Data curation activities for Space Surveillance and Tracking

Alfredo M. Antón
GMV, Spain

Jose Miguel Lozano
GMV, USA

Keiran McNally
GMV, Airspeed, 2, Harwell Science and Innovation Campus, Oxfordshire United Kingdom,

Sean C. Hannon
GMV, France,

ABSTRACT

Data curation is the concept that involves data quality assessment on SST systems. Data curation uses well known methodologies such as sensor calibration, observations correlation or orbit comparisons to provide a continuous monitoring, characterization and anomaly detection mechanism on the different elements composing the system, based on the data available. The presented methodologies can be either based on precise and trusted sources, such as GNSS satellites, laser calibration spheres, altimetry or geodesy satellites, etc., or based on statistical comparison of different (less accurate), sources of orbital data, such as the 18 SDS special perturbation (SP) catalogue. The first approach allows the accurate characterization of sensors and processes in terms of biases and noise, while the second approach allows anomaly detection, either in specific objects, either in the sensor or orbital information provider.

1. INTRODUCTION

Together with the increase of the number of man-made objects launched in the last few decades and the associated space debris generated by in orbit fragmentations, the concern of the space community has also grown in the same direction, giving a great importance to the Space Surveillance and Tracking (SST) systems used to protect space assets. In the last decade, hundreds of sensors have been deployed or repurposed from the scientific domain to answer the high demand of data required by these systems to fulfil their mission.

The availability of large amounts of data, in form of sensor observations or orbital information, represents great news for SST systems, however it opens a new discussion on the quality of the data received and the need for monitoring and validation processes to ensure this quality. The concept of data curation represents all the processes and tools required to ensure the quality of the data ingested or produced by the SST systems.

This paper presents several concepts and methodologies to perform the data curation activities on observational data from SST sensors and derived products of a SST system. These concepts and methodologies are applied to the ***FocuSST*** service, the GMV's commercial service for SST.

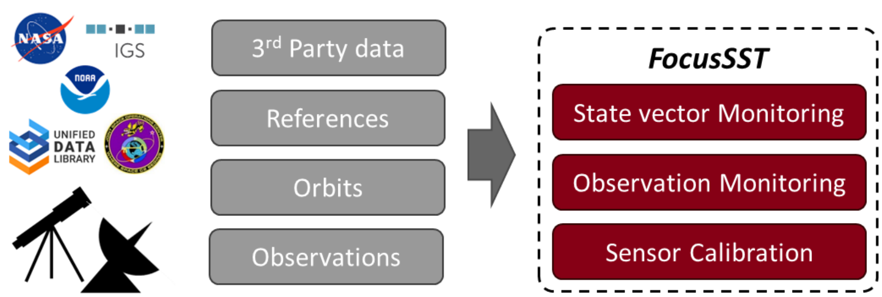

Figure 1: Data curation activities concept

2. PRECISION AND ACCURACY

Before discussing the data curation concepts and methodology is important to understand the concepts behind precision and accuracy. *Precision* is defined based on the observations themselves, defining statistical metrics accounting only for the distribution of these observations. On the other hand, *accuracy* refers to the proximity of those observations to the actual value of the magnitude observed, therefore the need of a validated and trusted reference to be compared is mandatory. In the space domain this reference is usually called "*ground truth*", resulting from on-ground processing considering the most precise models (dynamical and observational models) and data (GNSS, altimetry, laser ranging…) available. The data curation activities need to focus on both concepts, precision, evaluating the distribution of the observations themselves, and accuracy, validating the proximity of those observations to the measured magnitude.

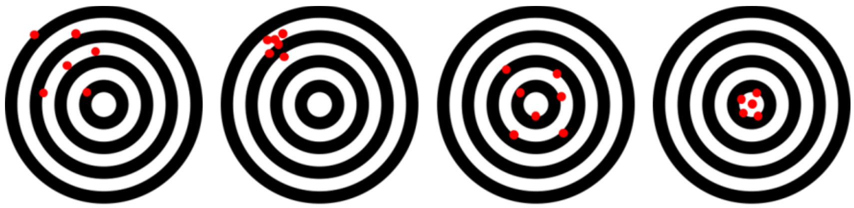

Figure 2: Precision and Accuracy

In addition to those two concepts, to describe some data curation methodologies, additional basic mathematical definitions are required:

- *Mean* (**μ**): Arithmetic mean, represents the statistically averaged value, computed by the sum of all the values divided by the number of values.
- *Standard deviation* (**σ**): Represents the distribution of the observations around the mean, providing a measure of the precision concept.
- *Bias*: Represents the systematic error of the observations against the reference and it is computed as the average of the deviations.
- *RMS* (Root Mean Square): Also known as geometric mean and it is computed as the square root of the mean of the squared elements of the distribution or function. This metric is commonly used in all engineering field, in particular, in data curation processes, it is usually applied to the observation residuals, difference between each observation and the reference value, as it provides a good measure of the composed precision and accuracy.

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2} \qquad Bias = \frac{1}{n}\sum_{i=1}^{n}(x_i - z_i) \qquad RMS = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - z_i)^2}$$

Where n is the number of samples, x_i each value and z_i the actual value of the observed magnitude.

3. CALIBRATION DATA SOURCES

As mentioned in previous section, in order to establish the accuracy of an observed magnitude, it is required to count on a trusted reference. This kind of information is not always publicly available, however there are some agencies, institutions and companies which provide regularly orbital data accurate enough to be used. Table 1 shows some public references that can be used for data curation services.

Table 1: Publicly available trusted data sources

Source	Link
Broadcast ephemerides for GNSS constellations	https://cddis.nasa.gov/Data_and_Derived_Products/GNSS/broadcast_ephemeris_data.html https://igs.org/mgex/data-products/#bce
Precise orbits for GNSS constellations	https://cddis.nasa.gov/Data_and_Derived_Products/GNSS/orbit_products.html https://igs.org/products/#orbits_clocks https://igs.org/mgex/data-products/#orbit_clock
Precise orbits for DORIS satellites	https://cddis.nasa.gov/Data_and_Derived_Products/DORIS/doris_idsorbit.html
Precise orbits for IRLS satellites	https://cddis.nasa.gov/Data_and_Derived_Products/SLR/Precise_orbits.html
Precise orbits for ESA's Sentinels satellites	https://scihub.copernicus.eu/

In addition to these public references, depending on the customer, GMV's ***FocusSST*** service could also make use of:

- ***FocusOC*** (https://www.gmv.com/en-es/products/space/focusoc) special perturbation ephemerides enhanced from 18 SDS through data sharing agreement.
- Precise orbits for GNSS constellations computed by the **GMV *GSharp*** [1] service (https://www.gmv.com/en-es/products/space/gmv-gsharp) combining public information and collected data from GMV owned stations.
- Precise orbits from commercial and institutional partners based on on-ground solution using GNSS and technologies.

4. SENSOR CALIBRATION

The goal of the sensor calibration activity inside the data curation concept is to correctly characterize the capabilities and performances of the sensor in terms of precision and accuracy. This goal is achieved through the observation of satellites which their orbits are precise and accurately known, such as the ones from the sources presented in previous section. The calibration process performs the estimation of biases and the characterization of the noise of the observations generated by the sensors, which are fundamental information for the adequate execution of the orbit determination processes.

The calibration process in ***FocusSST*** consists of the usual batch least squares algorithm applied on the observations received from the calibration satellites by considering the orbit fixed and only estimating the required sensor parameters. The algorithms are derived from the fundaments proposed by [2] and the implementation is shared with the Precise Orbit Determination (POD) service by GMV ***FocusPOD*** [3] operationally used by the ESA's Copernicus Programme [4]. This implementation includes propagation and observation models providing up-to centimeter precision and the possibility of estimating any model parameter, in particular the observation model ones which are in the interest for the sensor calibration application. In order to use this implementation in the SST domain, several additional observation models such as the telescopes, radars and passive ranging have been introduced on this development.

Table 2 contains the sensor types and observations are currently considered, including corrections and estimated parameters:

Table 2: Sensors and corrections considered in *FocusSST*

Sensor Type	Subtype	Applied corrections	Estimated Parameter
Telescope		Annual light aberration Diurnal light aberration Light travel time	Time bias Right Ascension bias & noise Declination bias & noise
Radar	Mono-static Bi-static	Tropospheric delay and refraction Ionospheric delay Light travel time	Range bias and noise Range rate bias and noise Azimuth bias and noise Elevation bias and noise
Laser		Tropospheric delay Light travel time	Range bias and noise
Passive Ranging		Tropospheric delay Ionospheric delay Light travel time	TDoA bias and noise FDoA bias and noise

Figure 3 below shows an example of the raw output from the computational algorithm.

```
Transponders Statistics. Number of transponders    5
              |-------RightAs (millideg)-------|-------Declina (millideg)-------|
 SAT-ID  TRAN |    #Obs/#Reject    Mean     RMS|    #Obs/#Reject    Mean     RMS|
 -------------+--------------------------------+--------------------------------+
 41175   MASS |     459/      0   0.091   0.133|     459/      0   0.092   0.151|
 41550   MASS |      61/      0   0.108   0.153|      61/      0   0.081   0.126|
 41861   MASS |     272/      0   0.090   0.116|     272/      0   0.114   0.148|
 41862   MASS |     194/      0   0.202   0.259|     194/      0   0.085   0.184|
 43056   MASS |      38/      0   0.139   0.175|      38/      0   0.103   0.148|
 -------------+--------------------------------+--------------------------------+

 Station Statistics. Number of stations    1
         |-------RightAs (millideg)-------|-------Declina (millideg)-------|
    ID   |    #Obs/#Reject    Mean     RMS|    #Obs/#Reject    Mean     RMS|
 --------+--------------------------------+--------------------------------+
  SENSOR |    1024/      0   0.114   0.163|    1024/      0   0.096   0.156|
 --------+--------------------------------+--------------------------------+
```

Figure 3: Raw sensor calibration output

The calibration activity has been usually performed in SST domain after every sensor maintenance or upgrade, or by means of periodic calibrations, assuming the sensors systems are stable enough to not require more often calibrations. The reality showed that clock synchronization issues, shutter delays and many other problems could appear at any time without being easily noticed from the observations collected. A small dedication to collect observations from the designated calibration satellites, e.g. 3 passes or slots per day, is enough to maintain the sensors correctly calibrated and characterized.

Currently, ***FocusSST*** performs three calibrations considering different intervals at each received track, first one considering only the received track, a second one including the last 24h of observations and finally a third one with the last 14 days of observations. These three executions provide at the same time a stable and well averaged result together with more reactive results. Results from each execution are stored into a database in form of time series that are lately represented visually by means of Grafana dashboards (Figure 4) where the alarms are automatically raised to the operator and the operator can then perform the analysis evaluating the different trends and possible discrepancies or anomalies.

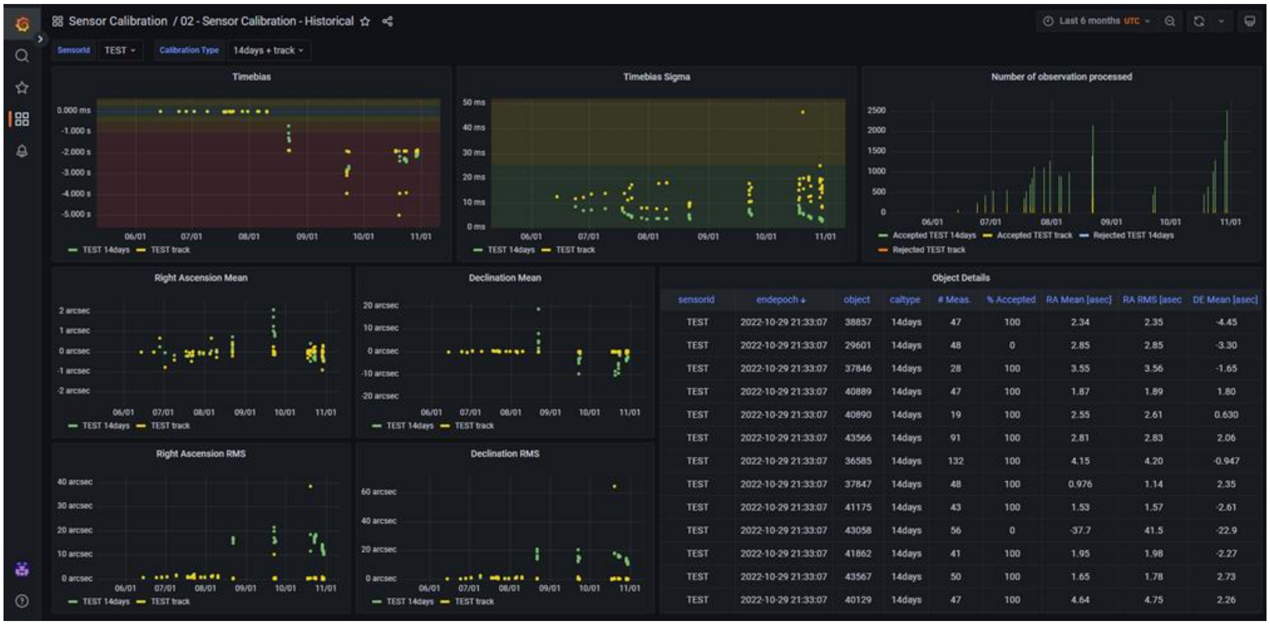

Figure 4: Example of visualization for calibration results (Grafana)

5. OBSERVATIONS AND STATE VECTORS MONITORING

In addition to the sensor calibration, another process in data curation activity consists of the monitoring of the different data and products received or produced by the SST System. Among this data, ***FocusSST*** current activities focus on the monitoring of observations and state vectors, however it could be extended to other type of data, such as conjunctions or maneuvers.

In case of the observations, the overall process is based on the computation of the residuals against a known reference, which can be a previous orbit determination or an external orbit, e.g. calibration satellites precise orbits or 18 SDS SP catalogue. The residuals are computed in a similar manner to any orbit determination process, by applying the measurement model on the reference orbital information, the mean, sigma, and RMS for each set of observations can be obtained and stored in a database for a later visualization. These residuals can also be normalized by using the Mahalanobis distance [6] & [7] concept, considering the calibrated noise from the sensor, and therefore providing a weighted evaluation of the error of the observations for each specific sensor.
This process, although not valid to characterize or calibrate the sensor due to the lesser accuracy of the reference orbits, is still very valuable to detect possible discrepancies or anomalies in sensors in near real-time, in particular when comparing observations from several sensors, or to detect changes in a specific object due to, for example, a maneuver or a propellant leak.

The other data curation activity performed is focused on the monitoring of the received or produced state vectors from the orbit determination process. The process is based on the computation of the orbital differences between the received state vectors and a reference orbit as before. The computation of the differences is performed in different coordinate systems and frames, for example, ***FocusSST*** current implementation considers cartesian differences in local orbital frame (LOF), keplerian and equinoctial in earth centered frame (GCRF), performing the required conversions in each case. In case the covariances are also present for both the analyzed state vector and the refence orbit, the Bhattacharyya distance [8] can be used, which includes the dependency of the covariance on the computed difference between the solutions. Similarly to the observation monitoring, the computed differences are stored in a database and used to present to the operator the results in several dashboards. The main dashboard shows the temporally accumulated values of the differences for different parameters and for each one of the satellites, such as the RMS for the cartesian LOF differences, or the semi-major axis, eccentricity, inclination and longitude. This dashboard can be setup with alarms which can be quickly reviewed by the operators and a more detailed analysis can be performed on the temporal evolution of the different computations.

6. IMPLEMENTATION

Finally, a brief description of the design and implementation of the ***FocusSST*** data curation service is presented in this section.

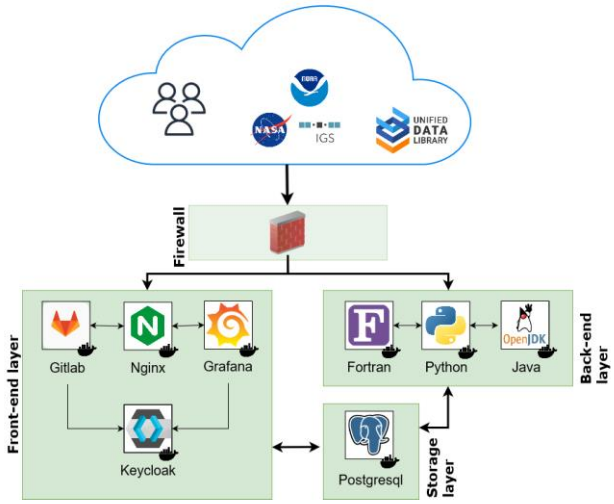

Figure 5: *FocusSST* data curation service high-level design

The overall system has been designed based on services implemented through a docker containerization which allows for horizontal scalability and high availability.
The system consists of three functional layers:

- **Front-end**: Containing the services which provide the operators and users with the monitoring dashboards based on Grafana and the issues tracking based on Gitlab. Both services are supported by a single sign-on service based on Keycloak.
- **Back-end**: Containing the computation services, based on several technologies such as Fortran, C++, Java and Python.
- **Storage**: Containing the databases based on the relational database technology Postgresql.

7. CONCLUSIONS

Data curation services represents a mandatory activity to ensure the quality of the data before and after being processed by the SST System. Data curation activities have been presented, describing calibration sources, processes and algorithms and the implemented architecture, together with his implementation in ***FocusSST***.

FocusSST data curation service has been tested successfully in the Commercial *Sprint Advanced Concept Training* (SACT). Currently it is also providing data curation services to the JTF-SD *Commercial Operations Cell* (JCO), processing sensor observations for telescopes, radars and passive ranging and state vectors for several providers.

8. REFERENCES

[1] D. Calle, L. Martínez, G. Tobías, GMV, Spain, ***magicGNSS*** Precise Product Provision for LEO POD Applications. IONGNSS 2020

[2] O. Motenbruck, E. Gill. Satellite Orbits - Models, Methods and Applications. Springer Berlin, Heidelberg https://doi.org/10.1007/978-3-642-58351-3

[3] Fernández Martín, C., Berzosa Molina, J., Bao Cheng, L., Muñoz de la Torre, M. Á., Fernández Usón, M., Lara Espinosa, S., Terradillos Estévez, E., Fernández Sánchez, J., Peter, H., Féménias, P., and Nogueira Loddo, C.:

FocusPOD, the new POD SW used at CPOD Service, EGU General Assembly 2023, Vienna, Austria, 24–28 Apr 2023, EGU23-1908, https://doi.org/10.5194/egusphere-egu23-1908, 2023.

[4] Fernández Sánchez, Jaime & Férnandez, Marc & Peter, Heike & Femenias, Pierre. (2022). Copernicus POD Service -Overview and Status.

[5] P. Mahalanobis, "On tests and measures of group divergence I. Theoretical formulae" *J. and Proc. Asiat. Soc. of Bengal* , 26 (1930) pp. 541–588

[6] P. Mahalanobis, "On the generalized distance in statistics" *Proc. Nat. Inst. Sci. India (Calcutta)*, 2 (1936) pp. 49–55

[8] Bhattacharyya, A. On a Measure of Divergence between Two Statistical Populations Defined by Their Probability Distributions. *Bulletin of the Calcutta Mathematical Society*, 35, (1943) 99-109.