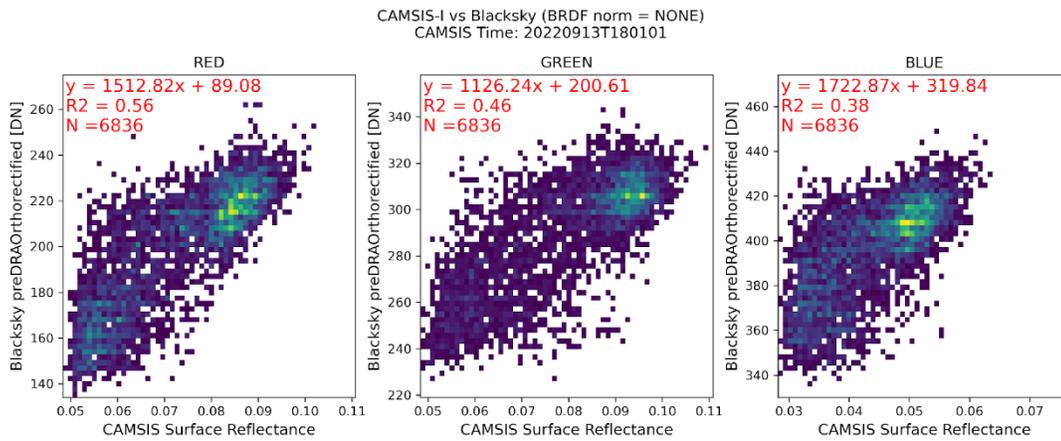
Figure 18. Comparison of BlackSky Proto product with CAMSIS surface reflectance at WLEF site.

We have further analyzed the BlackSky Proto products with simulated the Top of Atmosphere (TOA) normalized reflectance (reflectance x cosine of the solar zenith angle) using CAMSIS surface reflectance and 6S radiative transfer code and ancillary information obtained from MODIS overpass (see section 3.3, Temporal Stability). Figure 19 shows an example of the CAMSIS and BlackSky comparison.
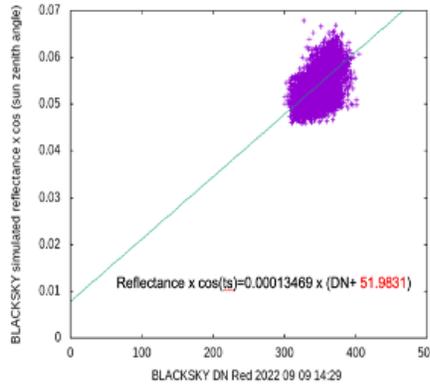
Figure 19. Comparison of the simulated TOA normalized reflectance observed by BlackSky in the Red with the BlackSky Product Digital Number. The dark count correction (number in red) is positive which is not expected (it should be negative) from uncalibrated data (note that this example uses default ancillary data, spectral response etc., the final calibration results for this day are in Table 3).

3.2 Signal-to-Noise Ratio (SNR)

3.2.1 Method

The SNR is typically derived from signals repeatedly measured within a fixed region of interest. The BlackSky data provided are in "DN" format, lacking conversion to TOA radiance or reflectance. Additionally, calibration (gain and offset) varies from scene to scene or acquisition to acquisition, with no information on the causes or governing algorithm for these calibration changes provided by BlackSky experts. Given these limitations, coupled with the very short time series of acquired images, accurately evaluating SNR from multi-temporal images (or temporal SNR) is challenging. As an alternative, spatial SNR is defined as the mean signal intensity divided by its variation across space. The spatial SNR is calculated as $\frac{mean\ signal\ (\mu)}{standard\ deviation\ of\ the\ signal\ (\sigma)}$, where signals are from adjacent pixels within small windows, assuming homogeneity in surface conditions within these windows. For this analysis, we've chosen Proto DN values within 9 pixels × 9 pixels windows over Libya-4. Libya-4's bright and relatively uniform characteristics make it suitable for spatial SNR analysis. In each scene, all pixels are regrouped into small windows, and SNR within these windows is calculated individually. The method is identical to the EDAP analysis for BlackSky. However, since the calculated $\frac{\mu}{\sigma}$ still contains the possible surface variability due to sand dunes and shadows, we assume that accurate sensor SNR can be derived from the average of 95-98th percentile values of $\frac{\mu}{\sigma}$, where the surface can be considered homogeneous. We exclude the highest $\frac{\mu}{\sigma}$ windows (98-100th percentile values) to avoid stripe artifacts in the image caused by adjacent dead detectors.

The SNR for BlackSky was also evaluated by two investigations of aquatic targets. One of these investigations looked at SNRs over clear water under typical illumination (Fig. 20) and found them to be 50-60 for the three visible bands, comparable to those of Landsat-7/ETM+ for corresponding visible bands [Hu et al., 2012]. Considering its ground resolution of ~1 m pixels, such SNRs are beyond expectations and sufficient to detect algae bloom features.
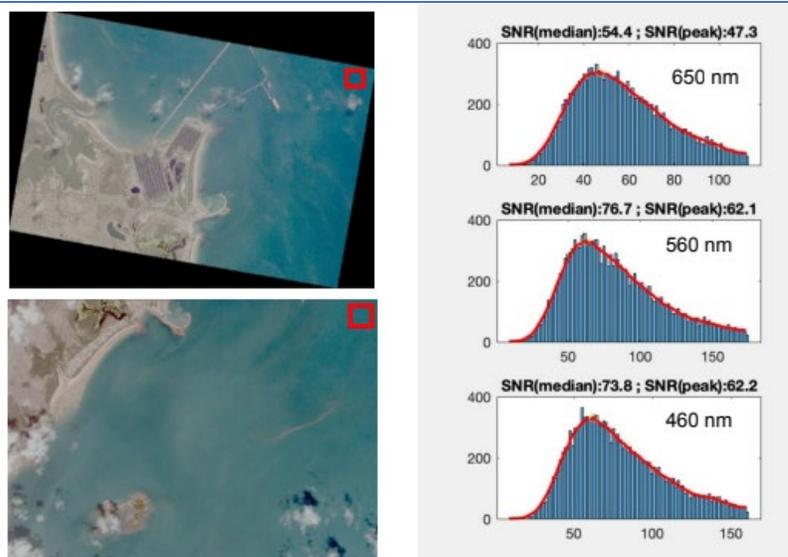
Figure 20. Left column shows the images where SNRs were evaluated (red squares), with results presented in the right column. The mode of SNRs for the three bands ranges from 47.3 to 62.2, while the median of SNRs ranges from 54.4 to 73.8. These are comparable to those of Landsat-7/EMT+ (Hu et al., 2012).

The other analysis was to investigate the SNR performance of the Black Sky Global (BSG) images thoroughly over water. The SNRs were estimated for each sensor, given that each instrument response can vary. The "on-demand" Proto products were utilized to compute SNRs, since the Anthro products went through a scene-by-scene non-linear dynamic range adjustment. This investigation estimated SNRs from the Proto data over Lake Erie, Crater Lake, and Lake Tahoe sites by carefully selecting calm waters (*i.e.,* void of wind-driven waves), little-to-no cloud and cloud shadow-impacted scenes. The SNRs were estimated by taking the ratio of mean (μ) and standard deviation (σ) of the DNs within "small" windows (9x9 pixels) and computing the average (*i.e.,* average of μ/σ). The 9x9-pixel windows were selected by investigating the histograms of the ratio (μ/σ) for 3x3, 5x5, 7x7 and 9x9 windows in a similar manner mentioned in Zanoni, et al., (2002) and Saunier (2021). At least 100,000 pixels from a BSG scene were used to compute the final SNRs. Means and standard deviations of the SNRs are presented in Figure 21, with error bars representing one-standard deviation. Note that at least four scenes were exploited where error bars are present, otherwise SNRs were computed from one scene (due to lack of good quality images). On average, SNRs for the blue and green band is ~ 50, while the SNR for the red band is < 40. SNRs are typically reported with the corresponding TOA radiance. Higher standard deviation bars may be due to the radiance differences between different scenes. Previously reported [10] over-water Landsat 8 Operational Land Imager (OLI) and Sentinel-2 multispectral instrument (MSI) SNRs are presented in Figure 8b, which show that the OLI SNRs are ~ 2 to 8 times higher than BSG, and BSG and MSI SNRs are approximately similar for the red band (664 nm). It is noteworthy that the BSG SNR values should be considered with caution since the radiance/reflectance values are unknown.
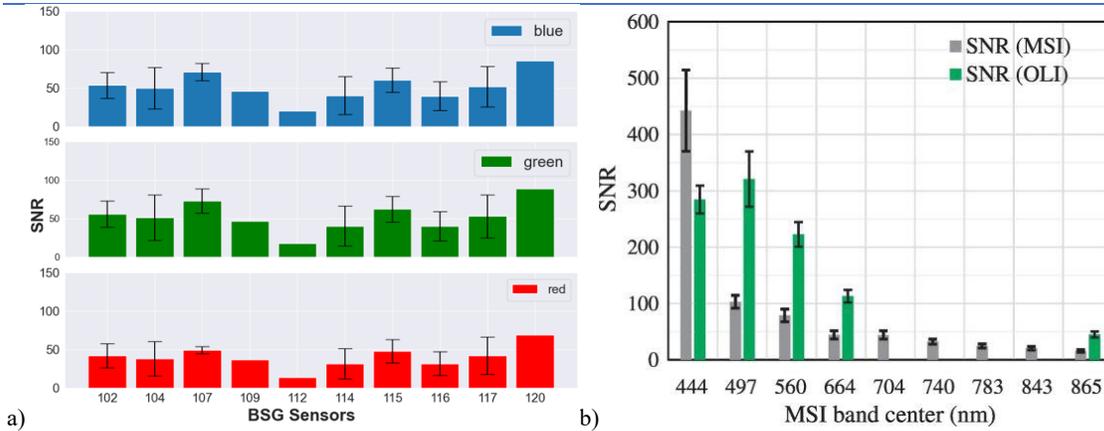
Figure 21. a) BSG SNRs over water computed from Proto product (N = 40), b) L8/OLI and S2/MSI SNRs calculated over clear, uniform bodies of water [10]. OLI SNRs are ~ 2 to 8 times higher than BSG, while MSI SNRs are ~ 2X BSG for the blue (497 nm) and red (560 nm) band.

3.2.2 Results Compliance

The $\frac{\mu}{\sigma}$ histogram, calculated from a single scene of Global-01 on May 21, 2022, is presented in Figure 22 for the three bands. The mean $\frac{\mu}{\sigma}$ values for the Blue band (96.70) are marginally lower than those for the Green (110.30) and Red (109.09) bands. Additionally, the distribution shape of Green and Red is very similar each other. As averaging the selected 95-98th percentile, we estimated SNR for Blue, Green, and Red bands as 142.53, 162.98, and 159.94, respectively.
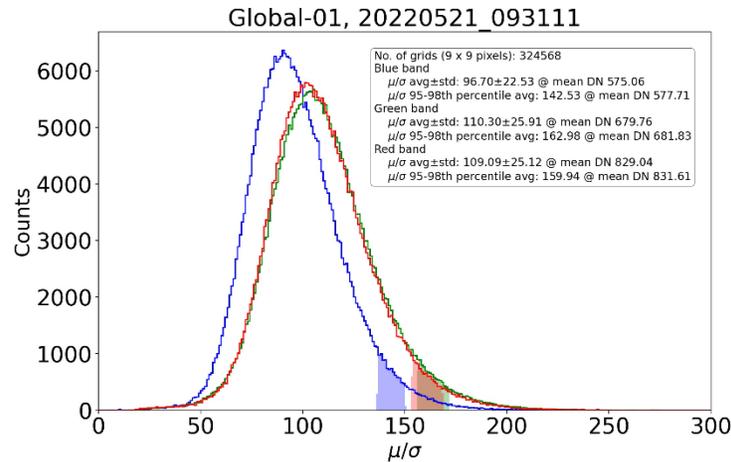
Figure 22. The $\frac{\mu}{\sigma}$ histogram computed for each spectral band (Blue, Green, and Red) for the product Global-01 on May 21, 2022, at 09:31:11 UTC over Libya-4. The color-shared ranges correspond to the 95-98th percentile for SNR estimation.

The same analysis was extended to other available BlackSky images covering Libya-4. The mean and standard deviation of the calculated SNR are summarized in Table 2. Notably, the variation across different scenes is relatively small, with standard deviations ranging from 2.57 to 11.37 for the Blue band. Consistently, a common characteristic across all satellites is the higher SNR (about 5-20) in the Green and Red bands compared to the Blue band. The Global-4 satellite exhibits the highest SNR (188-200), while the lowest SNR is observed in the Global-

13 satellite (128-136). SNR is typically associated with radiance levels and might be lower over darker scenes, such as oceans.

Table 2. Mean and standard deviation of SNR for Blue, Green and Red bands of Global-1 to -20.

Satellite	Scenes	Blue	Green	Red
G1	1	143.53	162.98	159.94
G2	1	147.99	167.91	165.89
G4	1	188.01	199.66	198.51
G8	2	143.32 ±3.73	150.03 ±3.35	149.45 ±3.07
G9	6	127.87 ±4.14	134.15 ±4.14	133.72 ±4.36
G12	2	140.18 ±3.28	146.27 ±3.46	145.43 ±3.48
G13	2	128.46 ±2.58	136.44 ±2.41	135.78 ±2.59
G14	10	142.82 ±11.37	149.78 ±11.79	149.35 ±11.78
G15	10	143.16 ±6.03	149.45 ±6.05	149.53 ±5.89
G16	3	135.85 ±4.77	143.03 ±4.58	142.40 ±4.27
G17	9	131.55 ±2.80	136.99 ±3.07	135.85 ±3.04
G20	4	132.54 ±2.57	137.86 ±2.46	136.99 ±2.41

3.3 Temporal Stability

3.3.1 Method

The method described in the previous section was generalized to all coincident observations between CAMSIS and BlackSky. The results are presented in Tables 3 to 5.

3.3.2 Results Compliance

The results are presented in Tables 3 to 5. The calibration coefficients are highly variable even when considering the same satellite (ex: G12 or G13). The dark currents are also variable and are mostly negative for the red which is abnormal.

Table 3: Calibration results in the red band

Date	Time	Sat	Slope	DN0
9/9/22	14:29:25	G12	0.00013077	-122.578
9/12/22	18:49:05	G12	0.00019551	-193.343
9/12/22	21:59:15	G8	0.00014234	-210.683
9/13/22	18:04:57	G13	0.00020318	-196.732
9/17/22	16:49:00	G12	0.00015693	-137.955
9/19/22	15:37:07	G13	0.0001095	-357.721
9/19/22	16:25:43	G2	0.00015112	-254.402

Table 4: Calibration results in the green band

Date	Time	Sat	Slope	DN0
9/9/22	14:29:25	G12	0.0001463	33.9115
9/12/22	18:49:05	G12	0.00021583	-66.4751
9/12/22	21:59:15	G8	0.00011617	-250.274
9/13/22	18:04:57	G13	0.00022416	-84.947
9/17/22	16:49:00	G12	0.00018529	50.2972
9/19/22	15:37:07	G13	0.00012763	-211.806
9/19/22	16:25:43	G2	0.00016896	-150.463

Table 5: Calibration results in the blue band

Date	Time	Sat	Slope	DN0
9/9/22	14:29:25	G12	5.08E-05	246.983
9/12/22	18:49:05	G12	1.02E-04	169.122
9/12/22	21:59:15	G8	5.36E-05	369.985
9/13/22	18:04:57	G13	1.23E-04	133.459
9/17/22	16:49:00	G12	9.92E-05	-19.1128
9/19/22	15:37:07	G13	5.68E-05	382.631
9/19/22	16:25:43	G2	9.75E-05	231.66

It is interesting that the BlackSky calibration slope in Red and Green bands are well-correlated (Figure 23) which might indicate that most of the variability is due to the data being collected with different integration times.
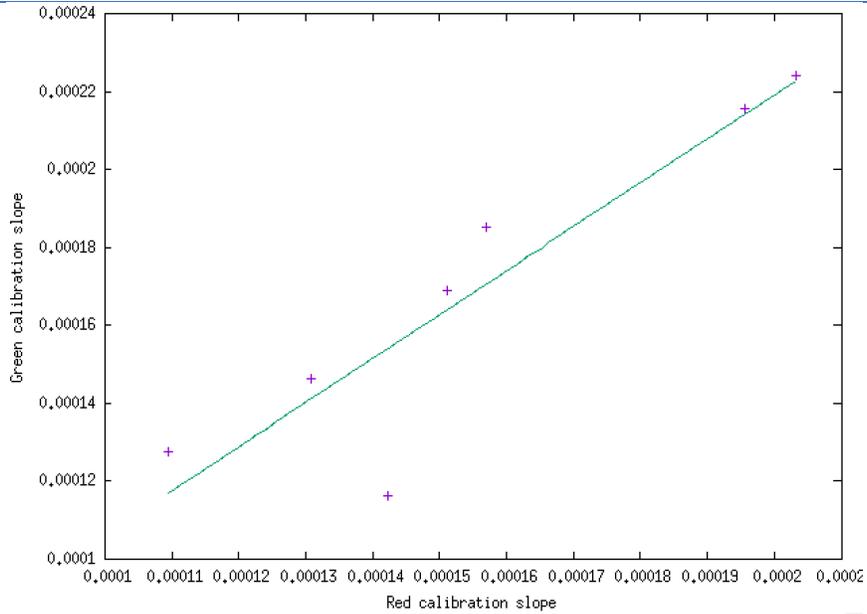
Figure 23: Calibration slope in the green vs red bands.

4 Detailed Validation – Geometric

4.1 Sensor Spatial Response (SSR)

4.3.2 Method

Assessment of sensor spatial response (SSR) is done with images over a variety of Cal/Val locations with a set of black and white squares slightly slanted with the image grid direction, each at least 20 m x 20 m in size. BlackSky images were requested from as many satellites as possible, and acquired for sensors Global 2, 4, 5, 8, 9, and 12 – 20. Sensor 7 did not image any of these locations, sensor 1 was taken offline before we started this evaluation. Images taken by sensors 2 and 8 were too noisy for the method described here. Most evaluations are assessed at the Baotou, China location and supplemented with Shadnagar, India and Big Spring, TX, USA where Baotou images were unavailable. BlackSky images used were non-standard product level of Proto (before the dynamic range adjustment is applied for the standard Anthro products).

For each direction along row or along column, assessment starts with reading in 20 x 20 rows and columns across the black/white (B/W) transition (Figure 24a). A transition line location is estimated based on visual inspection of each image. Pixels are then transformed from bins to distance (Fig. 24b) from the transition line with the equation $d_p = d*\cos(\Theta)$, where Θ is the angle between the line and vertical/horizontal, d is the vertical/horizontal distance from the pixel center to the transition line, and d_p is the perpendicular distance from pixel center to the transition line.
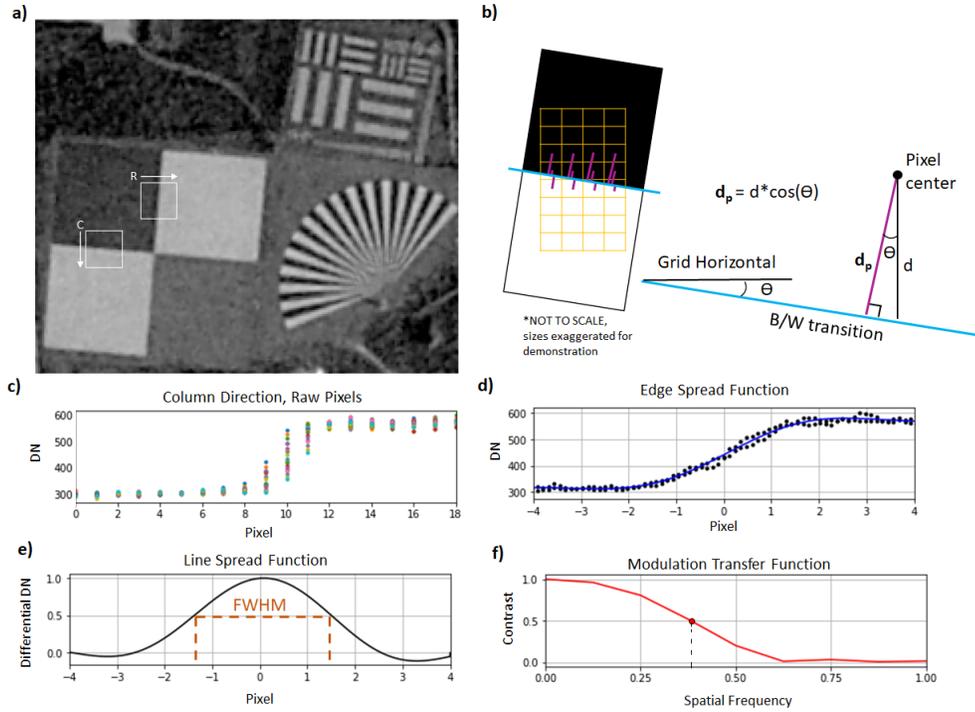
Figure 24. Visual demonstration of our SSR calculations. a) BlackSky red band image (BSG-115-20220220-001023-19247291) over Baotou, China site. White overlay squares indicate regions used in Edge Spread Function calculations in the row (R) and column (C) directions. b) Sub-pixel resolution is calculated based on pixel center distance (purple lines) from the black-to-white (B/W) color transition (blue line). Example demonstration is for the column direction. c) Raw pixels read in from box C in Figure22a, colored by column. d) Sub pixel edge response constructed based on Figure22b (black points) with a polynomial fit line (blue) for the Edge Spread Function (ESF). e) Line Spread function (LSF) is the derivative of ESF (Fig.22d). FWHM is shown in brown dashed lines, here it is 2.70 pixels. f) Modulation Transfer Function (MTF) is the Fourier Transform of the LSF (Fig.22e). Here, the x axis is normalized by Nyquist Frequency = 1 cycle in 2 pixels. GRD (ground resolved distance) is calculated based on the inverse of the frequency where MTF = 0.5 (red dot). Here, GRD is 1/.38 = 2.63 pixels.

Once the values are transformed into distance from the B/W transition, a polynomial is fit to the line, creating an Edge Spread Function (ESF) (Fig. 24d). The derivative of the ESF is the Line Spread Function (LSF) (Fig. 24e) in the direction of interest. Full Width at Half Maximum (FWHM) is found from this LSF to represent the sensor's effective footprint size. Finally, the Fourier transform of the LSF gives the Modulation Transfer Function (MTF) (Fig. 24f). Two more metrics for spatial response are found with the MTF curve, MTF value at Nyquist frequency and half wavelength where the MTF response reduces to half (we refer to this as the ground resolved distance [GRD]). We will evaluate BlackSky's spatial response with FWHM as its effective footprint size.

This evaluation includes 11 out of 15 BlackSky sensors. Their Globals -1, -2, -7, and -8 were not evaluated either due to noisy images or no passes over the target sites.

4.1.2 Results Compliance

BlackSky's sensors use an RGBW (red, green, blue, and pan) color filter array in an extended Bayer mosaic filter style. As their documentation states, "BlackSky uses advanced demosaicing algorithms to generate full resolution panchromatic and RGB images from the raw [RGBW color filter array] pixels" [BlackSky, 2022]. BlackSky indicated during a

technical exchange meeting that information from a neighborhood of 5 x 5 pixels from all available bands are used to demosaic values at each pixel when creating the full RGB and panchromatic images. As a result, their sensor spatial response is much better than expected (the footprint size expected for the blue and red bands to be 5 to 6 pixels, see Figure 23a). The average spatial resolution expressed in FWHM of LSF is 2.48 pixels in row direction and 2.63 pixels in column direction. Each band has a similar response, with blue band preforming the best and red preforming the worst. Pan and green bands often match, and both perform better than the red band but worse than the blue band. Each sensor preforms slightly differently than the other, see Appendix Table A2 for each sensor response evaluated.

One caveat to point out is that we are concerned with the radiometric performance of the BlackSky images due to the demosaicing algorithms they use to construct RBGW images. In Figure 25, the right image illustrates the artifact from such demosaicing scheme.
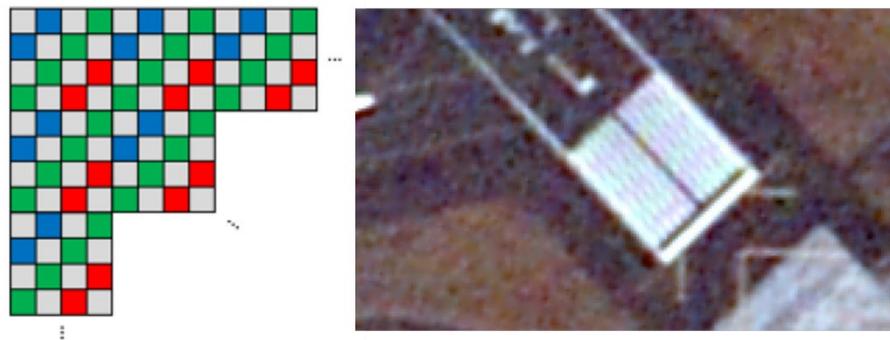
Figure 25. Left graphic shows BlackSky's extended Bayer color mosaic filter [from BlackSky, 2022]. The gray squares represent the panchromatic band. On the right is an image that Illustrates the color bleed problem, pink and green striping in the white lines (Image ID: BSG-104-20220106-133822-17364025).

There was a single tasked image BlackSky delivered outside their product specification with a pixel size of 3 m. Out of the 200+ images ordered, only one had this problem. The image was removed from our quota count and not used in the assessment.

BlackSky's spatial response is impressive, given their color mosaic filter. According to the guidelines, BlackSky's spatial sensor response grade is considered **Basic** because the ratio of the average sensor spatial response to pixel size is 2.48:1 in the row direction and 2.63:1 in the column direction. Both are higher than the recommended ratio of 2.0:1 [ESA-NASA, 2021].

4.1.2.1 Qualitative Assessment of Low Light Images

Low Light Anthro images were acquired at three USA locations. The quality of these images was examined by visual comparison to standard product images of the same locations. Comparisons were made between sensors of similar SSR performance (Appendix Table A3) and taken within one week of each other. Nine images were acquired, of which one had no distinguishable features due to a small signal to noise ratio. In general, low light images are of similar but lower quality than the standard products. Individual buildings and vehicles are still distinguishable, but lines in parking lots, roads, and football fields are not (Fig. 26). There is typically a blue dominance to the low light images, along with lower signal to noise ratio

manifesting as static. This combined with the reduced quality makes individual trees indistinguishable. Figure 26 shows comparisons of images from the same location for both low light and standard products.
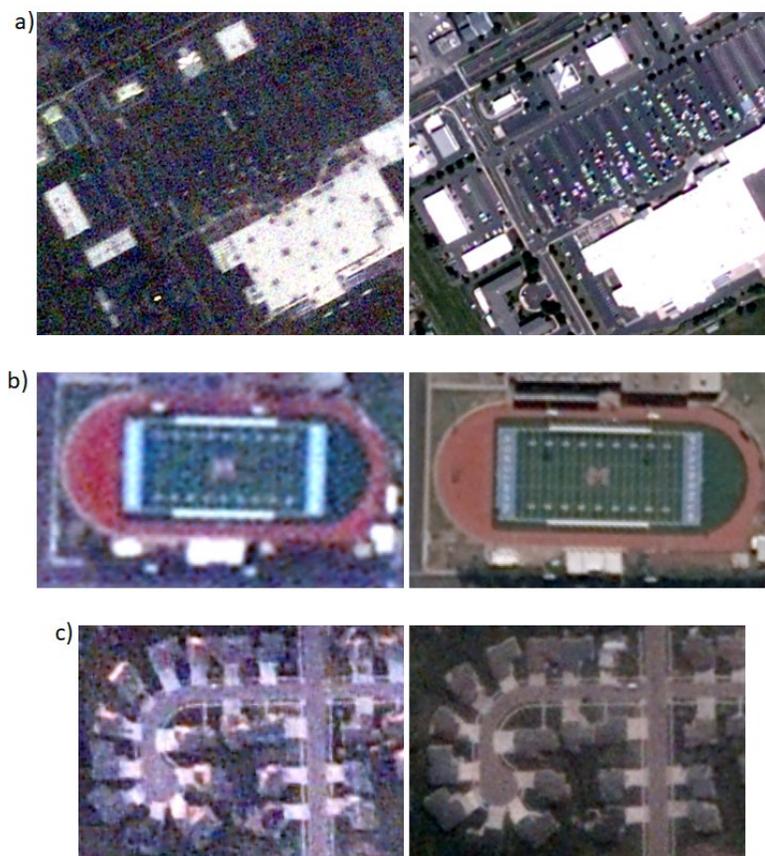
Figure 26. Visual comparison of low light images (left) with standard images (right) of the same locations in a) Pennsylvania, b) Texas, and c) South Dakota. a) Cars are visible in the low light image, but parking lot lines are not. b) In the low light image, football field lines are not visible, and lettering is illegible. c) In the low light image, individual houses are visible, but the sidewalks are not always distinguishable from the road. All low light images have more color static than the corresponding standard image.

4.2 Absolute Positional Accuracy (APA)

4.2.1 Method

Evaluation of BlackSky geolocation accuracy is a relative assessment with the panchromatic band of WorldView (WV) -2 and -3 images as the reference for most of this assessment. National Agricultural Imagery Program (NAIP) images were used to assess nine low light image acquisitions. WorldView imagery has a resolution of 0.3 m - 0.8 m and a CE90 of 5.4 m [DigitalGlobe Inc, 2016]. We orthorectify WV images with the 30 m Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM). BlackSky and NAIP data are delivered orthorectified, so we do not orthorectify BlackSky or NAIP data. NAIP has a resolution of 0.60 m and an accuracy of 1.80 m. We compared the red band of NAIP images to the red band of BlackSky's low light imagery.

Locations of interest for evaluation are either within cities or at airports. They all follow the criteria of at least 3 km^2 with minimal tree cover, low buildings, and no clouds. At the BlackSky pixel resolution of 1 m, changing tree texture and building shadows will interfere with the matching algorithm (described below). We would like to note that this is a best-case scenario evaluation done over locations with easy to match features. Many science investigations occur in remote locations that are harder to match and can expect geolocation accuracies that are lower than what we report here.

The assessment algorithm starts by determining the area of overlap between reference and target images. This overlapped region is then split into subset images of 250 m x 250 m image chips. Each chip in the target image has a matching chip in the refence image based on the image geolocation metadata. The algorithm then imposes offsets on the target chip of the pair and calculates the Pearson Cross Correlation (PCC) coefficient (a measurement of how well two images match). The offsets that give the best PCC are taken to be the geolocation offset between the chip pairs. Quality of chip co-registration is then calculated with a measurement uncertainty equation [De Luccia et al. 2016] and used to filter out poor quality chip matches [Semple et al, 2023].

Where possible, we acquired 5 WV images over the area of interest. WV images have inherent geolocation errors (5.4 m CE90), so we assessed the geolocation accuracy of this group of images to find which images clustered and which were outliers. The most central image of the cluster was used as our main reference image when evaluating BlackSky's geolocation accuracy. This increased our confidence in the accuracy of the reference data. Some locations evaluated had either poor clustering, or we were only able to acquire 1 image of the area of interest; these locations are marked with an asterisk (*) in the full results table in the Appendix (Table A1).

Most metrics we calculated were standard metrics, except CE90-demean. CE90-demean is similar to the standard CE90 calculation but assumes there is no systematic bias between the BlackSky and the reference images. CE90-demean measures the uncertainty of BlackSky's geometric calibration system. Figure 27 shows a visual demonstration of CE90 vs CE90-demean where the two values are different.
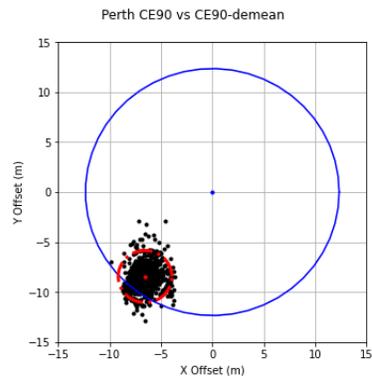
Figure 27. Perth, Australia's CE90 (blue circle) and CE90-demean (red dash-dot circle) difference shows grouping of evaluated BlackSky offsets (black points) off-center from the reference WV image.

4.3.3 Results Compliance

BlackSky claims a CE90 of 10 m in the USA and Australia, and 20 m CE90 everywhere else. Science users should be aware that this is well below the recommended grading for this document of within 80% of ground pixel size of about 1 m [ESA-NASA, 2021]. Our grade is based on BlackSky's CE90 claim.

We find that BlackSky's geolocation accuracy does not meet vendor's specification when considered globally, for CE90 (relative to WV) but does when considering CE90-demean (essentially relative to itself). Figure 26 shows global offsets when considering CE90 (Fig 28a, b) and CE90-demean (Fig. 28c, d). Overall, US and Australia regions have a CE90 of 23.6 m, and a CE90-demean of 9.0 m. Everywhere else has a CE90 of 23.2 m, and a CE90-demean of 16.4 m. Locally, CE90 is within specification at 12 out of 26 cities (Fig. 29a), while CE90-demean is within specification at 21 out of 26 cities (Fig. 29b). See the appendix for a table of all findings. Some locations are of significant note: Belo Horizonte, Brazil and Sapporo, Japan both have BlackSky images offset 60 m from other BlackSky images in the same location. Figures of their CE90 plots are in the Appendix (Fig.A1).
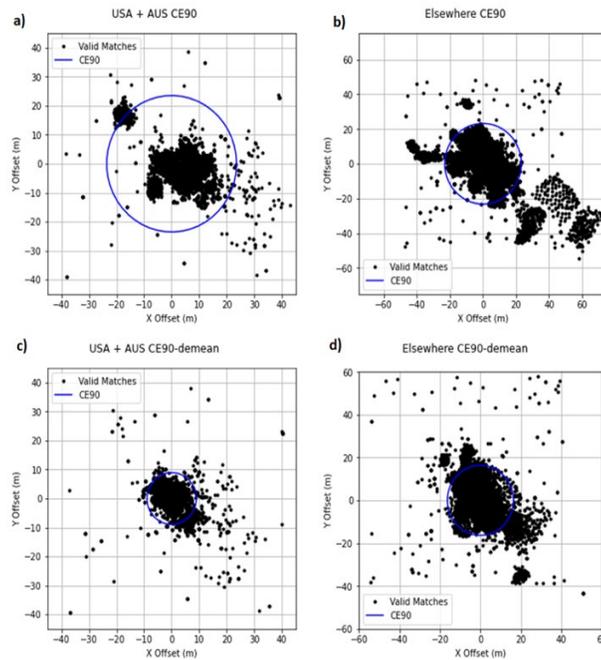
Figure 28. Plots of relative offsets for USA and Australia (a, c) and Elsewhere (b, d). Top row plots CE90, bottom plots CE90-demean. Valid matches are plotted as black dots, and the CE90 or CE90-demean as blue circles.
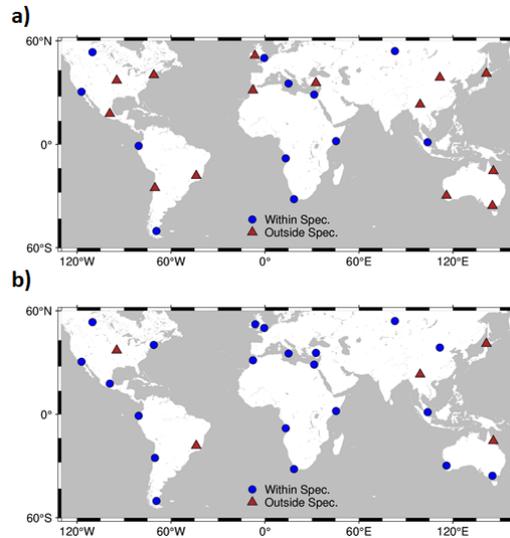
Figure 29. Global maps showing specification compliance for BlackSky for each location evaluated, considering a) CE90 and b) CE90-demean. Locations evaluated are marked with colored shapes, blue circles are within BlackSky's specification, brown triangles are outside BlackSky's specification. BlackSky's specification is CE90 < 10 m within the USA and AUS, and CE90 < 20 m elsewhere

Spot checks of BlackSky and reference WorldView images were performed on data from Brazil to confirm the source of this large offset. Another data set, Planet Labs SuperDove imagery, was used as an independent check because it has a different reference dataset from both BlackSky and WorldView. Planet Labs SuperDove imagery has a specified pixel size of 3 m. The false color image subsets in Figure 30 show the offsets between all three data sets. The BlackSky image is ~65m offset from both WorldView (Fig. 30a) and SuperDove (Fig. 30b) images. SuperDove imagery is ~1 m offset from WorldView imagery (Fig. 30c). We are confident that the Blacksky image is not well aligned with the Worldview and Planet images, while the Worldview and Planet images agree with each other.
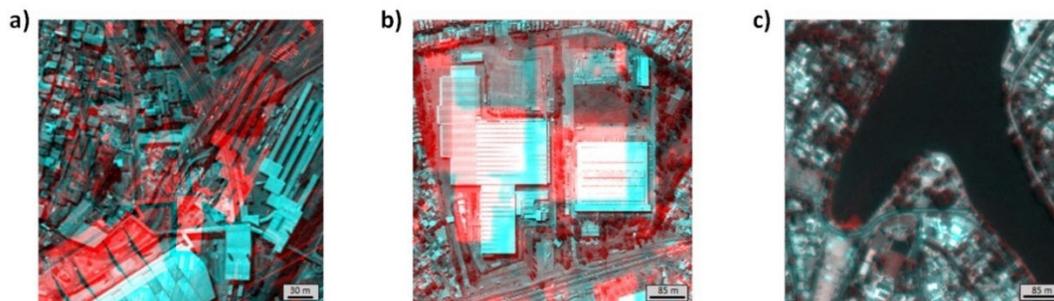
Figure 30. Subset images over various locations in Brazil, false colored to show image offsets. One image is colored red, the other cyan, where both are overlapped according to their geolocation metadata. Where red or blue is visible, there are offsets between the two images. a) A composite of WorldView as reference (red) with a BlackSky image (cyan) over a train station has a 65 m offset between the images. b) SuperDove as reference (red) with BlackSky image (cyan) over a warehouse. There is a 65 m offset between these images. c) A composite image of WorldView as reference (red) with SuperDove image (cyan) over a lake shoreline. No obvious red shapes indicate they are almost perfectly aligned. Image IDs used are, WorldView image ID: 22SEP08131526-P1BS-506969282100_01_P003, BlackSky image ID: BSG-120-20220629-161343-28457195, SuperDove image ID: 20220924_124349_64_249c.

The new BlackSky products, Burst (multiple images taken in rapid succession), Area (two images with 20% overlap), and Low Light (sun is just below the horizon) were also evaluated for

geolocation accuracy. Stereo imagery were also acquired, but could not be assessed with our methods, and thus were not evaluated. We selected two cities to evaluate the Area and Burst products, one with good geolocation performance, and one with poor geolocation performance. Low Light images were acquired in cities where night lights would be prominent but would not oversaturate. Findings are summarized here, but for a full table of results see the Appendix for Table A2.

Burst products were evaluated at Rio Gallegos, Argentina (8 acquisitions of 5 images) and Manta, Ecuador (8 acquisitions of 5 images). The relative geolocation accuracy of Argentina burst images to WV reference image performed slightly better than the standard product, with a CE90 = 11.1 m and a CE90-demean = 3.0 m. When grouped by Burst acquisition, the self-consistency within a Burst group improves to a CE90-demean of 1.9 m. The relative geolocation accuracy of Ecuador burst images is CE90 = 15.8 m and the CE90-demean = 4.9 m. When grouped by Burst acquisition, the CE90-demean remains the same, 4.9 m. Of significant note, two burst groups have a large internal variation, with CE90-demean of 7.0 m and 11.1 m. We found that the relative geolocation accuracy between BlackSky images within a Burst acquisition is much better than in the standard product. However, the offsets are larger than the pixel size (~1 m), so an end user would need to either rectify the images to a reference image or aggregate the pixels before processing.

For the Area product, we evaluated performance at San Diego, CA, USA (7 Area pairs) and Mexico City, Mexico (9 Area pairs). Visual inspection of two images (Pairs 1 & 5) reveals small, 1 - 3-pixel offsets along the image seam. Geolocation evaluation of image pairs shows an average relative stability of CE90-demean = 3.1 m over California and 2.9 m over Mexico. This is improved stability for both locations when compared to BlackSky's standard single acquisition product.

Low Light image products were acquired at three different US locations, Sioux Falls, SD (4); Lubbock, TX (1); and Philadelphia, PA (4) at a time when the sun was just below the horizon. Nine images were acquired, of which one was not usable because it was too dark. These products were evaluated for geolocation accuracy with NAIP imagery as the reference data set. The low light product performs within spec at all tested locations, with CE90 = 3.90 m (Table 6). The stability of these acquisitions (CE90-demean = 2.93) is similar to the stability of other USA and AUS acquisitions.

Table 6. NAIP vs Low Light imagery results.

A City Within:	# of Images	# of Valid Matches	X Offset (m)	Y Offset (m)	X StdDev (m)	Y StdDev (m)	X RMSE (m)	Y RMSE (m)	CE90 (m)	CE90-demean (m)
TX	1	709	-2.34	1.99	0.8	0.93	2.48	2.2	4.49	1.85
PA	4	3188	-0.39	1.88	1.49	1.05	1.54	2.15	3.7	2.56
SD	3	1365	-0.82	1.55	1.44	1.87	1.65	2.43	3.87	3.65
Total:	**8**	**5262**	**-0.76**	**1.81**	**1.54**	**1.31**	**1.72**	**2.23**	**3.9**	**2.93**

4.3 Band-to-Band Registration (BBR)

4.3.1 Method

Band-to-band registration (BBR) is assessed with the image matching algorithm that is described above in the APA section 4.2. Each band of an image is assessed with its corresponding panchromatic image for reference.

4.3.4 Results Compliance

BlackSky documentation makes no specific claims about BBR. Of the images examined for BBR, 100% of the green band is within 0.2 m of the panchromatic band in both directions, 100% of the blue band is within 0.4 m of the panchromatic band in both directions, and 95% of red band is within 0.4 m of the panchromatic band. 5% of the images had a 1 m east-west shift in the red band, resulting in the appearance of red and green stripes in the full RGB color image.

We gave BlackSky's BBR a grade of Excellent because the red band was found to be 1 m off in 5% of the images.

4.4 Temporal Stability

4.4.1 Method

Temporal stability is assessed with the image matching algorithm that is described above in section 4.2.1. Each location is assessed with the earliest BlackSky image in the set as the reference image. We acquired one BlackSky image per month, where possible, over three US cities for 1 year of temporal coverage. Snow cover in Boston, MA made some months not usable in this assessment. We obtained 11 images over Logan Airport in Boston, MA, 13 image over Kansas City Airport in Kansas City, MO, and 14 images over San Diego Airport in California.

4.4.2 Results Compliance

BlackSky does not make claims on temporal stability beyond the specification that geolocation accuracy within the USA and Australia is 10 m CE90, and 20 m CE90 elsewhere. Based on CE90-demean, a metric for BlackSky self-consistency, San Diego, CA (CE90-demean = 7.8 m) and Boston, MA (CE90-demean = 6.4 m) are within specification and preform similarly. Kansas City, MO is out of specification with a CE90 = 10.6 m. The average offsets of each image at San Diego, CA (Fig. A2a) vary from 0.1 m - 7.4 m, in the east-west (EW) direction and from 0.1 - 7.4 m in the north-south (NS) direction. The trend is mostly stable with a large jump in offsets to the southeast in the last two images. Boston, MA average offsets (Fig. A2b) trend to the southeast with the largest offsets at the end. Mean offsets here vary from 3.1 m - 11.3 m in the EW direction and 0.5 m - 14.1 m in the NS direction. Kansas City, MO shows no clear temporal trends (Fig. A2c). The average offsets here are generally larger than at the other two sites and vary from 2.3 m – 26.6 m in the EW direction and 0.1 m - 11.0 m in the NS direction. No locations perform well enough for end-users interested in time series analysis to process the data as delivered. Users will either need to rectify the images to a main image themselves or aggregate the pixels before preforming a time series analysis. Because of this, we give BlackSky's temporal stability a grade of Basic.

5   References

Alonso, K., Bachmann, M., Burch, K., Carmona, E., Cerra, D., de los Reyes, R., Dietrich, D., Heiden, U., Hölderlin, A., Ickes, J., Knodt, U., Krutz, D., Lester, H., Müller, R., Pagnutti, M., Reinartz, P., Richter, R., Ryan, R., Sebastian, I., & Tegler, M. (2019). Data Products, Quality and Validation of the DLR Earth Sensing Imaging Spectrometer (DESIS). Sensors, 19(20), 4471. https://doi.org/10.3390/s19204471

Choi, M., Lyapustin, A., Wang, Y., Tucker, C. J., Khan, M. N., Policelli, F., Neigh, C. S. R., & Hall, A. A. (2024). Calibration of Maxar Constellation Over Libya-4 Site Using MAIAC Technique. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17, 5460–5469. https://doi.org/10.1109/JSTARS.2024.3367250

De Luccia, F. J., Houchin, S., Porter, B. C., Graybill, J., Haas, E., Johnson, P. D., Isaacson, P. J., and Reth, A. D. (2016), "Image navigation and registration performance assessment tool set for the GOES-R Advanced Baseline Imager and Geostationary Lightning Mapper", Proc. SPIE 9881, Earth Observing Missions and Sensors: Development, Implementation, and Characterization IV, 988119 (2 May 2016); doi: 10.1117/12.2229059.

DigitalGlobe Inc. (2016), "Accuracy of worldview products". https://dg-cms-uploads-production.s3.amazonaws.com/uploads/document/file/38/DG_ACCURACY_WP_V3.pdf. Last accessed 15 June 2023.White Paper.

ESA, NASA (2021), "Earth Observation Mission Quality Assessment Framework - Optical Guidelines". Issue 0.1 (draft), 2021.

Hu, C., L. Feng, Z. Lee, C. O. Davis, A. Mannino, C. R. McClain, and B. A. Franz (2012). Dynamic range and sensitivity requirements of satellite ocean color sensors: learning from the past. Appl. Opt., 51:6045-6062.

Korkin, S., & Lyapustin, A. (2019). Matrix exponential in C/C++ version of vector radiative transfer code IPOL. Journal of Quantitative Spectroscopy and Radiative Transfer, 227, 106–110. https://doi.org/10.1016/j.jqsrt.2019.02.009

Krutz, D., Müller, R., Knodt, U., Günther, B., Walter, I., Sebastian, I., Säuberlich, T., Reulke, R., Carmona, E., Eckardt, A., Venus, H., Fischer, C., Zender, B., Arloth, S., Lieder, M., Neidhardt, M., Grote, U., Schrandt, F., Gelmi, S., & Wojtkowiak, A. (2019). The Instrument Design of the DLR Earth Sensing Imaging Spectrometer (DESIS). Sensors, 19(7), 1622. https://doi.org/10.3390/s19071622

Lyapustin, A. I. (2005). Radiative transfer code SHARM for atmospheric and terrestrial applications. Applied Optics, 44(36), 7764. https://doi.org/10.1364/AO.44.007764

Lyapustin, A., Wang, Y., Xiong, X., Meister, G., Platnick, S., Levy, R., Franz, B., Korkin, S., Hilker, T., Tucker, J., Hall, F., Sellers, P., Wu, A., & Angal, A. (2014). Scientific impact of MODIS C5 calibration degradation and C6+ improvements. Atmospheric Measurement Techniques, 7(12), 4353–4365. https://doi.org/10.5194/amt-7-4353-2014

Lyapustin, A., Wang, Y., Korkin, S., & Huang, D. (2018). MODIS Collection 6 MAIAC algorithm. Atmospheric Measurement Techniques, 11(10), 5741–5765. https://doi.org/10.5194/amt-11-5741-2018

Lyapustin, A., Y. Wang, M. Choi, X. Xiong, A. Angal, A. Wu, D.R. Doelling, R. Bhatt, S. Go, S. Korkin, B. Franz, G. Meister, A. M. Sayer, M. Roman, R. E. Holz, K. Meyer, J. Gleason, R. Levy (2023). Calibration of the SNPP and NOAA 20 VIIRS Sensors for Continuity of the MODIS Climate Data Records, Rem. Sens. Environ., 295, https://doi.org/10.1016/j.rse.2023.113717.

Saunier, S., "Technical Note on Quality Assessment for BlackSky," 2021.

Semple, A., Tan, B. and Lin, G. (2023), "Automation of Geometric Accuracy Assessment Algorithm", JACIE 2023 Workshop.

USGS Spatial Sites Catalog (https://calval.cr.usgs.gov/apps/spatialsites_catalog), Last accessed 15 June 2023.

Zanoni, V., R. Ryan, K. Holekamp, and M. Pagnutti, "IKONOS Signal-to-Noise Ratio Estimation," in High Spatial Resolution Commercial Imagery Workshop-IKONOS, 2002.

Appendix A. Additional Figures and Tables

A City Within:	# of Images	# of Valid Matches	X Offset (m)	Y Offset (m)	X StdDev (m)	Y StdDev (m)	X RMSE (m)	Y RMSE (m)	CE90 (m)	CE90-demean (m)
Boston, MA	10	1105	4.7	-5.4	3.2	3.6	5.7	6.5	11.3	6.4
Kansas City, MO	12	2225	9.4	-0.5	6.0	5.8	11.2	5.8	17.4	10.6
San Diego, CA	13	1433	-0.6	-0.5	3.0	2.9	3.1	2.9	7.6	7.8
Cairns	6	348	-1.2	0.5	9.9	9.6	10.0	9.6	16.7	16.9
Perth	6	773	-6.5	-8.5	1.1	1.3	6.6	8.6	12.4	2.6
Melbourne	6	1042	-17.6	16.1	1.6	1.5	17.7	16.2	25.9	2.8
Canada*	4	462	-1.9	-1.3	3.3	2.1	3.8	2.4	5.9	4.7
Mexico	6	3093	-5.3	6.3	6.3	5.1	8.3	8.1	24.2	16.1
Ecuador	7	1654	-0.6	5.0	9.0	6.2	9.0	8.0	17.9	19.8
Chile	5	2784	5.4	-3.7	6.7	6.3	8.6	7.3	23.8	17.3
Brazil	5	616	37.3	-23.4	16.5	13.6	40.8	27.1	71.8	28.6
Argentina	9	1464	5.4	-1.2	5.8	5.9	7.9	6.0	14.7	11.3
England	5	1123	-7.3	3.0	2.3	3.7	7.7	4.7	11.7	7.0
Ireland	5	823	-7.8	11.6	5.7	5.4	9.7	12.8	20.1	9.0
Sicily	5	244	-4.0	1.5	5.0	2.6	6.4	2.9	9.4	8.1
Turkey*	5	684	-35.9	4.9	5.5	3.2	36.3	5.8	43.1	9.2
Morocco	5	1994	7.0	-9.6	7.4	13.8	10.2	16.8	23.8	15.8
Angola	6	1023	-1.6	13.2	5.4	4.2	5.6	13.8	18.9	9.9
Egypt	6	1195	5.8	3.3	7.0	8.5	9.1	9.1	16.6	15.1
Somalia	6	464	-6.0	0.8	7.0	7.3	9.2	7.4	13.4	15.4
South Africa	5	1696	-6.9	-2.7	9.4	6.0	11.7	6.5	16.3	17.2
Singapore	6	2329	-3.9	4.2	4.6	5.3	6.0	6.7	11.3	11.0
Russia*	8	923	10.4	-3.2	3.5	5.5	11.0	6.3	17.4	9.6
Japan	10	1940	4.1	-6.7	7.7	16.7	8.7	18.0	23.2	22.5
Hohhot	6	944	-17.0	-2.3	2.5	3.2	17.2	4.0	21.3	7.4
Baoshan	6	1275	-4.2	11.0	3.4	8.9	5.4	14.2	34.5	22.8
Total	173	31846								
USA + AUS	53	6931	0.2	0.4	10.2	8.4	10.2	8.5	23.6	9.0
Elsewhere	120	26749	-0.8	0.6	12.2	10.8	12.2	10.9	23.2	16.4

Table A1. Relative geolocation accuracy assessment results. Red highlights indicate locations out of the specification US and AUS CE90 = 10 m (above dashed line), elsewhere CE90 = 20 m (below dashed line). Asterisk (*) by the city name indicate locations where either only 1 WV image was obtained, or the WV images did not group, leading to less certainty in the reference image's accuracy.

Burst	# Valid Matches	CE90 (m)	CE90-demean (m)
Argentina			
Group 1	82	8.4	1.9
Group 2	591	10.9	2.5
Group 3	172	9.9	1.4
Group 4	666	9.2	1.7
Group 5	1029	11.0	2.3
Group 6	1081	10.6	2.0
Group 7	711	9.6	1.7
Group 8	2008	11.5	2.2
Total	**6340**	**10.1**	**2.0**
Ecuador			
Group 1	1323	15.3	7.0
Group 2	1359	14.4	4.1
Group 3	1001	18.3	3.1
Group 4	936	15.4	3.0
Group 5	1801	13.9	3.6
Group 6	849	16.2	4.8
Group 7	1001	16.8	11.1
Group 8	947	12.7	2.6
Total	**9217**	**15.4**	**4.9**

Area	# Valid Matches	CE90 (m)	CE90-demean (m)
California			
Pair 1	517	5.2	5.0
Pair 2	336	2.7	1.6
Pair 3	297	3.2	2.3
Pair 4	365	3.8	2.9
Pair 5	686	5.4	4.6
Pair 6	380	3.3	1.9
Pair 7	446	3.2	3.1
Total	**3027**	**3.8**	**3.1**
Mexico			
Pair 1	705	6.4	3.7
Pair 2	795	8.0	3.2
Pair 3	750	6.6	2.2
Pair 4	473	5.9	2.4
Pair 5	516	8.1	2.9
Pair 6	557	7.5	3.2
Pair 7	534	7.6	2.4
Pair 8	487	7.1	2.5
Pair 9	366	8.1	3.9

Total	5183	7.3	2.9

Table A2. Burst and Area product geolocation (CE90) and stability (CE90-demean) results.

Sensor	Pixel Size (m)	Band	FWHM (pix)	GRD (pix)	MTF @ny	FWHM (pix)	GRD (pix)	MTF @ny
			Row Direction			Column Direction		
4	1.0	pan	2.88	2.60	0.005	2.85	2.70	0.005
		R	2.96	2.70	0.002	2.95	2.70	0.006
		G	2.88	2.60	0.003	2.89	2.67	0.006
		B	2.77	2.60	0.015	2.77	2.28	0.004
5	1.0	pan	2.15	2.00	0.025	2.25	2.20	0.004
		R	2.22	2.30	0.023	2.31	2.67	0.005
		G	2.16	2.00	0.025	2.23	2.20	0.004
		B	2.09	2.00	0.028	2.18	2.00	0.002
9	1.0	Pan	2.61	2.67	0.006	2.67	2.67	0.003
		R	2.67	2.67	0.006	2.75	2.70	0.002
		G	2.63	2.67	0.006	2.67	2.70	0.003
		B	2.53	2.30	0.006	2.60	2.67	0.004
12	0.9	Pan	2.44	2.00	0.009	2.40	2.30	0.012
		R	2.51	2.30	0.009	2.52	2.67	0.014
		G	2.47	2.00	0.009	2.45	2.30	0.012
		B	2.32	2.00	0.008	2.35	2.00	0.008
13	1.0	Pan	2.54	2.50	0.003	2.54	2.67	0.011
		R	2.71	2.70	0.001	2.71	2.75	0.003
		G	2.50	2.30	0.006	2.51	3.30	0.013
		B	2.37	2.00	0.004	2.38	2.00	0.026
14	1.0	Pan	3.11	2.67	0.006	3.27	2.80	0.004
		R	3.22	3.10	0.008	3.39	3.00	0.005
		G	3.11	2.50	0.005	3.29	2.80	0.003
		B	3.00	2.50	0.003	3.17	2.67	0.004
15	0.9	Pan	2.20	2.00	0.370	2.67	2.33	0.033
		R	2.29	2.30	0.033	2.70	2.63	0.023
		G	2.20	2.00	0.036	2.67	2.33	0.036
		B	2.10	2.00	0.044	2.55	2.00	0.044
16	1.0	Pan	2.59	2.56	0.005	2.55	2.50	0.003
		R	2.65	2.67	0.005	2.63	2.67	0.004
		G	2.59	2.56	0.006	2.59	2.50	0.006
		B	2.52	2.20	0.007	2.46	2.43	0.004

17	1.0	Pan	2.30	2.00	0.007	2.34	2.43	0.007
		R	2.40	2.00	0.004	2.45	2.43	0.004
		G	2.30	2.00	0.006	2.34	2.43	0.006
		B	2.21	1.70	0.013	2.21	2.00	0.009
18	1.0	Pan	2.38	2.00	0.009	2.41	2.40	0.009
		R	2.48	2.00	0.008	4.53	2.40	0.007
		G	2.39	2.00	0.008	2.43	2.40	0.008
		B	2.27	1.70	0.015	2.29	2.00	0.009
19	1.0	Pan	2.40	2.00	0.007	2.64	2.53	0.006
		R	2.50	2.20	0.006	2.76	2.67	0.005
		G	2.43	2.00	0.006	2.64	2.53	0.006
		B	2.34	2.00	0.009	2.56	2.50	0.009
20	0.9	Pan	2.20	2.00	0.040	2.41	2.43	0.040
		R	2.20	2.00	0.030	2.43	2.50	0.033
		G	2.20	2.00	0.050	2.41	2.43	0.050
		B	2.10	1.80	0.060	2.30	2.00	0.060
Total	1.0	Pan	2.48	2.25	0.04	2.58	2.50	0.01
		R	2.57	2.41	0.01	2.84	2.64	0.01
		G	2.49	2.22	0.01	2.59	2.55	0.01
		B	2.39	2.07	0.02	2.48	2.21	0.02

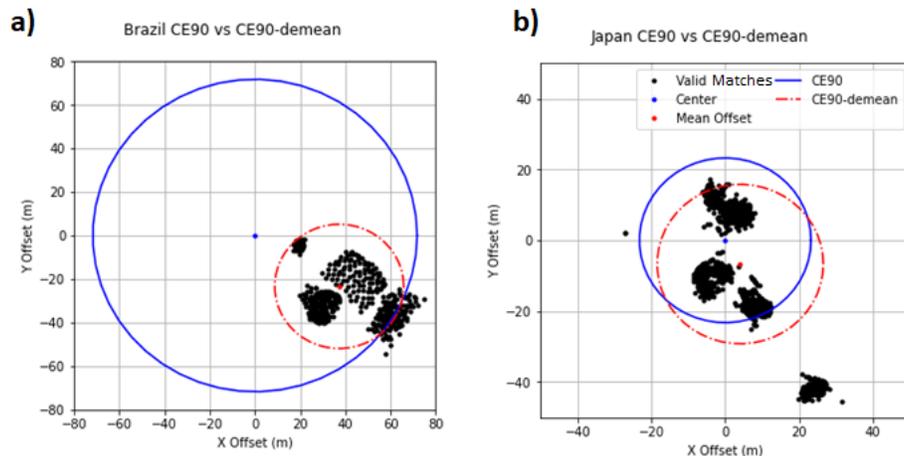**Table A3. Sensor spatial response for 12/16 sensors.**

Figure A1. CE90 (blue solid circle) vs CE90-demean (red dot-dash circle) for a) Belo Horizonte, Brazil and b) Sapporo, Japan. Both locations have the largest spread in offsets between BlackSky images. This is not representative of all locations but is surprising and should be noted. X is E-W direction, Y is N-S direction.
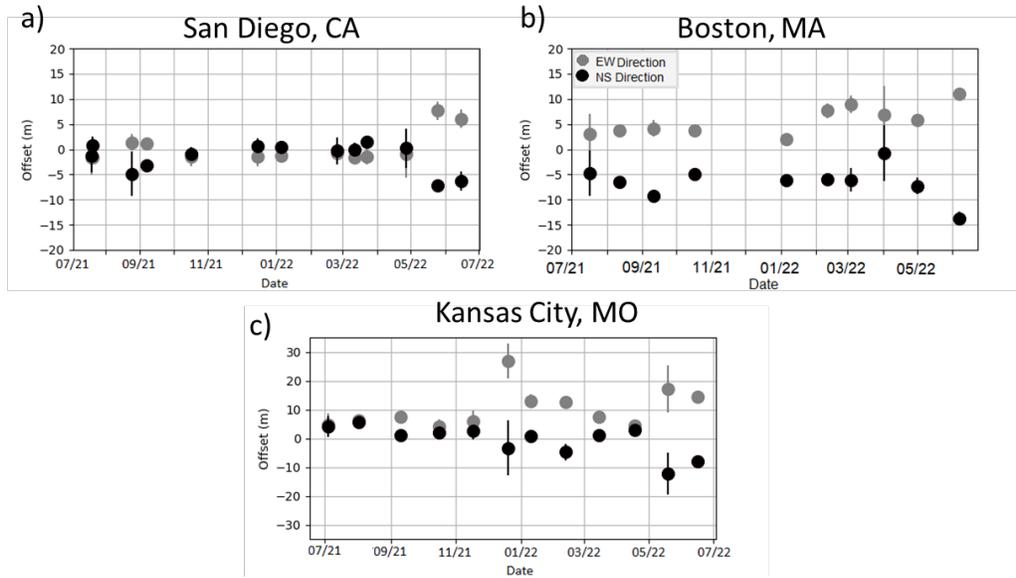
Figure A2. Time series stability plots for three USA cities, a) San Diego, CA b) Boston, MA, and c) Kansas City Airport, MO. Mean offsets in the north-south (gray) and east-west (black) directions are plotted as dots with standard deviations for each scene plotted as vertical bars. Note that the Kansas City Airport, MO plot has a different vertical scale from the other two plots.