Extending the Quality Standards for Non-Traditional Sensors: A Pathway to Increased Data Utilization

Steven Paligo
a.i. solutions, Inc.

ABSTRACT

The collective diversity and capability of space domain awareness (SDA) sensors is greater now than in the past. The existing data quality standards are inadequate to assess all the data those sensors currently generate in operation. The result is untrusted sensor observations, data underutilization, and potentially inaccurate results for SDA operators.

This paper describes, and proposes solutions to, three problems currently experienced by analysts performing data quality assessments of commercial SDA sensors. The goal is to convince policy makers to evolve the existing data quality standards so the SDA community can better leverage these new capabilities.

1. INTRODUCTION

Recent years have seen a significant shift in the space domain awareness (SDA) sensor landscape. Prior to this, space surveillance was dominated by militaries and research organizations. They used a small number of sensor types (primarily electro-optical, conventional radar, and phased arrays), each confined to a narrow set of configurations and use cases. For example, electro-optical (EO) sensors would be steerable, have high magnification, and narrow fields of view (FOV). They primarily tracked targets in deep space only.

Today's SDA sensor types, configurations, and data products are more varied. Many commercial organizations have emerged with SDA sensor offerings, and they perform space surveillance differently than it was done in the past. For example, passive radio frequency (RF) sensors gather data regardless of weather or lighting conditions, and they can effectively track emitting objects out to cislunar space. Short-wave infrared (SWIR) sensors can observe during local daylight [3]. Some EO sensors are being used for LEO tracking. Others perform wide-field surveys of LEO.

2. SENSOR QUALITY ASSESSMENT

2.1 Methodology

A key factor for determining an SDA sensor's quality is the accuracy of its astrometric ("metric") observations. A traditional methodology for assessing observation accuracy is a process that mimics sensor calibration. Specifically, a sensor tracks calibration satellites for which highly accurate reference ephemerides are available. Commonly, these objects include the International Laser Ranging Service (ILRS) satellites in LEO and the global navigation satellite system (GNSS) constellations in MEO and GEO. The differences between the reference ephemerides and the measured positions are determined. From those error residuals, noise and bias metrics are calculated and compared to established quality thresholds [2].

2.2 Sensor Quality Standards

Presumably many SDA data consumers adhere to established quality standards that define astrometric observation accuracy thresholds (though this is difficult to verify, as few standards documents are public). One public example is a 2019 Air Force Space Command Instruction (AFSPCI) titled, "Space Situational Awareness Metric Data Integration Guidelines for Non-Traditional Sensors" (AFSPCI 10-610) [1]. In it, they differentiate between "traditional" sensors (those confined to the legacy use cases and controlled by the U.S. Department of Defense) and "non-traditional" sensors (all others, including all commercial SDA sensors). They also define measurement-specific thresholds (e.g., less than 1 km range residuals sigma) for non-traditional sensors. Accuracy metrics must fall within those values for the data to be integrated into Air Force Space Command systems.

Although it is not explicitly stated, it appears the AFSPCI standards assume all sensors (both traditional and non-traditional) conform to the legacy use cases that existed before the surge in commercial SDA offerings. One evidence for that assumption is the 5 arcsecond angle threshold for angles-only sensors. At GEO distances, that translates to a cross-range error tolerance of approximately 867 meters. The standards also give a 360 arcsecond angle threshold for range-capable systems. At the 600-kilometer slant range that is explicitly stated in the document, it translates to a cross-range error tolerance of approximately 1,047 meters. From these two values, we see that Air Force Space Command was willing to accept approximately 1 kilometer of error from each angle measurement, regardless of orbit regime.

If we consider the case of an EO sensor tracking LEO objects and apply the AFSPCI standard's angles-only threshold, then it appears their angle error tolerance significantly drops to approximately 14.5 meters. That value appears to be an outlier when compared to the other tolerances, implying the thresholds do not consider this use case.

Another evidence that the AFSPCI standards are confined to the legacy use cases is that they makes no reference to passive RF sensors, SWIR sensors, or cislunar tracking.

Given the United States Air Force's long history in SDA and the incorporation of Air Force Space Command into the United States Space Force, it is reasonable to assume that many data consumers within the SDA community (including the Space Force) use similarly defined standards today. If so, then they lack guidance on how to evaluate the new class of commercial sensor capabilities. Without a standard, quality cannot be assessed, which in turn leads to a lack of trust and ultimately underutilization of available data.

3. REVISITING THE METHODOLOGIES AND STANDARDS

The intent of this paper is to persuade policy makers in the SDA community to revisit the methodologies and standards used for sensor quality assessment. It is assumed that they use standards similar to the Air Force Space Command document.

The following sections describe problems and solutions identified by *a.i. solutions, Inc.* while performing quality assessments for the United States Space Force's (USSF) Joint Commercial Operations (JCO) program, as well as other commercial organizations.

3.1 Expand the Number of Usable Satellites

The ILRS and GNSS calibration satellites are typically used for quality assessments because their accepted positions are known with sub-meter accuracy [5]. However, they do not represent the diversity of on-orbit objects in terms of orbit type, size/reflectivity, GEO longitude, etc. Additionally, there are relatively few of them, considering the large volume of space around Earth and the coverages of the commercial sensors. Table 1 lists the small quantities of ILRS and GNSS satellites available to *a.i. solutions, Inc.* for quality assessments.

Table 1: The number of ILRS & GNSS satellites used by *a.i. solutions, Inc.* for sensor quality assessments

Type	Orbit Regime	Quantity
ILRS	LEO	5
ILRS	Other	1
GNSS	MEO	109
GNSS	GEO (inclined)	13
GNSS	GEO (non-inclined)	12

Because of these limitations, analysts from *a.i. solutions, Inc.* have witnessed observability problems for some sensors during the quality assessment process. For example, wide-FOV electro-optical sensors that are configured to track LEO objects are often unable to observe the ILRS satellites. Although the satellites are designed to be highly reflective for laser ranging (see Fig. 1), their small sizes (less than 1 meter) make them very dim for optical observation.

Another example of limited observability is electro-optical sensors configured to stare at high interest objects in the GEO belt. Their high-magnification and narrow-FOV configurations, coupled with the small quantity of available GEO GNSS satellites, means they likely cannot consistently track any GEO GNSS satellites while maintaining custody of their observational targets. They must rely on serendipitous passes of MEO GNSS satellites through their view.

A solution to these two problems is for data consumers to expand what objects they consider acceptable for use in quality assessments. Starlink satellites are plentiful (6,290 at the time of writing) and are often bright, despite SpaceX's attempts at light pollution mitigation [4]. While they do make frequent orbit adjustments (affecting knowledge of their true positions), their abundant numbers mean plenty could be available for LEO-focused assessments after filtering those that are maneuvering.

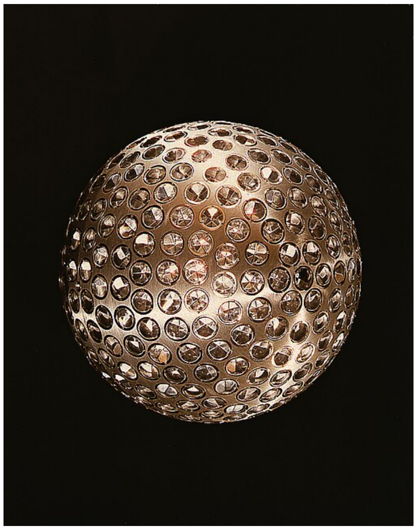

Fig. 1: LAGEOS-1 satellite (ILRS) designed to provide an on-orbit laser ranging target (NASA/Marshall, Public domain, via Wikimedia Commons)

Another option is to use defunct rocket bodies in graveyard orbits. They are non-maneuverable, often have densities that minimize solar radiation pressure effects, and are above the influence of atmospheric drag. It is possible to use observations of these rocket bodies from traditional well-calibrated sensors to generate accurate orbital states. Applying a high-fidelity propagation model to those states makes it possible to form reference ephemerides suitable for quality assessment of the non-traditional sensors.

Analysts from *a.i. solutions, Inc.* used special perturbations (SP) state vectors from the USSF 18th Space Defense Squadron to identify 61 rocket bodies whose predicted positions were sufficiently similar to subsequent states. This implies the rocket bodies' orbits are both stable and predictable enough for ephemeris generation. The result was a 44% increase in the number of calibration satellites suitable to assess GEO staring sensors, thereby increasing the probability of serendipitous tracks.

The Starlink satellites and defunct rocket bodies are just two possibilities to increase the number of trackable objects for quality assessments. Given the predictive nature of their reference ephemerides, they do not have the sub-meter accuracies of the traditional calibration satellites. Therefore, to accommodate cases like this, it is recommended to amend assessment report formats to support data provenance statements and confidence metrics. That way data consumers can understand the processes used to generate assessment results and the associated errors present in them.

3.2 Standardize Thresholds

The AFSPCI standards define quality thresholds in measurement-specific terms. For example, angles have tolerances specified in arcseconds and degrees. Range thresholds are given in kilometers and time thresholds are given in seconds. However, the document justifies some of those tolerances in terms of their contribution to positional error, implying that positional error is the actual threshold to measure against [1].

The second proposal is to redefine all thresholds in terms of positional error. That is to say, each measurement (time, range, angles, etc.) cannot yield more than a specified amount of positional error. However, it should refrain from enumerating a finite list of measurement types.

Some benefits of this approach would be:

1. It becomes easier to compare the relative quality of sensors that have different measurement types. This is a current problem in operations.
2. It eases the adoption of new sensor and measurement types because the standards would not have to be amended for each one. It would also immediately fix the problem of measurement types currently in operation that are not supported by the document (e.g., difference of arrival).
3. It decouples the sensor and measurement types from the orbital regimes
4. It allows operators to work in familiar measurements and units (i.e., distances and kilometers). For example, it

is unintuitive to evaluate the effect on satellite position of a time difference of arrival (TDOA) error measured in microseconds.

Because some quality thresholds are justified in terms of the contribution to positional error, this proposal seems like a common sense evolution of the definitions. However, care should be taken to make the thresholds scalable as the community expands its sights into cislunar SDA.

3.3 Allow Concurrent Quality Assessments

Sometimes a commercial sensor that is configured for a specific purpose unintentionally collects good quality observations of other objects. An example is the EO sensors that stare at subsections of the GEO belt. They are focused to observe objects at GEO distances, but they occasionally observe MEO objects that pass through their fields of view. Since the sensors are focused for a different distance, the MEO observations may be blurry or dim, resulting in noisier observations. Also, if their data processing pipelines are optimized for points of light from GEO objects, they may have difficulty with the streaks of light from MEO objects (see Fig. 2).

The traditional quality assessment process assumes a single set of sensor-wide noise and bias metrics. Therefore, when applied to this example, it restricts operation to one of two suboptimal choices. The first is to discard the MEO data. It could be argued that this is a logical choice since the sensor is not optimized for MEO targets. However, using recent observations from commercial sensors matching this example, analysts from *a.i. solutions, Inc.* found this course of action would discard 5.4% of the sensors' observations. Policy makers would have to decide if this is an acceptable loss. By inspection, analysts have seen that the error residuals from the MEO objects usually fall within the established AFSPCI thresholds, so they should not be carelessly dismissed.

The second choice is what some analysts use in commercial operations today. That is, combining both the MEO and GEO data to calculate a single set of quality assessment metrics for the sensor as a whole. This avoids discarding data, but it mixes data that perhaps should not be.

Figure 3 illustrates this. It shows the results of calculating the declination bias in different ways for 15 EO sensors that observed both GEO and serendipitous MEO objects from 1-14 August 2024. For each sensor, bias values were calculated using MEO-only data, GEO-only data, and the combined MEO and GEO data. Two facts stand out:

1. In all cases, the MEO-only values are significantly different from the GEO-only values. That shows the two data sets have different statistical properties.
2. Some sensors (specifically numbers 7, 8, and 15) have a combined bias that differs significantly from both the MEO and GEO values. That shows the bias generated from combined data may not be a good representation of either data set. When that happens, the decision not to discard the MEO data has negatively impacted the sensor's primary use case (GEO staring).

The overall impact of combining the MEO and GEO data to calculate sensor-wide metrics is that the results may not accurately reflect the properties of the sensor or its observations. When that occurs, data consumers would be subtracting out incorrect bias values from the observations or incorrectly weighting them in their orbit determination (OD)/differential correction (DC) processes. The net effect would be inaccurate orbital states of high interest objects.

Fig. 2: An image from an EO sensor showing a serendipitous track of a MEO object (top left) while attempting to track a GEO object (center).

Instead of having to choose between those two options, the third proposal is to:

1. Stop relying on a single set of sensor-wide metrics and allow the quality assessment process to generate multiple sets of metrics that apply concurrently, and
2. Change the quality metrics reporting formats so they instruct data consumers when each set of metrics apply.

For this example of GEO staring sensors with serendipitous MEO tracks, one set of quality metrics would be generated for the GEO observations and one for the MEO observations. However, the same idea could be used for sensors with different operational modes, side-by-side comparisons of experimental data processing pipelines, etc.

These updates would allow data consumers to generate more accurate data products from the sensor observations and lead to better informed SDA operators.

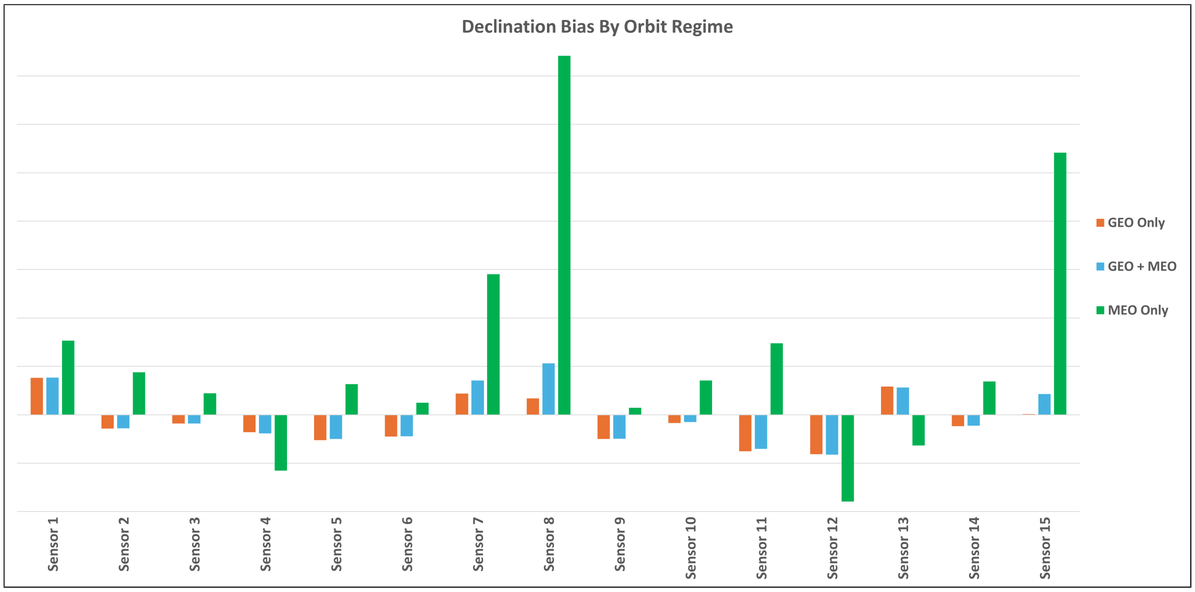

Fig. 3: The results of calculating sensor declination biases using different subsets of observations (GEO-only, GEO+MEO, and MEO-only). The sensor names are anonymized and the Y-axis (bias magnitude) is omitted to protect the commercial sensor providers' proprietary information.

4. CONCLUSION

Commercial organizations are bringing new capabilities to the SDA community. We need policy makers to ensure sensor data quality standards evolve to keep pace with them. Otherwise, the sensors may suffer from a lack of trust and be underutilized by traditional SDA data consumers. By expanding the definition of what is an acceptable assessment satellite, standardizing the accuracy thresholds into common terms, and allowing concurrent quality assessments per sensor, the standards-making bodies can ensure that military, civil, and commercial operators maximize the impact and effectiveness of the resources we spend on SDA.

REFERENCES

[1] Air Force Space Command. Space Situational Awareness Metric Data Integration Guidelines for Non-Traditional Sensors. `https://static.e-publishing.af.mil/production/1/afspc/publication/afspci10-610/afspci10-610.pdf`, 2019.

[2] Antón et al. Data curation activities for Space Surveillance and Tracking. `https://amostech.com/TechnicalPapers/2023/Poster/Anton.pdf`, 2023.

[3] Shaddix et al. Daytime Optical Contributions Toward Timely Space Domain Awareness in Low Earth Orbit. `https://amostech.com/TechnicalPapers/2021/SSA-SDA/Shaddix.pdf`, 2021.

[4] Grace Halferty, Vishnu Reddy, Tanner Campbell, Adam Battle, and Roberto Furfaro. Photometric characterization and trajectory accuracy of Starlink satellites: implications for ground-based astronomical surveys. *Monthly Notices of the Royal Astronomical Society*, 516(1):1502–1508, 2022.

[5] Peter Steigenberger and Oliver Montenbruck. Consistency of MGEX Orbit and Clock Products. *Engineering*, 6(8):898–903, 2020.