DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof. Reference herein to any social initiative (including but not limited to Diversity, Equity, and Inclusion (DEI); Community Benefits Plans (CBP); Justice 40; etc.) is made by the Author independent of any current requirement by the United States Government and does not constitute or imply endorsement, recommendation, or support by the United States Government or any agency thereof.

ORNL/SPR-2025/3771

Maxar Initial Satellite Characterization

Orrin Thomas
Chris Tate
Kjell Ryerson
Hongwei Zhu

Approved for public release.
Distribution is unlimited.

June 2025

OAK RIDGE National Laboratory

ORNL IS MANAGED BY UT-BATTELLE LLC FOR THE US DEPARTMENT OF ENERGY

DOCUMENT AVAILABILITY

Online Access: US Department of Energy (DOE) reports produced after 1991 and a growing number of pre-1991 documents are available free via https://www.osti.gov/.

The public may also search the National Technical Information Service's National Technical Reports Library (NTRL) for reports not available in digital format.

DOE and DOE contractors should contact DOE's Office of Scientific and Technical Information (OSTI) for reports not currently available in digital format:

US Department of Energy
Office of Scientific and Technical Information
PO Box 62
Oak Ridge, TN 37831-0062
Telephone: (865) 576-8401
Fax: (865) 576-5728
Email: reports@osti.gov
Website: https://www.osti.gov/

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

ORNL/SPR-2025/3771

Geospatial Science and Human Security Division

MAXAR INITIAL SATELLITE CHARACTERIZATION

Orrin Thomas
Chris Tate
Kjell Ryerson
Hongwei Zhu

June 2025

Prepared by
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, TN 37831
managed by
UT-BATTELLE LLC
for the
US DEPARTMENT OF ENERGY
under contract DE-AC05-00OR22725

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ASV	Assessment Specification and Validation
CORS	Continuously Operating Reference Station
CPU	central processing unit
CSV	comma-separated values
ECEF	Earth-centered, Earth-fixed
ENU	East, North, Up
GNSS	Global Navigation Satellite System
GPU	graphics processing unit
HPC	high-performance computing
IC	Intelligence Community
IGS	International Global Navigation Satellite System (GNSS) Service
IID	independent and identically distributed
LSR	local space rectangular
NEON	National Ecological Observatory Network
ORNL	Oak Ridge National Laboratory
RAM	random-access memory
RMSE	root mean square error
SSECCOR	Satellite Systematic Error Calibration and Correction
VCE	Variance Component Estimation
WV01	WorldView-1
WV02	WorldView-2
WV03	WorldView-3

ABSTRACT

This documents reports on the initial Assessment Specification and Validation (ASV) of Maxar satellite imagery as a geolocation system. Validation is a prerequisite for using geolocation for applications that require quantified uncertainty. A Variance Component Estimation (VCE) calibration of the provider's covariance matrices concluded that they were exaggerated. However, we currently lack enough independent validation data to quantify the improvement from the VCE. The current ASV process is slow and expensive. It will not be able to keep up with growing numbers of satellite systems. We proposed a method to automate ASV using controlled imagery. The proposed method will allow us to keep up with new sensors and generate unbounded amounts of calibration data, leading to unprecedented accuracy.

1. INTRODUCTION

Sensor characterization is shorthand for the geolocation system Assessment Specification and Validation (ASV). "Assessment" refers to measuring the system's performance and is reported as the best estimate and least upper bound of the 90^{th} percentile of the error. "Specification" refers to publication of assessment results in official documents and software that allow a geolocation system to be used for specific applications. "Validation" tests the hypothesis that a system's published accuracy specification is accurate [6].

Describing accuracy and performance at the system level is fundamentally different from the surveying and mapping industry approach which uses independent validation data to measure each product's accuracy [1]. The Intelligence Community (IC) deals in country and global scale production and rapid refresh needs. Validation data is rarely available and wouldn't be practical for such wide extents or fine temporal scales; hence, the need to characterize measurement systems rather than individual products.

Geolocation system assessment is done using order statistics that do not require knowledge of the underlying probability distribution. Rather, it is assumed that the data is independent and identically distributed (IID), i.e., uncorrelated and drawn from the same probability distribution. In practice, there are guidelines for including some correlated samples, which are summarized in Table 1 [6, 27]. Hence, understanding these correlations is essential to planning the necessary control data.

Table 1. Approximate numbers of permissible samples per area and required minimum number of areas for validation [6, 27].

Correlation among ground points	Maximum Samples per area	Minimum number of areas
20%	20	5
50%	5	20
80%	3	34
95%	2	50

In addition to ASV, linear least squares Variance Component Estimation (VCE) (hereafter simply VCE) [29] is used as a calibration method to correct systematic errors in the input covariance information and thus improve the assessment. VCE models the measurement covariance matrix, Σ, as a linear combination of a set of given component matrices, Σ_i. Each component covariance matrix pertains to a subgroup of the measurements, and the VCE model allows for each component matrix to be in error by a positive scale factor s_i.

$$\Sigma = s_1\Sigma_1 + s_2\Sigma_2 + s_3\Sigma_3 + ...$$

These scale factors, s_i, are iteratively updated during processing until the weighted sum of the residuals for each measurement group equals the expected value, and the total weighted sum of the squared residuals is simultaneously minimized. The final scale factors are estimates of the bias in the input covariance components and are used in subsequent processing to achieve improved accuracy and error propagation. As the data used to perform the VCE contains control points and the subsequent processing typically will not, VCE can be thought of as a way to carry forward information learned from adjustments that had the benefit of control.

The measurements in our VCE were divided into four groups to be consistent with historical work:

- Sensor initial values
- Control point coordinates, e.g., from Continuously Operating Reference Station (CORS), National Ecological Observatory Network (NEON), or other sources.
- Manual image point observations of control points.
- Automatic image point observations of matching image features (i.e., 'autoties')

The measurement grouping is customizable, and optimum groupings are unknown. For example, we debated splitting the autoties into groups by area or based on their residual magnitudes. We also invented a way to mix correlated measurements into different VCE groups so that sensor parameters could be grouped arbitrarily.

The accuracy of the VCE scales depends on many factors (e.g., the validity of the IID assumption, the suitability of the math models, the accuracy of the component covariance matrices, absence of systematic errors, etc.). Among the sources of error, there is a hard theoretical limit. The weighted sum of squared residuals is modeled as following a chi-squared distribution, from which the expected value used to compute the scale is defined. The confidence interval of the expected value has a domain that depends solely on the redundancy: the higher the redundancy, the narrower the range, and, hence, the more confident the scale estimate. The scale estimates cannot be better than the limit implied by this confidence interval. For example, to achieve a 10% error bound *floor* on the scale requires a redundancy of just shy of 2,200.

As a calibration method, VCE is limited. First, it cannot correct systematic measurement errors (which is arguably the fundamental expectation of a 'calibration'). VCE also does not correct unmodeled or erroneous correlations. VCE only corrects systematic scale biases in the measurement's covariance data.

Our commercial ASV efforts began with Maxar's WorldView-1 (WV01), WorldView-2 (WV02), and WorldView-3 (WV03) sensors. Their high resolution, revisit rates, collection history beginning in 2007, and availability through existing contracts made them the obvious choice. Table 2 lists the active WorldView sensors, their launch dates, published accuracy information, and revisit times.

Table 2. Maxar sensors used in this work [19, 17].

Sensor	Launch Date	GSD (m)	CE90 (m)	LE90 (m)	Revisit Time (days)
WV01	2007	0.5	4.0	3.7	1.7
WV02	2009	0.46	3.5	3.6	1.1
WV03	2014	0.3	3.5	None Given	1.0

We anticipate Maxar remaining relevant due to the longevity of the active WorldView sensors and their new Legion constellation. Maxar launched the first four of six planned Legion satellites in 2024; the final two launches are planned for 2025. WorldView Legion satellites have a ground resolution of 0.3 m and revisit rates up to 15 times per day [18]. Coupled with currently available Maxar assets, this constellation

will ultimately triple Maxar's 0.3 m class imagery collection capability to over 6 million square km per day [18].

Our ASV work used Level 1B imagery products. These are unprojected products delivered with geometric sensor model parameters [5, 17]. Horizonal CE90 and vertical LE90 specifications for WorldView products are 5.0 m. [19, 17]. Previous validation work has shown that the CE90 and LE90 are slightly better than the product specifications (see Table 2). [19, 17]. The VCE work seeks to further improve the accuracy by correcting systematic errors in the provided covariance data.

2. INPUT DATA

Control data and multiple overlapping images that observed the control are required for VCE. In this case, the control data consisted of ground points with geodetic positions known to sufficient accuracy. Three sources of control data were used in this work CORS, NEON, and ground survey.

The CORS data is from the International Global Navigation Satellite System (GNSS) Service (IGS). The IGS is a global network of 500+ GNSS stations, which collect, archive, and distribute GNSS observation data sets for a wide range of applications [26]. For our purposes, we needed only the GNSS antenna coordinates. Hence, we used the most recent cumulative reprocessing of all the data called "repro3" [24]. The IGS absolute horizontal and vertical accuracy of geocentric station positions are given as less than or equal to 2 mm and 5 mm, respectively [26]. The stations used in this work are given in Table 3.

Table 3. IGS stations used in this work.

IGS Station	Latitude	Longitude
ALIC00AUS	-23.67011	133.88552
BAKO00IDN	-6.491055	106.84891
IITK00IND	26.5214214	80.2321334
PPPC00PHL	9.77293253	118.740234

NEON sites were a second source of control data. These sites are part of an observation network and data are collected over each site yearly, including high resolution lidar [21]. We used these lidar data to derive ground control points. First, we identified sites with sufficient non-transient ground features that could be identified in satellite imagery (e.g., building corners). These sites were the Blue River ("BLUE"), the Central Plains Experimental Range ("CPER"), the Marvin Klemme Range Research Station ("OAES"), and the San Joaquin Experimental Range ("SJER") (see Table 4).

NEON lidar data was collected from an airborne platform flying at 1000 m above ground level at 100 knots in flight lines with 37% overlap [12]. Direct georeferencing of each detected laser shot makes use of GPS data, IMU data, instrument mirror angle, range to target, and calibration information, including lever arms and boresight misalignments. Uncertainties within the point cloud positions are calculated by making use of the Law of General Propagation of Variances with the uncertainties associated with each subsystem measurement and the direct georeferencing equations [12]. NEON lidar point cloud accuracy requirements are stated as less than or equal to 0.58 m horizontal and 0.15 m vertical [12]. In practice, uncertainties on individual laser spots are typically much less than these requirements.

Table 4. NEON sites used in this work.

NEON Site	Latitude	Longitude
BLUE	34.444218	-96.6242
CPER	40.815536	-104.7456
OAES	35.410599	-99.05878
SJER	37.10878	-119.7323

We acquired all the WorldView-1, -2, and -3 panchromatic images (© 2024 Maxar) available at the time of query for each control site. All images whose ground features are not obscured by clouds or other image quality effects were used. The number of images used, as well as their date ranges, are given in Table 5.

Table 5. WorldView imagery associated with each control site.

Site	Number of Images	Date Range
ALIC00AUS	63	2008 - 2023
BAKO00IDN	7	2011 - 2022
BLUE	18	2010 - 2023
CPER	24	2009 - 2023
IITK00IND	35	2008 - 2023
OAES	21	2010 - 2022
PPPC00PHL	38	2012 - 2022
SJER	33	2009 - 2023

The number of image ranges from 7 to 63, with an average of 30. The temporal coverage ranges from 10 to 15 years, with an average of 13.

The eight sites with CORS and NEON control data were used for VCE. Four sites with ground survey control data were reserved to validate the VCE. They will be discussed in a classified addendum.

3. DATA PROCESSING

The simultaneous VCE requires matching points and ground control measurements for each site.

3.1 MEASURING IMAGE POINTS

The image point measurement process followed the steps below. Steps 1 through 9 are accomplished by the Oak Ridge National Laboratory (ORNL)-produced software package named "Suture".

Suture is written in C++ and published as a container so that it can be run on any hardware. We are running it on a single high-performance computing (HPC) node with 256 central processing unit (CPU) cores and two terabytes of random-access memory (RAM), and eight NVIDIA A100 GPUs. Graphics processing units (GPUs) are not required in the current version, but many of the modern matching algorithms are impractically slow without them.

1. **Each image has 14 regions targeted for matching (by default).** These area are near image edges to maximize coverage.
2. **Each image region is opened and resampled to have consistent scale and rotation.** Research has consistently shown that matchers perform better if they don't have to support scale and rotation invariance [23, 11, 10].
3. **Feature matching is done among chains of image pairs to create initial matches.** Currently this is done using upright SURF with minimal scale invariance from the OpenCV library [22]. Modern deep-learning-based matcher performance is being tested in related research project; from among those matchers, the R2D2 [25] and ASLFeat [16] algorithms were selected for implementation.
4. **Feature matches are winnowed to remove outliers.** The off-the-shelf tool for winnowing is RANSAC [7] algorithm that uses the 'eight-point algorithm'. The eight-point algorithm was derived for frame images and proved slow and insensitive for small tiles of satellite imagery. We invented a three-degree-of-freedom mathematical model (a five-dimension reduction) that is faster and correctly categorized 98% of inliers and outliers in a series of many thousands of simulations.
5. **Bundle adjust the inlier feature matches.** This bundle adjustment was local to the region being matched and used the typical 2D Affine image space correction model [28, 8, 2, 3, 30, 15]. Huber maximum likelihood estimation [13] was used to manage the influence of any higher residual points.
6. **Predicted missing matches.** The bundle results are used to predict missing matches. Every autotie with at least two inliers was used to predict matches through the whole stack of images.
7. **Least squares Matching was used to verify predicted matches and refine initial matches.** Least squares matching is well-known for sub-pixel matching and provides estimates of the matching accuracy as a standard error in pixels [20, 9, 14]. The initial matches were replaced with least square matching results only if the predicted quality of match was high enough (less than 0.2 pixels). The threshold for verification of a predicted match was higher (less than 0.4 pixels).
8. **The bundle adjustment was repeated using all the verified and refined ties from the least squares matching.**
9. **A user-specified number of ties in each region was saved.** To prevent an impractical flood of tie points, users specify the final number of points they want save in each image region (in this case, five). Points measured in the most images were preferred when choosing which few to save. Results were saved in SOCET GXP® file formats (specifically, IPF and GPF) because these were believed to be a stable and supported exchange format. We are also adding a GDAT-M comma-separated values (CSV) format.

10. **Ground control points were manually observed in all images.** This was done using the SOCET GXP® multi-sensor triangulation interface [4].
11. **The autoties and control measurements were merged into a single set of files.** This combined set of files can be read as input for the VCE bundle adjustment.

3.2 SIMULTANEOUS VARIANCE COMPONENT ESTIMATION

Satellite Systematic Error Calibration and Correction (SSECCOR) is the application that performs the simultaneous VCE. It was also developed in-house at ORNL. SSECCOR's name may seem grandiose considering VCE is its only current function, but the vision behind the ambitious name will be explained in the "Future Work" section.

The current version of SSECCOR doesn't require GPUs. However, the linear algebra systems can become large. Each area included in the processing adds about 7,000 parameters and a quarter million equations. Also, VCE requires a full matrix inversion, making it impractically slow on most desktop workstations for more than a few areas to be processed simultaneously.

The VCE processing workflow:

1. **Simultaneous bundle adjustment of autotied calibration areas from the same geolocation system.** The parameters are the sensor models and ground coordinates of each measured image features. The measurements are the initial values of the satellite sensor model parameters, the initial values of control points, and image space observations of control and tie points. Bundle adjustment tweaks the values of the parameters to minimize the weighted sum of the squared errors in the measurements.
2. **Scale each group of measurements' weight matrix so that the sum of the weighted squared errors equals the expected value.** For a group of measurements, i, with weight matrix W_i, and residual vector, V_i, the weighted sum of squares errors is $V_i^T W_i V_i^T$. $V_i^T W_i V_i^T$ follows a χ^2 distribution and so has an expected value $E(V_i^T P_i V_i^T) = \nu_i - 2$, where ν_i is the degrees of freedom and is equal to the redundancy contributed to the system by the group of measurements i. The scale correction is thus:
$$s_i = \frac{\nu_i - 2}{V_i^T W_i V_i^T}$$
3. **Repeat steps 1 and 2 until the scale updates are negligible.** The scaling of the component weight matrices (P_i) shifts the optimum solution to Step 1 (the simultaneous bundle adjustment). Adjusting the parameters in Step 1 in turn changes the residuals vectors, V_i, which requires $E(V_i^T W_i V_i^T) \neq \nu_i - 2$. Hence, Steps 1 and 2 iterate until the scale corrections in Step 2 are small enough to ignore.

SSECCOR outputs are saved as a reports. The native reports are simple text and CSV format, but we are working on Python scripts for better visualizations.

The pixel space reporting from the bundle adjustment attempts to represent the quality, quantity, distribution, and completeness of the image measurements. Some of this reporting is typical and some is more experimental.

The object space reporting deals with adjustment motion and errors on the ground surface. Each of these residuals is reported in a local space rectangular (LSR) East, North, Up (ENU) coordinate system. This choice was made because SSECCOR is intended to process vast amounts of data scattered all over the land mass of the planet, thus no single LSR would do, and we believed Earth-centered, Earth-fixed (ECEF) would be harder to interpret.

There are also some reports that are simple listings of the final adjusted sensor and ground parameters.

Individual reports:

- **The VCE report defines each of the measurement groups, their redundancy, final cumulative weight matrix scale factor, and confidence intervals of that scale factor.** These final cumulative weights scales, are the calibration parameters that can be used in subsequent adjustments.
- **The residuals report pertains to the autoties quality, quantity, distribution, and completeness.** The report begins with a summary of the combined residual distribution including the 1^{st} quartile, median, 3^{rd} quartile, 97^{th} percentile, and root mean square error (RMSE). Similar data is given image by image at the bottom. There is also a list of the 250 worst image measurements and a breakdown of the points by ray count. A complete list of all residuals is provided in a separate file in CSV format.
- **The control report lists the ground space movement of the control, and image space residuals of the control observations.** Motion in this report is in ENU local to each point. In addition to typical line and sample residuals, this report includes "Object Space Missclosure"—the distance of nearest approach from the observation ray to the final control point coordinates.
- **The ground point CSV file contains all the final ground space coordinates.** The coordinates are ECEF.
- **The sensor parameter CSV contains the final corrections to the sensor parameters.** This report is a simple CSV for ease in passing on the results for use in coordinate mensuration, orthorectification, mosaicking, or epipolar resampling.
- **The coverage report can be run before the bundle adjustment to assess the completeness of the autotie coverage.** This report begins with 2D "overlap matrices" for each separate area. The overlap matrices show the number of ties among all the images in the area. There are also individual image breakdowns that report success placing autoties in each of the 14 key regions around the image edges.

4. RESULTS

4.1 INPUT VALIDATION

Input validation verifies that the standard set of adjustment assumptions are reasonably satisfied. We expect our measurements to be free from blunders and have systematic errors corrected, and thus be approximately normally distributed. Recall that the satellite sensor model and ground control initial parameter values are modeled as measurements, and thus must also be verified.

We checked the input image sensor models and ground control by analyzing the the pre-adjustment control reports (pre-adjustment reports can be generated by setting the SSECCOR parameter "in.bundle.max.iterations" to zero). The expectation was that the image measurement residuals of control points would be symmetrically distributed around the origin and roughly equivalent among the areas. The pre-adjustment residuals of the ground control observations are shown in Figure 1. The distribution is free from large systematic errors and blunders.

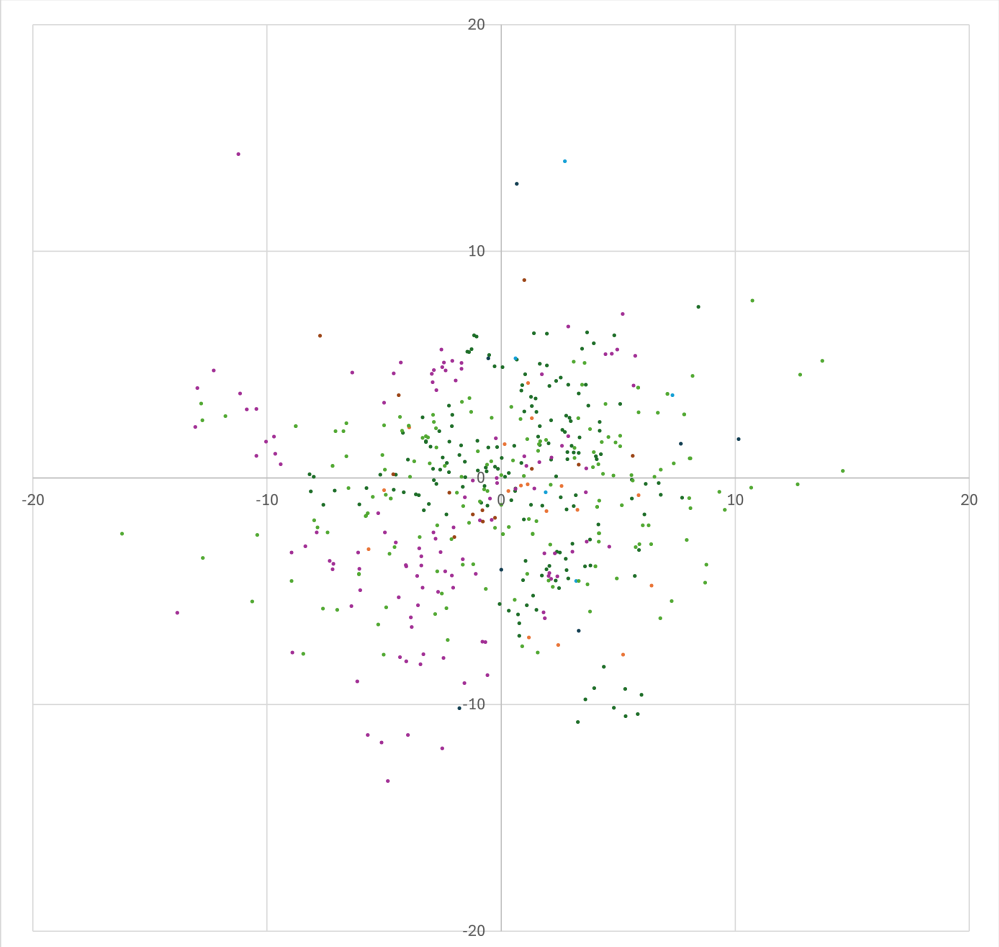

Figure 1. Pre-adjustment (i.e., input validation) residuals of image observations of control points in the eight study areas (pixels).

The control reports also have an "object space missclosure" for each observation of a control point. This is the linear distance (in meters) of closest approach from the control point coordinate to the 3-dimensional

ray implied by the image measurement. Thus, it is a way to translate the image space residuals from pixels to meters. The cumulative probability density function for the object space missclosures are given in Figure 2. The 90^{th} percentile of the 3D object space missclosure was 4.5 m. This is consistent with the published spec [17] and the operator's previous validation work [19], and thus, we concluded the sensor parameters and ground control coordinates for the study areas were ready.

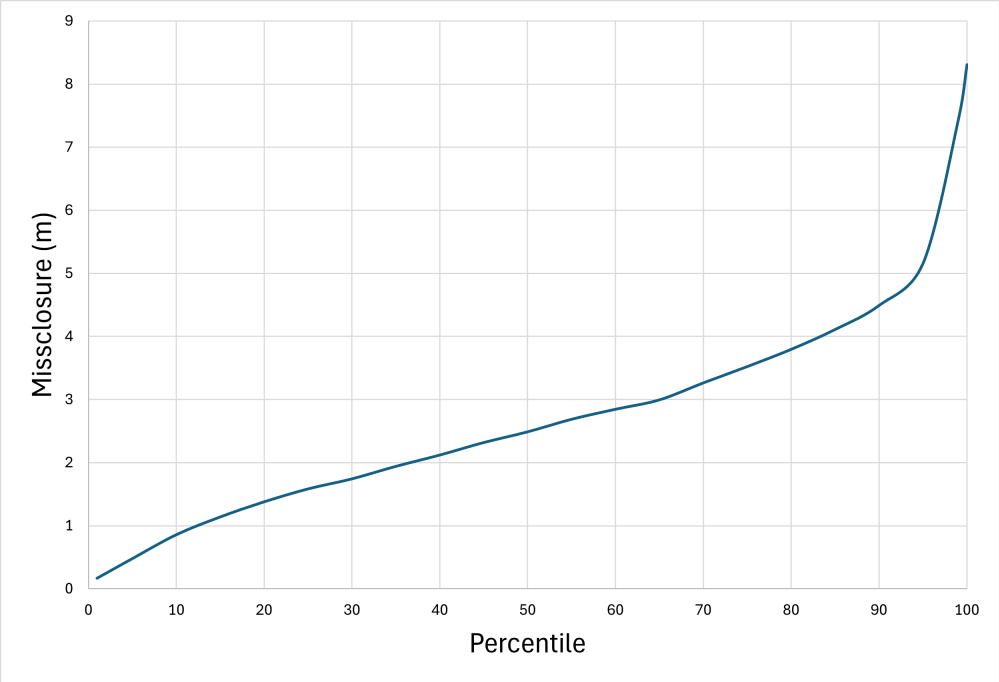

Figure 2. Pre-adjustment (i.e., input validation) distribution of object space missclosures (meters) of image observation of control points.

To validate the autoties coverage and completeness we generated coverage and residual reports for each area individually. The number of ties among images were given as matrices at the top of the coverage report. The data was summarized in Table 6. 4439 of 4483 (99%) overlapping image pairs were successfully autotied together. We have not yet investigated why CPER had the majority of failures (40 of the 44). The median number of ties ranged from 27 to 613. The number of ties tends to increase with the number of images, but this signal is complicated by how challenging the images are for the matching algorithm.

Table 6 also includes information about the the distribution of ground points observed in two images (2-ray points), observed in three images (3-ray points), and observed in more than three images. If it can be avoided, we do not want our adjustment dependent on two-ray points because they have a fundamental geometric weakness that points with more rays do not. There were no 2-ray points in any of the areas. The percentage of 3-ray points ranges from 4% to 68%.

The measurement completeness data in Figure 3 adds context to the ray count distribution information. BAKO00IDN had a lot of 3-ray points (compared to other projects), but it was also the most complete set of measurements. In contrast, PPPC00PHL had among the lowest percentage of 3-ray points, but also the worst completeness curve. In general, we caution against reading too much into Figure 3. It may not be reasonable to expect a measurement to be in dozens or hundreds of photos. 3-ray points are sufficient to link a set of images and can be generated much faster than deeper measure stacks. However, we predict that comparing the completeness curves among autoties generated with different versions of Suture and different autotie algorithms will help us track progress.

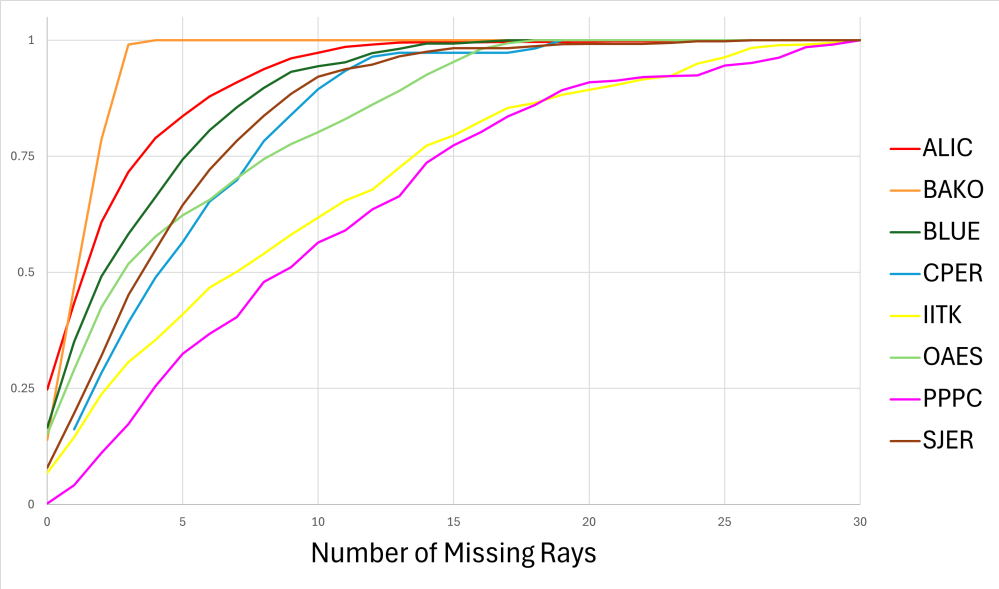

Figure 3. Cumulative distribution of missing ground point observations rays among the eight study areas.

Table 6. Common points among image pairs.

	Image Pairs		Number of Common Ties						Ray Counts		
Area	**Total**	**Failed**	**Mean**	**Min**	**Q1**	**Med**	**Q3**	**Max**	**2-Rays**	**3-Rays**	**More**
ALIC00AUS	1953	0	613	95	401	613	803	1415	0%	4%	96%
BAKO00IDN	21	1	43	0	18	27	78	104	0%	68%	32%
BLUE	231	0	550	13	113	166	250	550	0%	25%	75%
CPER	276	40	64	1	13	45	97	285	0%	30%	70%
IITK00IND	561	0	180	7	101	164	243	631	0%	37%	63%
OAES	210	0	288	18	197	299	373	624	0%	19%	81%
PPPC00PHL	703	3	69	0	37	58	90	271	0%	19%	81%
SJER	528	0	414	45	297	423	537	925	0%	20%	80%

The final input validation was to check the consistency of image measurements. For this, we ran bundle adjustments of the 8 areas individually. The distributions of the image measurement residuals are given in Table 7. In five areas, the 97^{th} percentile of 2D image residual vectors lengths was never more than 3.3 pixels. The median 2D residual vectors was no more that 0.7 pixels. The RMSE of single-axis residuals ranged form 0.62 to 1.46 pixels.

4.2 VARIANCE COMPONENT ESTIMATION

The VCE adjustments were run for each area individually, and in a combined simultaneous adjustment. The combined adjustment included 242 images, 10,202 ground points for a total of 35,446 parameters. There were 268,143 total equations. The processing time was about 8.5 hours, and peak memory usage was approximately 100 gigabytes. Our current compute node could run a system about 20 times as large without optimization. We note that the code base is young and largely unoptimized.

The Maxar sensor parameter weight scale 90% confidence interval was (0.20-0.21). This weight scale is the principal result of the VCE calibration that is carried forward and used in subsequent adjustments.

Table 7. Distributions of 2D image measurement vector lengths.

Site	count	Quartile 1	Median	Quartile 3	97 Percentile
ALIC00AUS	60623	0.3	0.4	0.7	1.5
BAKO00IDN	729	0.4	0.7	1.2	3.3
BLUE	977	0.4	0.6	1.0	2.2
CPER	3943	0.2	0.4	0.6	1.3
IITK00IND	16657	0.3	0.6	0.9	1.9
OAES	5768	0.3	0.4	0.7	1.4
PPPC00PHL	6028	0.3	0.5	0.8	1.6
SJER	28068	0.6	1.0	1.6	3.2

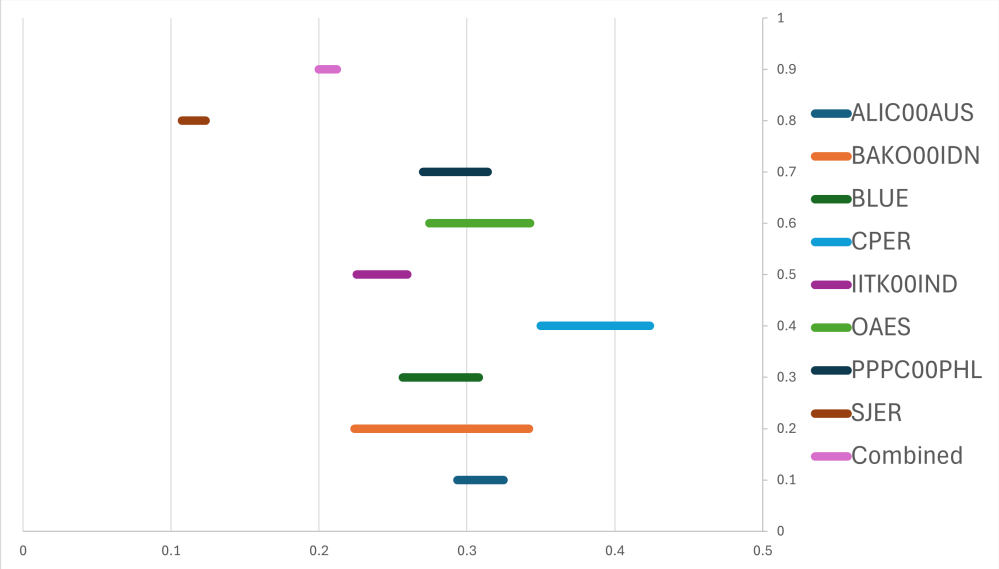

Figure 4. 90% confidence intervals of the sensor parameters weight scales.

The weight scale being less than 1.0 implies the covariance information provided by the sensor operator is exaggerated. This combined adjustment assigned all 20 of the sensor parameters in a single VCE parameter group. The the redundancy of the group was 1347.8—far less than the target of 2400.

Figure 4 shows a comparison of confidence intervals of the sensor weight scales computed in the individual and combined VCE adjustments. All of the adjustments agreed that the initial values were too heavily weighted (i.e., the weight scales are less than 1.0). There was reasonable agreement among seven of the areas (SJER was the most distinct). The combined adjustment was centered between the SJER and other seven areas. Recall that these are confidence interval floors—based on uncertainty in the expected values of the sum of weighted change in the sensor parameters. The lack of overlap among some of them implies the floor may be quite conservative.

The iterative changes to the control point weight scales during the combined VCE told an interesting story. When the the bundle first converged, the control moved by up to a few centers, but the predicted sum of weight squared error was below the expected value. Hence, the weights on the control initial values were increased. In this same adjustment, the sensor parameters errors were larger than expected and hence in the next bundle adjustment they were weighted less heavily. As a result of these two scales moving in opposite directions, the control point coordinates were adjusted less and sensor parameter slightly more in

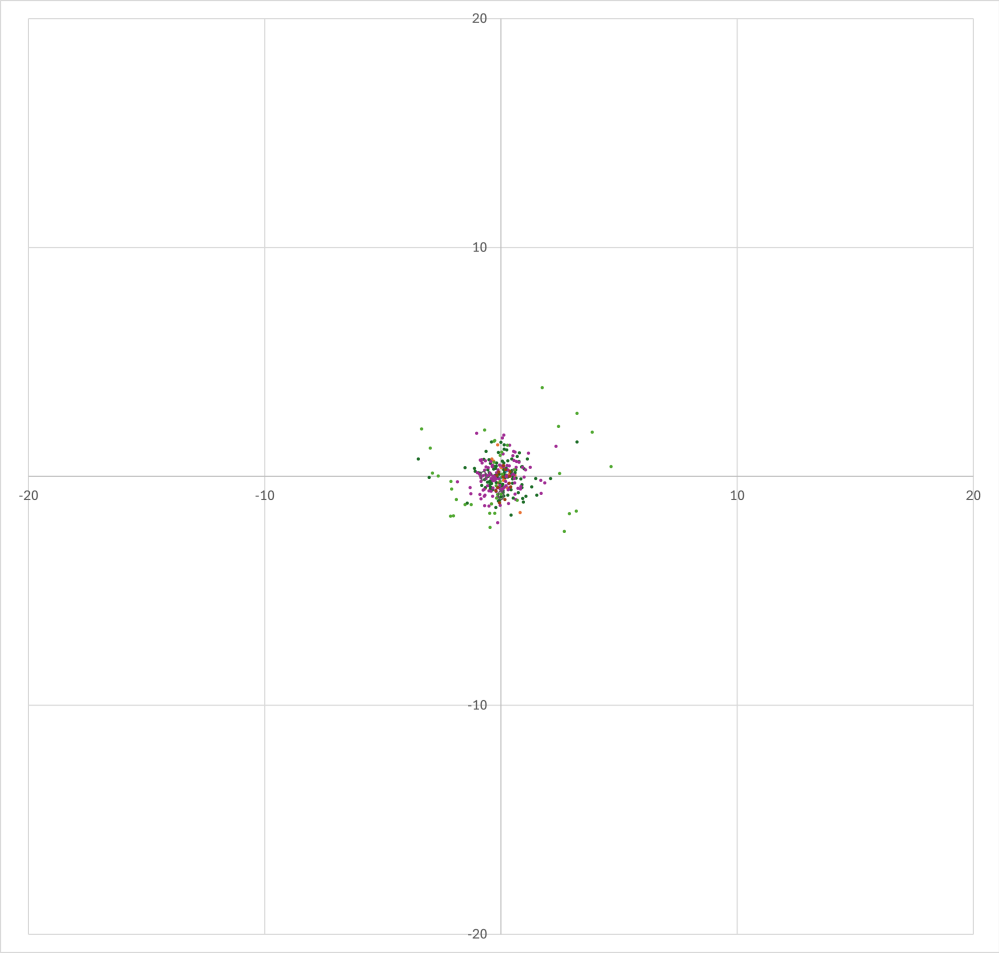

Figure 5. Post-adjustment residuals of image observation of control points in the eight study areas (pixels).

second bundle adjustment. This in turned caused the weight to be adjusted in same directions. This cycle repeated through seven adjustments before the senor parameter scale stabilized. At the end, the control was weighted so heavily that its motion was negligible, and the correction ratio calculation was becoming numerically ill-conditioned.

The short version of the ground control weight scale story is that the VCE iterated toward holding it constant. This is a reasonable result. It implied that the collection of satellite images with no better than 0.5 meter pixels cannot credibly update the coordinates of CORS or even to the lidar-derived NEON control. In future VCE, we recommend holding the ground control constant after rigorous input validation. The ground control corrections are not plotted because they were negligible (below the millimeter level).

The final residuals of image observations of control points are shown in Figures 5, 6, and 7. For ease of comparison, the axis scales in Figure 5 (post-adjustment residuals) are the same as Figure 1 (pre-adjustment residuals). In Figure 5 there are 11 points with residuals over 3 pixels that are obviously outside the cluster at the origin. The same data is plotted as cumulative distributions in Figure 6. In the cumulative distribution plot, the larger residuals appear as tails with greater magnitude than expected for a normal distribution fit to the central data. These tails indicate difficult to measure features. The initial and final object space missclosure are shown in Figure 7.

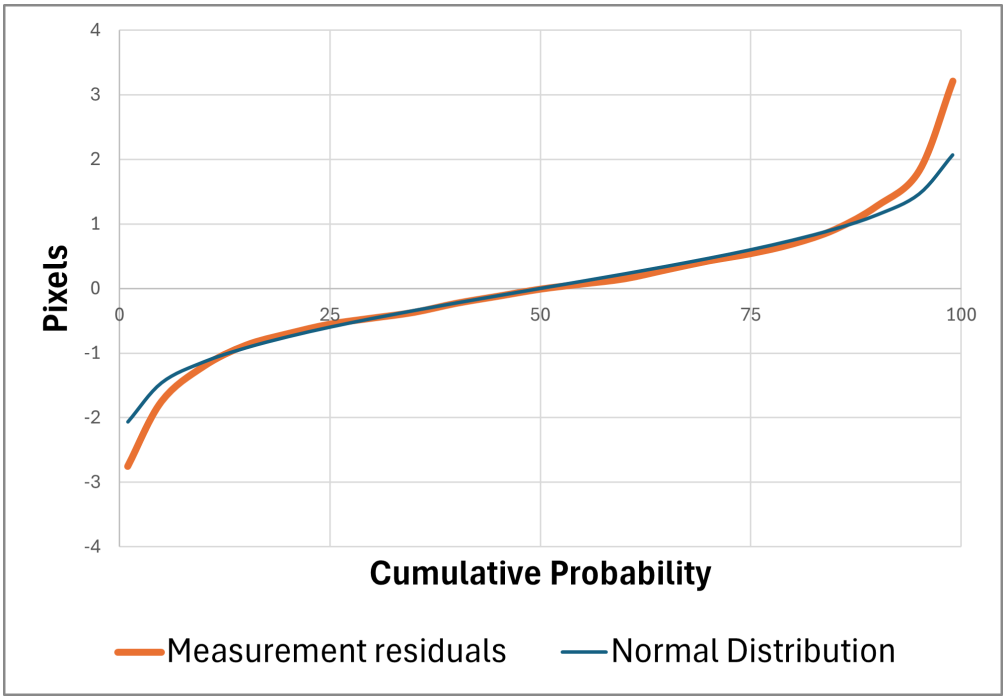

Figure 6. The cumulative probability curve of the control point observations residuals compared to a normal curve.

Note that the VCE adjustment's trend to hold the ground control as constants did not unreasonably increase the residuals of the control point observations. The standard deviation of the normal curve fit the central 80% of control point observation residuals 0.9 pixels (Figure 6). The 90^{th} percentile of the object space missclosure was 1.4 m, or no more than about 3 pixels (Figure 7).

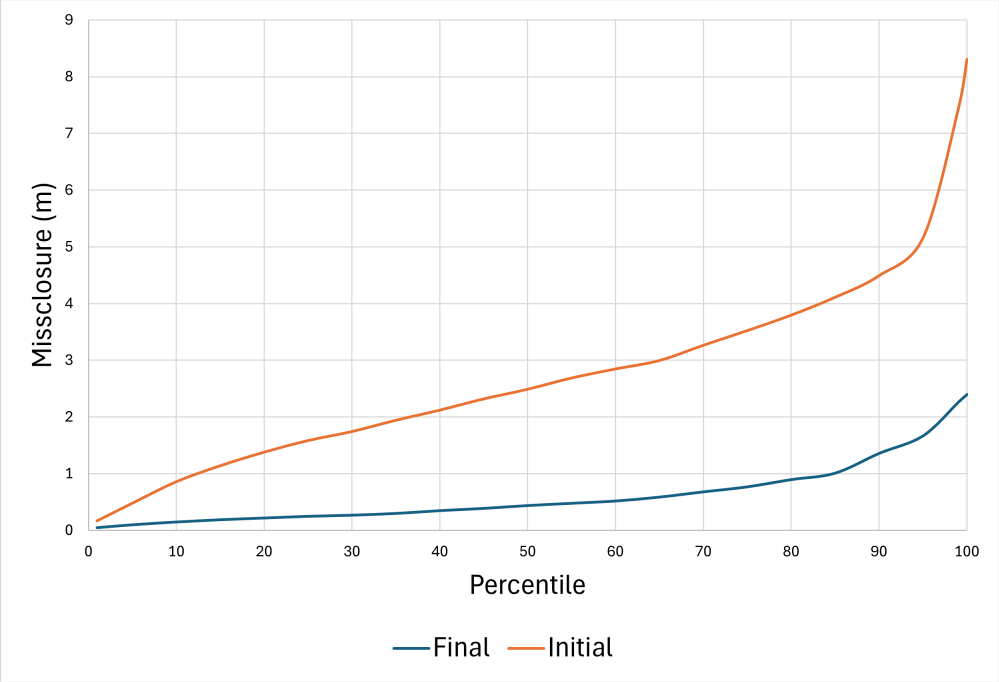

Figure 7. Initial and final cumulative distributions of object space missclosures.

5. FUTURE WORK

5.1 FUTURE VARIANCE COMPONENT ESTIMATIONS

The sensor model parameters of the 270 images contributed a total redundancy of 1347.8 to the system. This is not enough to begin splitting the parameters into VCE sub groups. However, it might be reasonable to only optimize a subset of parameters. Maxar's provided weight matrices have eleven parameters with weights less than 10^4 and nine with parameters with weights greater than 10^9. It is worth experimenting with modeling the heavily weighted parameters as constant. The autoties contributed about 150,000 to the redundancy. Each sub area was easily over 2200 and could be its own VCE group.

Experimentation with more and varied VCE groups or differing parameterizations cannot be validated or compared given our current data. More independent test data is required. Currently, we have only two sites reserved for evaluation and testing the VCE performance. Even with the low correlations among ground coordinates in individual sites, at least 3 more sites, with 20 control points each, are required for proper assessment [6] [27]. There are many hundreds more CORSs, and about a dozen more NEON sites. However, finding CORS whose GNSS receiver antennas are visible in satellite imagery is a slow manual process. NEON have a similar high manual cost to preparation; the areas have to be manually searched for features in the lidar that can be extracted as control points. In mid fiscal year 2024, the decision was made to pivot to procedure that promised automation and hence unbounded data.

5.2 AUTOMATIC ASSESSMENT VALIDATION AND SPECIFICATION

The science of automatic image tie point measurement is mature and widely successful. There is still a large number of manual hours required to identify and measure the control points in VCE and geolocation system validation. This too could be automated if we shifted from using control points to using controlled imagery.

Controlled imagery has many advantages; foremost, the same mature image matching technology can tie new geolocation system images to controlled imagery and thus automate all the measurements required for VCE and validation. Imagery also has a longer useful life than ground control. A control point may be obscured by vehicles, pedestrians, perspective lean, or shadows. They may also be destroyed by construction, erosion, or disaster. In contrast, controlled images cover large areas with many features, and this project has demonstrated that images can be reliably automatically matched even after a decade. Short, of cataclysmic disasters controlled imagery's usefully life span is open ended.

Controlled imagery is also significantly cheaper. Vexcel Imaging is a leading manufacturer of metric aerial cameras. They also regularly collect and control imagery over many human settles in the USA, Europe, Australia, Japan, and a few other areas. Their Bluesky product is built from metric aerial images with a 7 cm ground resolution. The published accuracy specification is 15 cm, which is sub-pixel for Maxar VCE and validation. The control aerial images can be purchased for dollars per square kilometer. We estimate that purchasing a hundred 25 km^2 areas of controlled imagery from Vexcel will cost about $\$10,000$. A survey crew would burn though that budget in a handful of days likely only collecting a single site.

We do not intend to take Vexcel's word regarding their product's quality; it needs to be validated using the same procedures as any other geolocation system. For this validation, we identified 100 sites where Vexcel's Bluesky product was collected over operational CORS around the world. We expect that Vexcel is actually understating their accuracy, but in the worst case scenario we'll control the Vexcel imagery to the local CORS stations.

5.3 SATELLITE SYSTEMATIC ERROR CALIBRATION AND CORRECTION (SSECCOR)

With automation and abundant cheap control information, it will become practical to generate unbounded amounts of unsupervised training data. With this wealth of data, we plan to truly calibrate geolocation systems. This will include learning how correlations among images decay over time and removing systematic errors. Among the uses for such through calibration is created a network of remotely sensed control points and global foundation geospatial data with unprecedented accuracy.

6. REFERENCES

[1] Qassim Abdullah. "The ASPRS Positional Accuracy Standards, Edition 2: The Geospatial Mapping Industry Guide to Best Practices". In: *Photogrammetric Engineering & Remote Sensing* 89.10 (2023), pp. 581–588.

[2] Mehmet Alkan et al. "Geometric accuracy and information content of WorldView-1 images". In: *Optical Engineering* 52.2 (2013), p. 026201. DOI: 10.1117/1.OE.52.2.026201. URL: https://doi.org/10.1117/1.OE.52.2.026201.

[3] Luigi Barazzetti et al. "GEOREFERENCING ACCURACY ANALYSIS OF A SINGLE WORLDVIEW-3 IMAGE COLLECTED OVER MILAN". In: *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLI-B1 (June 2016), pp. 429–434. DOI: 10.5194/isprs-archives-XLI-B1-429-2016.

[4] K De Venecia, Rick Racine, and A Stewart Walker. "End-to-end photogrammetry for non-professional photogrammetrists". In: *Conference Proceedings, 2005 ASPRS Annual Conference*. Citeseer. 2005, pp. 7–11.

[5] John Dolloff and Principal Scientist. "RPC UNCERTAINTY PARAMETERS: GENERATION, APPLICATION, AND EFFECTS". In: 2012. URL: https://api.semanticscholar.org/CorpusID:61558386.

[6] John Dolloff, Henry Theiss, and Brian Bollin. "Assessment, Specification, and Validation of a Geolocation System's Accuracy and Predicted Accuracy". In: *Photogrammetric Engineering & Remote Sensing* 90.3 (2024), pp. 157–168.

[7] Martin A. Fischler and Robert C. Bolles. "USAC: A Universal Framework for Random Sample Consensus". In: *Communications of the ACM* 24.6 (1981), pp. 381–395. DOI: https://doi.org/10.1145/358669.358692.

[8] Jacek Grodecki and Gene Dial. "Block Adjustment of High-Resolution Satellite Images Described by Rational Polynomials". In: *Photogrammetric Engineering Remote Sensing* 69.1 (2003), pp. 59–68. ISSN: 0099-1112. DOI: doi:10.14358/PERS.69.1.59.

[9] Armin Gruen. "Adaptive least squares correlation: a powerful image matching technique". In: *South African Journal of Photogrammetry, Remote Sensing and Cartography* 14.3 (1985), pp. 175–187.

[10] Jared Heinly, Enrique Dunn, and Jan-Michael Frahm. "Comparative Evaluation of Binary Features". In: *Computer Vision – ECCV 2012*. Ed. by Andrew Fitzgibbon et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 759–773. ISBN: 978-3-642-33709-3.

[11] Yuhe Jin et al. "Image Matching across Wide Baselines: From Paper to Practice". In: *CoRR* abs/2003.01587 (2020). arXiv: 2003.01587. URL: https://arxiv.org/abs/2003.01587.

[12] Keith Krause and Tristan Golden. *NEON L0 to L1 Discrete Return LiDAR Algorithm Theoretical Basis Document (ATBD)*. Tech. rep. NEON, Mar. 2022. URL: https://data.neonscience.org/data-products/DP1.30003.001.

[13] Wu Li and JJ Swetits. "The linear L 1 estimator and the Huber m-estimator". In: *SIAM Journal on Optimization* 8.2 (1998), pp. 457–475.

[14] Frank Liebold and Hans-Gerd Maas. "Computational Optimization of the 3D Least-Squares Matching Algorithm by Direct Calculation of Normal Equations". In: *Tomography* 8.2 (2022), pp. 760–777.

[15] Ana-Maria Loghin, Johannes Otepka-Schremmer, and Norbert Pfeifer. "Potential of Pléiades and WorldView-3 Tri-Stereo DSMs to Represent Heights of Small Isolated Objects". In: *Sensors* 20.9 (2020). ISSN: 1424-8220. DOI: 10.3390/s20092695. URL: https://www.mdpi.com/1424-8220/20/9/2695.

[16] Zixin Luo et al. “Aslfeat: Learning local features of accurate shape and localization”. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2020, pp. 6589–6598.
[17] Maxar. *Accuracy of WorldView Products*. Tech. rep. Maxar Technologies, May 2021.
[18] Maxar. *WorldView Legion*. Tech. rep. Maxar Technologies, Sept. 2024.
[19] Maxar. *WorldView-3*. Tech. rep. Maxar Technologies, Aug. 2020.
[20] Chris McGlone. “Basic Computer Vision Techniques”. In: *Manual of photogrammetry, 6th Edition*. Ed. by Chris McGlone. Bethesda, Maryland, USA: American Society for Photogrammetry and Remote Sensing, 2016. Chap. 6, pp. 559–564.
[21] National Ecological Observatory Network (NEON). *Discrete return LiDAR point cloud (DP1.30003.001)*. en. 2024. DOI: 10.48443/HJ77-KF64. URL: https://data.neonscience.org/data-products/DP1.30003.001/RELEASE-2024.
[22] Frazer K Noble. “Comparison of OpenCV’s feature detectors and feature matchers”. In: *2016 23rd International Conference on Mechatronics and Machine Vision in Practice (M2VIP)*. IEEE. 2016, pp. 1–6.
[23] Yuki Ono et al. “LF-Net: Learning local features from images”. In: *Advances in neural information processing systems* 31 (2018).
[24] Paul Rebischung et al. “The IGS contribution to ITRF2014”. In: *Journal of Geodesy* 90 (Apr. 2016). DOI: 10.1007/s00190-016-0897-6.
[25] Jerome Revaud et al. “R2D2: repeatable and reliable detector and descriptor”. In: *arXiv preprint arXiv:1906.06195* (2019).
[26] Internation GNSS Service. *IGS Products*. 2020. URL: https://igs.org/products/#about (visited on 12/23/2024).
[27] GWG World Geodetic System and Chair Geomatics Focus Group. *NGA STANDARDIZATION DOCUMENT Accuracy and Predicted Accuracy in the NSG: Specification and Validation Technical Guidance Document (TGD) 2c*. 2023. URL: https://nsgreg.nga.mil/doc/view?i=5601.
[28] Tee-Ann Teo. “Bias Compensation in a Rigorous Sensor Model and Rational Function Model for High-Resolution Satellite Images”. In: *Photogrammetric Engineering Remote Sensing* 77.12 (2011), pp. 1211–1220. ISSN: 0099-1112. DOI: doi:10.14358/PERS.77.12.1211.
[29] Peter JG Teunissen and AR Amiri-Simkooei. “Least-squares variance component estimation”. In: *Journal of geodesy* 82 (2008), pp. 65–82.
[30] Z. Xiao, B. Yang, and H. Zhang. “Positioning Accuracy Analysis and Application for Worldview-1 Stereo Imagery”. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-4 (2014), pp. 311–314. DOI: 10.5194/isprsarchives-XL-4-311-2014. URL: https://isprs-archives.copernicus.org/articles/XL-4/311/2014/.