Establishing an Independent Data Quality Analysis Framework for UDL Published Datasets

A. J. Gonring, J. Won, D. Chen
The Aerospace Corporation

M. A. Mendiola, J. Y. Cruz, S. Bergonzi, A. L. Chervenak, V. J. Kong
The Aerospace Corporation

ABSTRACT

With the rise in proliferated satellite constellations and commercial space ventures, there is a desire to make tracking data available from more diverse sources than in the past. The Unified Data Library (UDL) is an online data repository for all data pertaining to Space Domain Awareness (SDA). Commercial data providers, including LeoLabs, Inc., are invited to publish their tracking data to the UDL. In this report, a framework for processing, reviewing, and analyzing commercial data pertaining to SDA is discussed. A sample of data from LeoLabs was used in developing the framework, and results from the analyses are presented. First, LeoLabs' radar observations and state vectors are checked for self-consistency through an orbit determination process. Next, reference data from a calibration satellite is used to analyze the states and observations provided by LeoLabs. Lastly, a model to propagate covariances more accurately is developed to enhance the utility of the data available from the UDL.

1. INTRODUCTION

The Unified Data Library (UDL) is an online data repository, envisioned to be a one-stop shop for all data pertaining to Space Domain Awareness (SDA) [1]. The UDL is being developed for the Air Force Research Laboratory and Air Force Space and Missile Systems Center by Bluestaq. The vision is to have data from government, commercial, and foreign data providers be accessible from a single platform and with a common format. While some commercial entities, such as LeoLabs, Inc. and Numerica, Inc., have already begun pushing their data to the UDL, this data is not currently being incorporated into official space catalogs maintained by government entities. Before this data can be used by authoritative sources, its accuracy must be independently verified. In this study, a framework was developed to analyze radar observation data and state vectors published to the UDL for accuracy and utility. A sample of data from LeoLabs was used while developing this framework; however, the capabilities discussed in this paper apply to data from any provider.

First, LeoLabs' radar observations and state vectors are checked for self-consistency. The observations are used in an orbit determination run using TRACE, the Aerospace Corporation's high-accuracy computer program for orbit determination, trajectory analysis, and ephemeris modeling. The resulting orbit is compared to LeoLabs' state vectors to ensure that the state vectors are reflective of the observations used to generate them. The observation residuals are also compared to the published observation uncertainties.

Next, a calibration satellite for which high-accuracy ephemerides are publicly available was selected. LeoLabs state vectors and radar observations are compared against the high-accuracy ephemerides from the calibration satellite to determine baseline accuracy, detect biases in the observations, and compare differences to published uncertainties.

Third, a model to more accurately propagate covariances is developed. Using calibration satellite data, the process noise to be used in the covariance propagation model is adjusted to account for errors in the data and dynamic model. Development of this capability greatly expands the utility of the data acquired from UDL for various SDA activities, such as conjunction analysis (CA).

2. LEOLABS RADAR STATIONS AND DATA PRODUCTS

LeoLabs is a private company that tracks objects in low Earth orbit (LEO) using multiple radar stations across the globe and publishes data products to its own data platform as well as the UDL. At the time the analysis was performed, LeoLabs operated two phased-array ultra high frequency (UHF)-band radar stations at Poker Flat, Alaska (Poker Flat Incoherent Scatter Radar, PFISR) and Midland, Texas (Midland Space Radar, MSR) [2]. They have recently added an S-band radar station in New Zealand (Kiwi Space Radar, KSR), but data from this station was not used in the analysis. The UHF radars track thousands of objects in LEO, including objects as small as 10 cm [2].

From these radar stations, LeoLabs provides a number of data products to the UDL, including location information of their radar stations, range and range-rate observations and uncertainties, and estimated state vectors of tracked objects. These state vectors are generated using observations and orbit determination processes from LeoLabs and contain additional information regarding the implemented dynamic models and corresponding solved-for state parameters. LeoLabs data is available for download on the UDL after requesting and receiving access to their products.

3. TRACE

TRACE is The Aerospace Corporation's trajectory analysis and orbit determination program, and is the primary tool used in this analysis. Originally developed in the 1960s, TRACE has been used regularly to support a variety of high-fidelity analyses, including orbit propagation, ephemeris generation, orbit determination, and covariance analysis [3].

A consistent feature of TRACE across all modes of use is the level of detail and configurability available for the user. Users have the ability to model orbital trajectories with a high-order numerical integrator or a selection of analytic propagators. Use of the numerical integrator allows the user to configure many detailed force models to appropriately characterize the dynamic environment for their scenario. Additionally, TRACE performs orbit determination and covariance estimation. It is able to ingest and simulate many types of measurements and estimate various states and biases. Output files provide ephemerides in the coordinate frame of choice, along with trajectory partial derivatives, observation residuals, and covariance results.

TRACE has been used as the benchmark for independent verification and validation of other commercial astrodynamics tools and orbit propagators. It continues to be used extensively in support of various U.S. government space programs.

4. SELF-CONSISTENCY OF THE DATA

To check the self-consistency of LeoLabs' data, the range and range-rate observations from LeoLabs' radar stations were used in a differential correction (DC) analysis to estimate a state vector for the same satellite that LeoLabs had issued the previous day. The polar orbiting, LEO satellite, SPOT-6, was identified for analysis due to the quantity of observations available from LeoLabs.

TRACE was used to perform the DC process. In order to align dynamic models as closely as possible with those used by LeoLabs in their state estimation process, the following models were used in TRACE:

- 42x42 Earth gravity model 2008 (EGM-08)
- Solar and lunar third-body gravity forces
- Flat plate solar radiation pressure (SRP) model
- 2000 Mass-Spectrometer-Incoherent-Scatter (MSIS) Neutral Thermosphere atmosphere model
- Daily estimates of atmospheric flux indices
- Solid earth tide forces
- Nutation, precession, and polar motion models

Additionally, station range and range-rate biases were estimated before being input into TRACE. LeoLabs presents daily station bias information outside of UDL. This information was taken into account; however, given the time span of the data used, station data from consecutive days were grouped into non-overlapping bins and assigned a shared station bias.

The TRACE DC process fits an orbit to the observations, biases, dynamics, uncertainties, and initial state as provided or described by LeoLabs. It returns an updated state vector at the same epoch as the initial state estimated by LeoLabs. The difference between the LeoLabs' state and the state generated by TRACE for SPOT-6 is shown in Table 1 along with the published uncertainty of the state vector, also provided by LeoLabs. The position and velocity differences and uncertainty are both listed in the radial, in-track, and cross-track (RIC) frame, with the uncertainty listed as a 1-sigma value. The differences are commensurate with the published uncertainties and are likely caused by effects due to reduced but unresolved differences in the physical models, estimated biases, and estimation methods used by LeoLabs versus TRACE.

Table 1: Results from TRACE DC process compared to LeoLabs' state uncertainty

	R [m]	I [m]	C [m]	dR/dt [m/s]	dI/dt [m/s]	dC/dt [m/s]
Difference	-21.4	87.0	30.5	-0.07	0.02	0.16
1-σ Uncertainty	8.5	39.3	43.8	0.05	0.01	0.19

Final observation residuals from the differential correction are shown in Figures 1, 2, and 3. These plots show the differences between the published observations from LeoLabs and the expected observations based on the resultant orbit from the TRACE DC process. With the exception of range-rate residuals from the Poker Flat station, the residuals are distributed as expected based on the published observation uncertainties. Across these figures, blue dots represent residuals associated with MSR and green crosses with PFISR. The shading of the colors represents the binning of passes for estimating station biases.

Fig.1 shows the range residuals from the differential correction process from both stations. These results are in agreement with the 1-sigma measurement uncertainties published by LeoLabs. Each pass had its own unique measurement uncertainty, with the range uncertainty varying between 9-14 meters for MSR and 13-16 meters for PFISR.

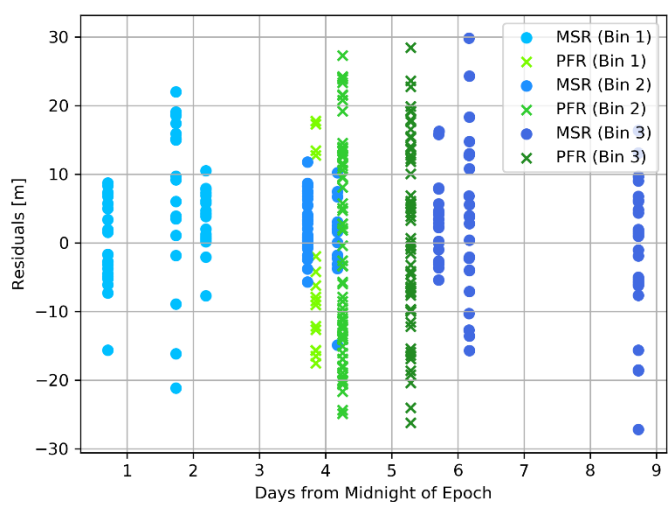

Fig.1: Range residuals from DC analysis

Figs. 2 and 3 show the range-rate residuals, with Fig. 2 showing the residuals from both stations. The residuals from PFISR are much larger than MSR, which was expected based on the uncertainties provided by LeoLabs; however, the PFISR residuals from our analysis still exceeded the uncertainties from LeoLabs for some passes. The posted 1-sigma Doppler measurement uncertainties per pass for PFISR ranged between 2-4 m/s, while for MSR it was

between 0.19-0.29 m/s. Fig. 3 shows the residuals from only MSR. The choice of bins to reduce station biases was made to prioritize resolving range biases over range-rate biases. This is more apparent in Fig. 3 where some MSR biases are still visible.

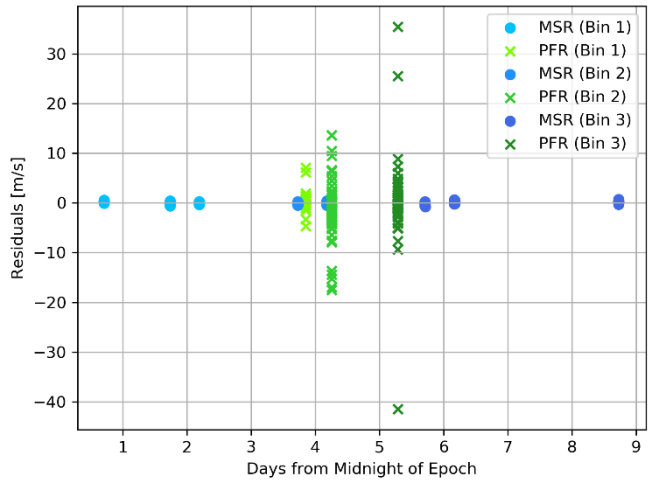

Fig. 2: Range-rate residuals from DC analysis

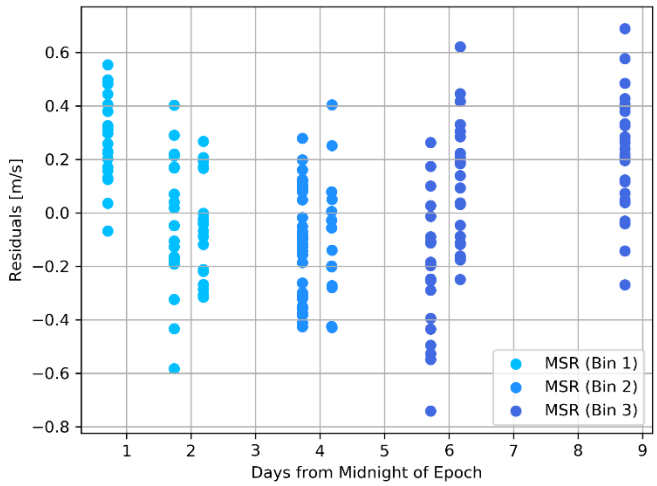

Fig. 3: Range-rate residuals from DC analysis from only MSR

5. CALIBRATION AGAINST REFERENCE EPHEMERIDES

To assess the accuracy of LeoLabs' data, observations and state vectors were calibrated against publicly available precision ephemerides for the European Space Agency's (ESA) Swarm B satellite [4], which is tracked via satellite laser ranging [5]. Swarm B was selected due to the ease of access to mission data and, as with SPOT-6, the quantity of tracking data available from LeoLabs.

Before the accuracy of the LeoLabs data could be assessed, the baseline accuracy of TRACE's orbit propagation for the dynamic environment of the specific satellite had to be determined. This was accomplished by comparing TRACE-propagated ephemerides to the reference ephemerides from ESA. Using the same models as in Section 3, differences with the reference ephemeris were brought down to the order of 10 meters over a 20 hour arc. Fig. 4 shows these differences in the RIC frame. Data available from ESA did not have drag or SRP coefficients for the satellite, so this propagation included constant drag and SRP coefficients that were solved for in an earlier differential correction process, which mimicked the process discussed below.

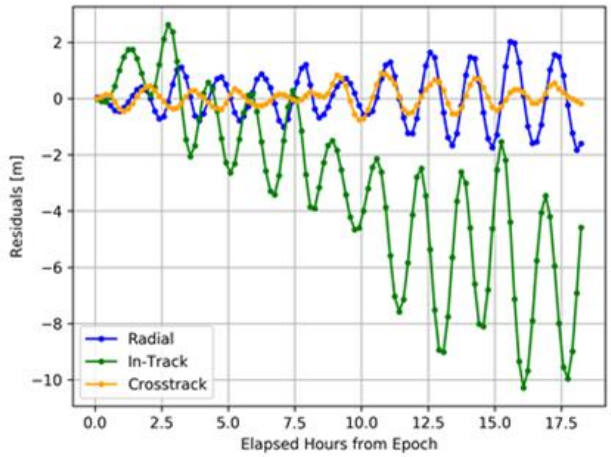

Fig. 4: Differences in position between TRACE-propagated ephemeris and ESA ephemeris

The TRACE DC analysis used a state from the reference ephemeris as the initial condition and a down-sampled set of states from the same ephemeris to represent observations. The goal was to identify an orbit that aligned as closely to the reference ephemeris as possible, while allowing the initial state and drag and SRP coefficients to change as needed. Fig. 5 shows the final differences between the ESA ephemeris and the TRACE orbit solution.

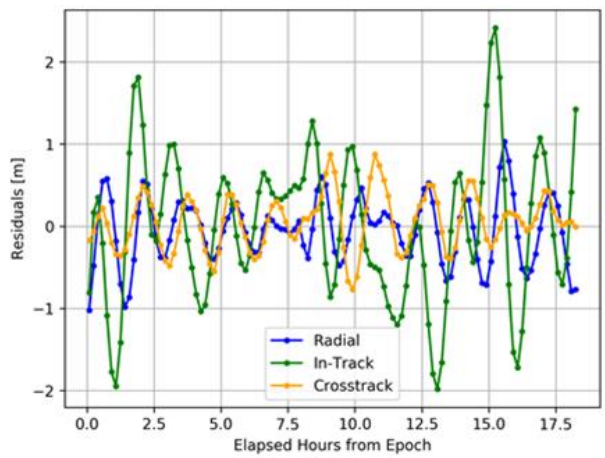

Fig. 5: Differences in position between TRACE DC solution and ESA ephemeris

Table 2 shows the changes in the initial state that were made to determine this orbit during the DC process. These changes are well within the stated accuracy of the reference ephemeris, which was 7 meters (1-sigma).

Table 2: Initial state changes from DC analysis

	R	I	C
Position [m]	1.2	1.5	0.2
Velocity [mm/s]	-1.88	-1.25	0.04

Through the propagation and differential correction analyses, the TRACE results were shown to align well with the reference data. With this agreement demonstrated, TRACE and the reference data were used to assess the accuracy of the sampled LeoLabs data.

LeoLabs' estimated state vectors were first reviewed. The same methods that were used to evaluate the reference data, propagation and differential correction, were used to analyze the state vectors. First, a reference state was

propagated to the epoch of the LeoLabs state vector and compared directly. The reference ephemeris listed state vectors every second, and so the latest vector preceding the selected LeoLabs state was chosen and propagated for less than one second. The difference between this propagated reference vector and the LeoLabs state vector are shown in the first row of Table 3.

Table 3: State differences with reference data compared to LeoLabs uncertainty

	R [m]	I [m]	C [m]	dR/dt [mm/s]	dI/dt [mm/s]	dC/dt [mm/s]
Difference (Propagation)	10.6	9.4	2.6	-8.0	-11.3	-8.8
Difference (DC)	11.8	10.9	2.8	-9.9	-12.6	-8.8
1-σ Uncertainty	85.1	177.8	53.3	194.7	98.4	69.0

Next, a differential correction analysis was used to estimate the state error. State vectors from the ESA reference ephemerides were used as observations, and the LeoLabs state vector was used as the initial state estimate. The drag and SRP coefficients that were solved for when evaluating the reference ephemerides were used as the initial values but were allowed to adjust in the DC process.

The difference between the solved-for initial state from the DC and the initial LeoLabs state is shown in the second row of Table 3. The differences from both methods compare well with one another. LeoLabs also provides a one standard deviation uncertainty on their published state in an earth centered inertial frame. This uncertainty was rotated to the RIC frame to match the previously computed differences and is shown in the last row of Table 3. For this individual state published by LeoLabs on the UDL, it was found that the differences between it and a reference state compared well against its associated uncertainty.

LeoLabs also provides measurement data and its uncertainties to the UDL. These measurements were compared to the resultant orbit of the DC process using the reference ESA states, which generated an orbit to closely align with the calibration data. The LeoLabs measurements and station biases were incorporated into the TRACE analysis to produce the differences between the LeoLabs measurements and the expected measurements from the orbit fitted to the calibration satellite data. These residuals are plotted in Figs. 6 and 7, which show residuals for only one pass from one station, MSR. Fig. 6 shows the residuals for the range observations and Fig. 7 shows the residuals for the Doppler observations.

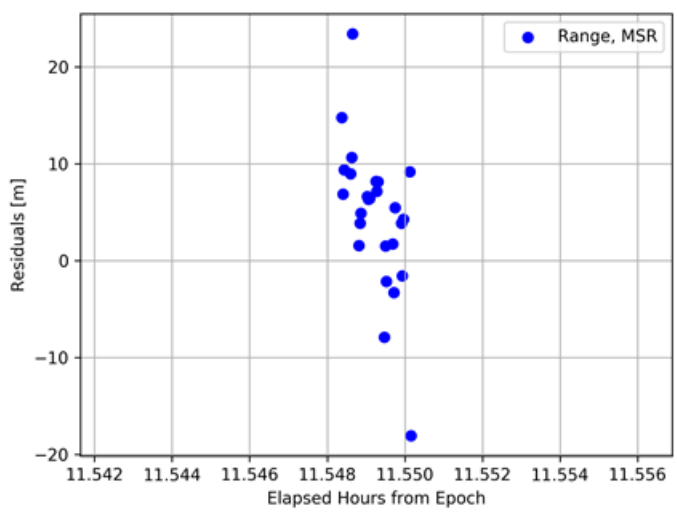

Fig. 6: Range residuals for one MSR pass compared to an orbit fitting the calibration satellite data

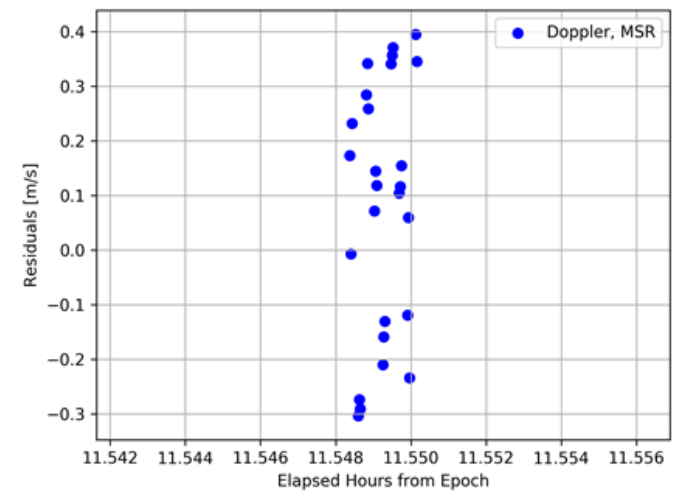

Fig. 7: Doppler residuals for one MSR pass compared to an orbit fitting the calibration satellite data

The posted measurement uncertainties compared well with the residuals shown in Figs. 6 and 7. For the range measurements for the pass of interest, the posted 1-sigma uncertainty was 22 meters, and for the Doppler pass the uncertainty was 0.23 m/s.

Finally, a check on the propagation of the state vector was performed. The LeoLabs state vector was propagated for 20 hours using TRACE and compared to the ESA calibration satellite states. Fig. 8 shows the differences in the RIC position between the calibrated satellite ephemeris and the LeoLabs state propagated in TRACE. Differences in model parameters in the propagation contribute to the observed differences, as well as the initial error in the LeoLabs state vector. For the case in Fig. 8, the SRP and drag coefficients used were the ones provided by LeoLabs in their published state vector.

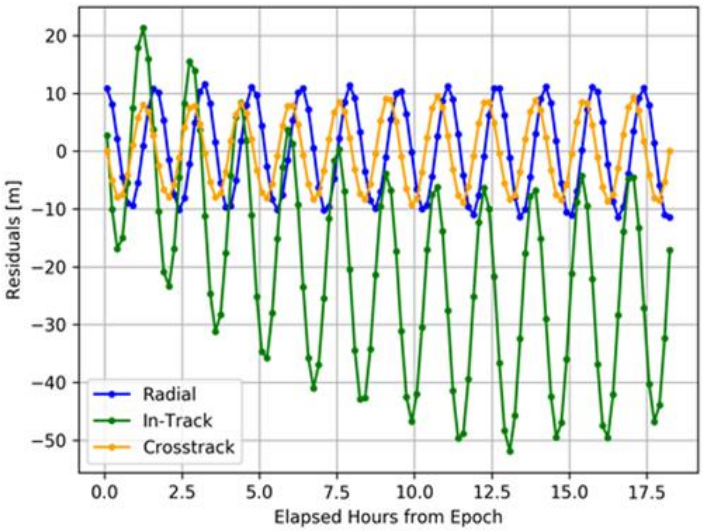

Fig. 8: Difference in position in propagated LeoLabs state with calibration satellite ephemeris.

The various capabilities discussed in this section can be used to assess the quality of the data published to the UDL. This framework includes the abilities to ingest various forms of data posted on UDL, compare UDL data to outside reference data, and analyze posted data for biases and consistency. These tools can also be used to suggest possible improvements or changes to be made by the publisher in how the data is produced and what products should be published.

6. CALIBRATION OF PROPAGATED COVARIANCE AGAINST REFERENCE EPHEMERIDES

An important potential use of LeoLabs' state vector data is in conjunction analysis (CA). This analysis requires realistic propagated covariances of the objects involved, so an assessment and calibration of the covariance propagation model is needed. As shown in Section 5, slight differences were noticed between the propagation model used in TRACE and the reference ephemerides from ESA. Therefore, in anticipation of future CA efforts, the covariance propagation model was calibrated against the reference ephemerides.

The covariance was calibrated by tuning an additive process noise matrix, similar to a Kalman filter. This tuning was performed *after* including the effects of uncertainties in the following physical parameters as consider parameters:

- Solar radiation pressure coefficient
- Drag coefficient
- Earth's central gravity term GM
- Earth's second gravitational harmonic term J2

The process noise matrix was tuned by comparing the analytically propagated covariance against the covariance of the difference between a propagated Monte Carlo spread (80 000 points) and the reference ephemerides from ESA.

The analytically propagated covariance was calculated as follows:

$$P(t) = \phi(t, t-1)P(t-1)\phi^T(t, t-1) + Q(t, t-1) \tag{1}$$

where $P(t)$ is the covariance at time t, $\phi(t, t-1)$ is the state transition matrix from time $t-1$ to time t, and $Q(t, t-1)$ is the process noise matrix from time $t-1$ to time t. The state transition matrix was calculated by integrating the variational equations (partial derivatives of the trajectory equations with respect to the trajectory parameters) together with the satellite trajectory. Both the covariance and state transition matrices included terms for the consider parameters.

The Monte Carlo samples were drawn from the same initial covariance distribution as the analytic propagation. Each sample was propagated using the same dynamic model. At each time step, every sample $\vec{x}_i$ was differenced with the ESA reference ephemeris $\vec{x}_{ref}$, and the covariance $\bar{P}$ of those differences was calculated.

$$\bar{P} = \frac{\sum_i (\vec{x}_i - \vec{x}_{ref})(\vec{x}_i - \vec{x}_{ref})^T}{num\ samples} \tag{2}$$

It should be noted that the covariance was calculated by using the reference ephemeris as the mean, rather than the sample mean. This produces a distribution that is biased relative to the reference ephemeris due to the imperfect dynamic modeling of the propagator. The covariance calculated from this biased distribution therefore accounts for the dynamic modeling errors in the propagated ephemeris. If the dynamic modeling errors are zero-mean Gaussian in nature, then the bias itself has a random walk pattern for a given propagated timeseries. Thus, over many propagation attempts, the bias would be expected to be zero-mean Gaussian distributed. Unfortunately, since the propagation model itself is deterministic, and there is only one "truth" data point per time for a given satellite, it is not possible to generate statistics on the distribution of the bias at a single given timestep. Rather, the bias statistics must be built up over many propagations of different initial states. This proved to be too computationally expensive for the present study, so only a single propagated timeseries was used, but a truly rigorous covariance calibration would perform the calibration process using many propagated timeseries of the same satellite.

The process noise matrix entries which were tuned were the diagonal entries for radial, intrack, and cross-track velocities, the SRP coefficient, and drag coefficient. Since propagation was performed in the ECI frame, the velocity process noise was rotated from RIC to ECI before being added in equation (1). These matrix entries were tuned by using a numerical solver to minimize the root-mean-square error between $P(t)$ and $\bar{P}(t)$. In order to prefer over-bounding the covariance rather than underestimating it, errors in which $P(t)$ was smaller than $\bar{P}(t)$ were weighted more heavily in calculation of the error. The particular solver used in this case was a genetic algorithm, but any black-box numerical algorithm could have been used.

While the intention of covariance calibration is to correct for physical model errors, several other error sources enter into the calibration because the absolute truth orbit is unknown. Baseline noise in the reference ephemerides means that the initial state being propagated has some error in it already. Additionally, since the propagated states are compared back to the reference ephemerides, this baseline noise also affects the apparent errors in the propagated states. The scope of effects that play into the covariance calibration is shown in Fig. 9.

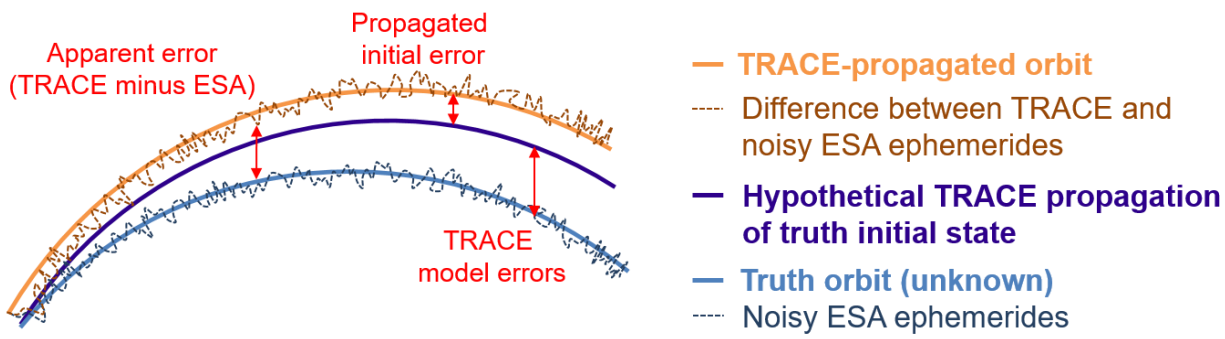

Fig. 9: Different effects captured by the covariance calibration process

Results of the covariance calibration are shown in Figs. 10, 11, and 12 for the radial, in-track, and cross-track position and velocity uncertainty components. The dashed black lines show the covariance calculated from Monte Carlo samples, the solid orange lines show the analytically propagated covariance prior to calibration, and the solid green lines show the calibrated covariance.

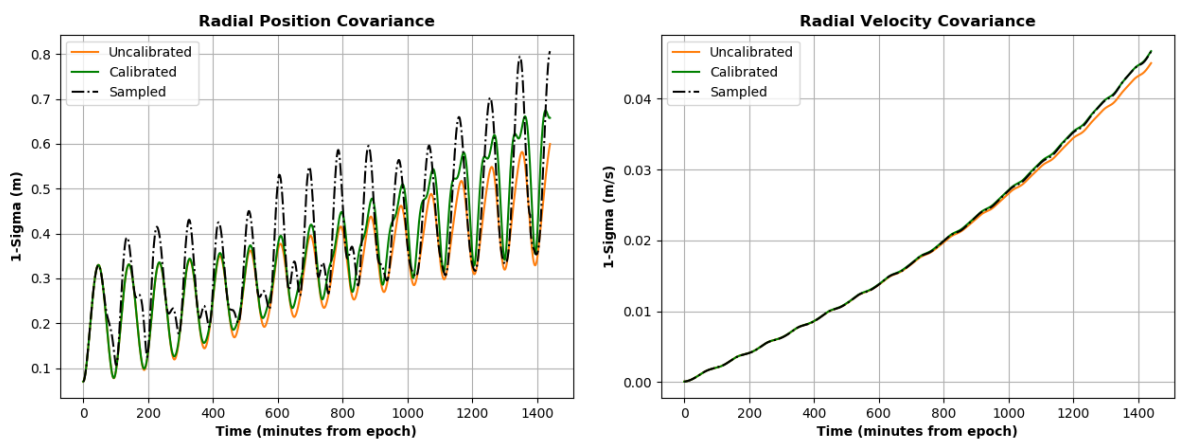

Fig. 10: Radial position and velocity covariance comparison

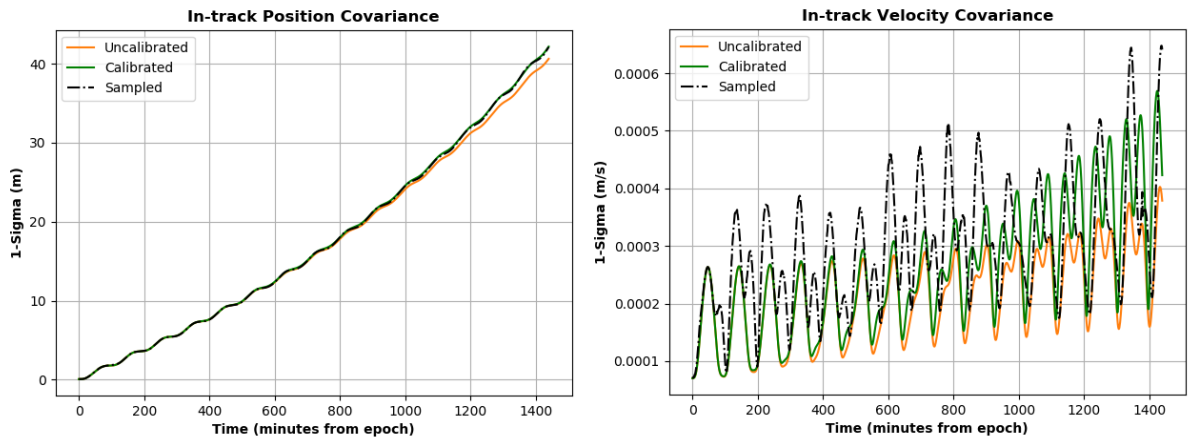

Fig. 11: In-track position and velocity covariance comparison

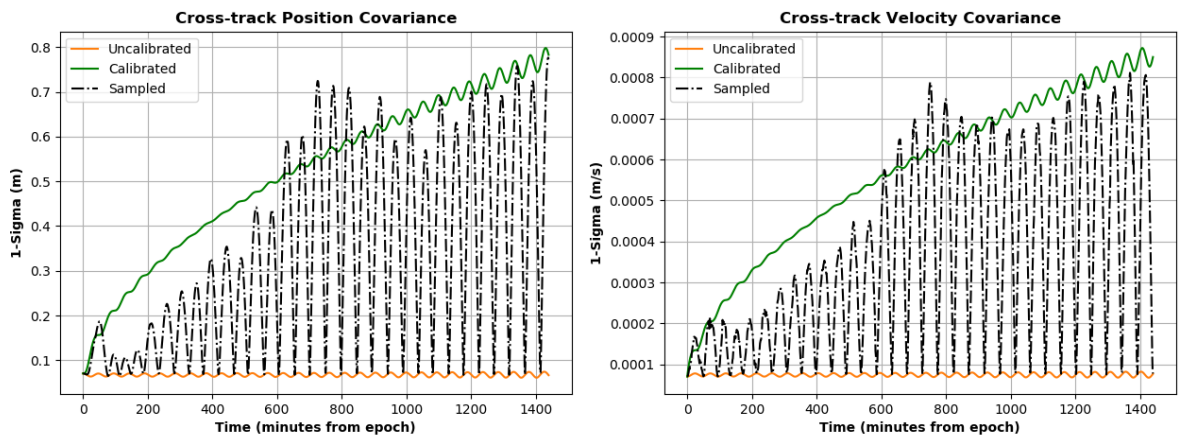

Fig. 12: Cross-track position and velocity covariance comparison

As shown in Figs. 10-12, there is a noticeable difference in the analytically propagated covariance compared to the Monte Carlo sampled covariance, even with the addition of the calibrated process noise matrix. The dominant components of the covariance – in-track position and radial velocity – were able to be calibrated closely to the sampled covariance. While calibration did improve the other components as well, the simple process noise model used for calibration could not match the shape of the sampled covariance curves. Since the cross-track covariance could not be matched, it was calibrated instead to closely bound the sampled covariance.

The calibrated covariance can be used to improve the fidelity of conjunction analysis assessments using satellite states and covariances which are propagated from the state vectors obtained from the UDL in the previous sections. The conjunction analysis using UDL data can be compared to official conjunction data messages (CDMs) to assess the utility of the data for conjunction analysis and collision avoidance purposes. The conjunction analysis itself is left for future work.

7. CONCLUSION

The analyses presented above demonstrated a process by which the quality of state vectors and observation data pushed to the UDL by commercial companies can be assessed. The specific LeoLabs data that was analyzed showed generally good consistency between the stated uncertainties and the actual accuracy and precision. The only outlier was the range-rate observations from the PFISR, which had much larger observation residuals than the stated

uncertainty. The calibrated covariance modeling technique showed improved performance in propagating covariance, adding another tool in support of CA that provides additional utility for UDL data. Altogether, the analyses presented here comprise a framework for assessing and utilizing commercial SDA data published on the UDL. Future work includes development of a pipeline to increase automation of these capabilities and facilitate machine to machine interaction.

8. REFERENCES

[1] "LeoLabs." https://www.leolabs.space/ LeoLabs Inc. (accessed 2020).
[2] "Unified Data Library." https://unifieddatalibrary.com/ (accessed 2020).
[3] R. H. Prislin, D.C. Walker, and R. J. Mercer, "TRACE66 Trajectory Analysis and Orbit Determination Program, Volume 1: General Program Objectives Description, and Summary," The Aerospace Corporation, August 1971.
[4] "Swarm Data Access." https://swarm-diss.eo.esa.int/ European Space Agency. (accessed 2020).
[5] "About ILRS" https://ilrs.gsfc.nasa.gov/ International Laser Ranging Service. (accessed 2020).