US-EUSST Data Exchange for Improved Orbital Safety

Felix Hoots
The Aerospace Corporation
Valentin Baral; Florian Delmas
Centre National d'Etudes Spatiales
Alejandro Cano; Santiago Martínez
GMV
Cristina Pérez
CDTI
Matthew Hejduk; Patrick Ramsey; Kerstyn Auman
The Aerospace Corporation

ABSTRACT

Data sharing and exchange among different Space Situational Awareness (SSA) data collectors/providers is frequently discussed as a mechanism to improve the accuracy and precision of orbital safety products. To collect empirical data on the improvements achieved through sharing of measurement data, the US Office of Space Commerce and the European Union Space Surveillance and Tracking (EUSST) Support Framework have performed a joint experiment to collect and exchange satellite metric observations. The experiment covers a 60-day collection period on 12 sample satellites from Low Earth Orbits (LEO), Medium Earth Orbits (MEO), Highly Eccentric Earth Orbits (HEO) and Geosynchronous Orbits (GEO). The SST Partnership EU Member States are represented through their national designated entities: Austria (FFG), Czech Republic (MDCR), Denmark (RDAF), Finland (FMI), France (CNES), Germany (German Space Agency at DLR), Greece (NOA), Italy (ASI), Latvia (IZM), the Netherlands (EZK), Poland (POLSA), Portugal (PT MoD), Romania (ROSA), Spain (AEE) and Sweden (SNSA). The US and EUSST have performed orbit determination with their own observations, the other entity's observations, and with both observation data sets combined. The product results are then compared with each other, showing differences attributable to the three different data groups and to the two entities' different approaches to orbit determination.

1. INTRODUCTION AND PROJECT HISTORY

Since the beginning of the space age in 1957, both the US and Russia have maintained their own catalogs of earth-orbiting, human-made objects. As a consequence of a series of US/Russian Space Surveillance Workshops beginning in the mid-1990's, the two nations exchanged satellite catalogs on a quarterly basis for several years. This was the first instance of large-scale space surveillance information sharing between nations. Today many nations have significant assets in space and have developed their own satellite tracking networks and catalog maintenance processes, so there are many independently-maintained satellite catalogues and derivative SSA products.

For decades the possibility of inadvertent collisions of orbiting objects has been routinely examined and alerts to operators have been issued. With the recent significant increase in large commercial constellations, the US and other nations are developing systems dedicated to Space Traffic Coordination: both the accurate identification of potentially high-risk conjunctions and the codification and implementation of a rule-base for prescribing and allocating mitigation actions to reduce such risks. The continued growth in space congestion increases the urgency of a cooperative posture among the national space surveillance capabilities to reduce the risk of avoidable collisions among satellites.

In 2018, at a meeting between the European Union Space Surveillance Tracking (EUSST) Consortium and the United States Department of State, Department of Defense, and the National Aeronautics and Space Administration (NASA) representatives, the goal of cooperative collection and sharing of space surveillance data between the EUSST and the US was first articulated. Follow-up periodic technical discussions reached the conclusion that the most effective way of achieving improvement in flight safety assessments is by sharing data at the observation level, allowing effective

and straightforward data fusion. In contrast, the alternative of sharing ephemeris data rather than observation data is substantially more complicated and, while theoretically possible, technically unproven. This goal came to fruition in February 2021 in the form of a joint agreement between the US Department of Commerce and the EUSST to perform an observation sharing experiment to understand its benefits.

The European Union established the Space Surveillance and Tracking (SST) Support Framework in 2014 with the Decision 541/2014/EU of the European Parliament and the Council [1]. This Decision foresaw the establishment of an SST Consortium, which evolved since its creation into a Consortium of seven EU Member States. The SST Consortium EU Member States are represented through their national designated entities: France (CNES), Germany (German Space Agency at DLR), Italy (ASI), Poland (POLSA), Portugal (PT Ministry of Defense), Romania (ROSA) and Spain (CDTI). In addition, eight new Member States are now joining the EU SST Partnership: Austria (FFG), Czech Republic (MDCR), Denmark (RDAF), Finland (FMI), Greece (NOA), Latvia (IZM), the Netherlands (EZK) and Sweden (SNSA). Within EUSST, the two entities involved in the analysis conducted herein are CNES through its Centre Opérationnel de Surveillance de l'Espace (COSE) and the Spanish Space Surveillance and Tracking Operations Centre (S3TOC).

By early 2022, funding and procedures were in place to perform the experiment with the following general steps:

- Identify a set of satellites for the experiment and secure approvals for data exchange for these satellites. A couple of these satellites have external truth orbits, while many others do not.
- Perform joint tracking of these satellites over a fixed period of interest; the period chosen here is 1 JUN - 31 JUL 2022
- Exchange observational data and supporting information (sensor locations, calibration, &c.)
- Construct orbits from the exchanged observational data, with each entity using its intrinsic orbit determination (OD) capabilities. Three combinations of OD approaches are to be accomplished by each entity: US data only, EUSST data only, and the two datasets combined.
- Examine the set of results and draw conclusions
- Form a preliminary assessment of the benefits of data sharing

The satellites chosen for the experiment were a combination of active spacecraft and inert spacecraft (dead payloads, rocket bodies, and debris), spanning the four standard orbital regimes (LEO, HEO, MEO, and GEO). For a more durable final experiment, it may make sense to choose a much larger set of objects; but this particular set was appropriate to standing up the basic machinery of the experiment, working through data exchange difficulties, and producing insights that would allow the enlightened selection of a larger experimental set. The particular satellites selected are given in Table 1 below.

Table 1. Experiment Satellites

Orbit Regime	Satellite ID	Perigee / Apogee (km)	OD Fit-Span (Days)	Reference Orbit?
LEO	1982-096B	1,496 x 1,504	10	N
	1984-108B	371 x 386	5	N
	1986-094B	1,499 x 1,505	10	N
	2016-011A	807 x 808	10	Y
MEO/HEO	1999-036A	518 x 39,848	12	N
	1999-025EQ	19,910 x 20,459	12	N
	2013-026B	2,664 x 63,621	14	N
	2017-079C	23,214 x 23,233	15	Y
	2020-018B	19,388 x 19,710	12	N
GEO	1967-066F	33,264 x 33,665	15	N
	1974-033A	36,196 x 36,321	15	N
	1999-018A	36,284 x 36,341	15	N

Similarly, particular tracking sensors were selected by both entities to provide tracking data for the experiment. For the US, policy decisions restricted the participating sensors to ground-based sensors only. For the EUSST, certain sensors were unable to participate at the time, while others could only share suboptimal data due to similar policies. It therefore can be said that neither entity furnished all of the observational data that could possibly be mustered for the experiment satellites, but that response was significant enough to enable meaningful analysis. The sensors from both entities that contributed to the experiment are given in Table 2 below.

Table 2. Experiment Tracking Sensors (US sensors at left; EUSST sensors at right)

List of US sensors involved			
ID	Name	Location	Type
211	SOCORRO CAM1	Socorro, NM	Optical
212	SOCORRO CAM2	Socorro, NM	Optical
213	SOCORRO CAM3	Socorro, NM	Optical
231	MAUI CAM1	Maui, Hawaii	Optical
232	MAUI CAM2	Maui, Hawaii	Optical
233	MAUI CAM3	Maui, Hawaii	Optical
241	DIEGO GARCIA CAM1	Diego Garcia	Optical
242	DIEGO GARCIA CAM2	Diego Garcia	Optical
243	DIEGO GARCIA CAM3	Diego Garcia	Optical
310	Space Fence KW	Kwajalein	Radar
334	ALTAIR	Kwajalein	Radar
335	TRADEX	Kwajalein	Radar
338	HEH	Western Australia	Radar
344	FYLINGDALES A	RAF Fylingdales	Radar
345	FYLINGDALES B	RAF Fylingdales	Radar
346	FYLINGDALES C	RAF Fylingdales	Radar
354	ASCENSION (12.15)	Ascension Island, UK	Radar
369	MILLSTONE L-Band	Westford, MA	Radar
382	CLEAR (SW)	Clear AFS, AK	Radar
383	CLEAR (N)	Clear AFS, AK	Radar
386	CAPE COD (NE)	Cape Cod AFS, MA	Radar
387	CAPE COD (SE)	Cape Cod AFS, MA	Radar
388	BEALE (S)	Beale AFB, CA	Radar
389	BEALE (NW)	Beale AFB, CA	Radar
393	COBRA DANE	Eareckson AFS, Alaska	Radar
394	THULE (SE)	Thule AB Greenland	Radar
395	THULE (N)	Thule AB Greenland	Radar
396	PARCS	Cavalier AFS, ND	Radar
398	EGLIN DS	Eglin AFB, FL	Radar
399	EGLIN NE	Eglin AFB, FL	Radar
979	CUSTODY RAVEN	Maui, Hawaii	Optical

List of EUSST sensors involved			
ID	Name	Location	Type
411	ES_BOOTES-5	Baja California, Mexico	Optical
412	ES_TFRM	Lleida, Spain	Optical
413	ES_TJO	Lleida, Spain	Optical
414	ES_IAC-80	Tenerife, Spain	Optical
416	FR_TAROT-CALERN	Caussols, France	Optical
417	FR_TAROT-CHILI	Coquimbo, Chile	Optical
418	FR_TAROT-REUNION	Réunion	Optical
419	FR_GRAVES	Revest-du-Bion, France	Radar
426	IT_MFDR-LR	South Sardinia, Italy	Radar
427	RO_TEEMO-T03	Bucharest, Romania	Optical
428	IT_SPADE	Matera, Italy	Optical
429	PL_OBORNIKI	Oborniki, Poland	Optical
466	IT_MFDR-MR	Nuoro, Italy	Radar
468	IT_PDM-MITE	Rome, Italy	Optical
472	PL_BEATA	New Mexico	Optical
473	PL_COAST	Tenerife, Spain	Optical
475	PL_MOONBASE	Khomas Region, Namibia	Optical
476	ES_S3TSR	Seville, Spain	Radar
477	PL_PST2	Sonoita, AZ	Optical
478	PL_POLONIA	Antofagasta, Chile	Optical
479	PL_SLR2	Sutherland, South Africa	Optical
480	PL_SLR3	Coonabarabran, New South Wales	Optical
481	FR_SATAM-R1	Retjons, France	Radar
482	FR_SATAM-R2	Sommepy-Tahure, France	Radar
483	FR_SATAM-R3	Ventiseri, France	Radar
486	DE_TIRA	Wachtberg, Germany	Radar
488	PL_RANTIGA	Reggio Emilia, Italy	Optical
490	RO_BERTHELOT	Tismana, Romania	Optical
491	RO_NEEMO-T05	Bucharest, Romania	Optical
943	ES_CENTU-1	Ciudad Real, Spain	Optical
944	ES_TRACKER-1	Ciudad Real, Spain	Optical

2. MEASUREMENT DATA FUSION CONSIDERATIONS

Data fusion at the measurement level, in which sensor metric observations from different entities' sensors are combined into a single OD fit called the fit-span, seems straightforward; and indeed it is in comparison to proposed methods for direct fusion of SSA data products, such as ephemerides and CDMs. However, there are a number of potential issues that arise when exchanging measurements generated by someone else's sensor system. Many of these issues have been known by experienced SSA practitioners for years, but it is helpful to catalogue them here, as they were encountered in this experiment and in many cases not resolved without deliberate and non-trivial effort.

1. **Sensor Observation Type**. Different sensor phenomenologies can produce different observables; for example, radar sensors typically produce with each measurement at the least a range-to-target and two angular measurements, whereas optical sensors simply produce two angular measurements. Standard measurement types such as these are accommodated by all major OD systems. However, the EUSST GRAVES radar is a bistatic radar system that furnishes range-rate as well as angular measurements. Because the US SSN does not use bistatic radars to produce range rate measurements (a US angles-only interferometer was decommissioned in 2013), its operational system code is not presently configured to accept bistatic range-rate measurements. In the present experiment, the US processing used only the angles measurements from this radar, which produced a notable divergence in results for satellites for which GRAVES furnished substantial amounts of information.
2. **Sensor Observation Corrections**. There are a number of corrections to sensor measurement data that are applied between the collection of the raw measurement and the use of the measurement data in OD; and the particular point at which these corrections are applied is largely a question of preference or local custom. These corrections themselves are important and vary by sensor phenomenology: for optical sensors, aberration corrections (annual and diurnal) to account for the displacement in stellar relative position due to the motion of the earth; for radar sensors, ionospheric and tropospheric refraction and wavefront delay that can affect both range and angular observables; and for all sensors, measurement biases in all observables due

to systematic errors. It is important to understand which of these corrections have already been applied to the observations, and which remain to be applied, when the data are exchanged between maintenance entities. During the conduct of the present experiment, despite *a priori* attempts at US-EUSST measurement correction reconciliation, iterative efforts were required during the actual experiment conduct to reach a point at which there was full mutual understanding on measurement data corrections and the consistent use of similarly-corrected data on both entities. The final consensus reached was to share observations with all relevant corrections and biases applied, a task best done by the entity operating the sensor, having knowledge on its inner workings.

3. **Data Formats**. US SSN systems typically use a "B3" observation format, which was built in conformity to punch-card constraints and has been in use since the 1960's. The EUSST sensors, which are of more recent vintage, export information in the Consultative Committee for Space Data Systems (CCSDS) Tracking Data Message (TDM) format. The TDM format forces the specification of many of the aspects of the measurement in order to reduce ambiguity and makes provision for substantial metadata. Each entity was required to become conversant with the other's format.
4. **Sensor Data Rates**. Radar sensors in the US SSN typically use tracking filters; these filters collect radar hit data and, through a sequential estimator, output a smoothed observation at either a regular time interval or when a particular cumulative SNR threshold has been reached. The GRAVES radar does not do this in the same way, essentially outputting data at the rate of the individual radar hits. This difference in data density between radars, which can be several orders of magnitude, causes difficulties in OD: if a single sensor submits vastly more data than the others in the correction, its dataset heavily dominates the fit. Some sort of compensation mechanism for this disparity must be introduced; in this experiment, an approach called Track Weighting was employed by the US cataloguing system. Because correlation between subsequent measurements is very high, each additional observation taken during a tracking session provides rather little information gain; so it is usually better to think of measurements in terms of single measurement collection sessions conducted during a single pass of the satellite over the ground-based sensor (a clear concept in LEO but admittedly more problematic in GEO). Track weighting adds additional weighting factors to the observation set in order that the data contribution from each tracking session, regardless of the number of observations taken during that session, may hold the same weight in the OD; this gives a GRAVES track the same influence as one by a US SSN sensor and smooths out the disparate initial weighting brought about by the difference in measurement density. Such a Track Weighting approach is not available in the EUSST system.
5. **Sensor Locations and Calibration Information**. In order to use a sensor's measurements in an OD, it is necessary to know the location of the sensor and the general quality of that sensor's observations. The latter is typically measured as the variance of residuals from comparing a sensor's measurement data on a calibration satellite to an as-flown, precision reconstructed ephemeris for that calibration satellite. Subtleties exist in both specifying sensor locations (*e.g.*, astronomical versus geodetic latitude and longitude, positioning of the sensor local vertical for the topocentric coordinate system) and sensor residual variances (*e.g.*, outlier exclusion, sub-binning by correlating attribute such as time of day or elevation angle); and these issues need to be worked through before exchanged measurement data can be used properly. Indeed, certain abiding sensor location and calibration issues required multiple iterations to resolve satisfactorily.

The important takeaway here is that the attendant technical issues to simply understanding another entity's measurements are considerable, and notable engineering effort and schedule needs to be allocated to this step of the data fusion process. The use of modern message formats for the data exchange, such as the CCSDS TDM, by both entities would simplify the situation substantially, as such a format allows many of the above details to be specified; but there are also coordination items that are not resolved by message format alone that will need to be confronted in any such exchange situation.

3. GENERAL METHODOLOGY DESCRIPTION

The orbit determination approach used by both the EUSST and US systems employs a Gauss least squares process to minimize the sum of the squares of the residual differences between the measurements and the within-fit-span ephemeris prediction. This compatibility facilitates a much more direct way to agree on data processing and results comparison, which was governed by the following agreements:

1. Each entity will provide observational data removing all known biases from their data (aberration, refraction, &c.).
2. For each observing site, each entity will provide nominal measurement weights applicable during the experiment period.
3. Each entity will use the same time span for fitting the observations, modified as appropriate for each orbit regime. Solution for parameters such as ballistic or radiation pressure coefficient is left to the practice of each entity.
4. For each satellite, the fitting window will be performed with state vector epoch on the whole day at the conclusion of the window. The window fit will slide in one day increments until the data is exhausted.
5. This process will be repeated using US only data, EUSST only data, and combined US and EUSST data producing three separate analysis sets for each of the test satellites.

The key metric that will determine possible benefits of data sharing is the prediction accuracy using combined data, compared to the data coming from either the EUSST or the US systems alone. This accuracy is estimated at the 1-, 2-, and 3-day prediction intervals. We quantify the accuracy by computing the vector magnitude difference between the predicted position and the corresponding position on the reference orbit. To avoid time-phased periodic effects, we computed the root mean squares (RMS) of the vector magnitude differences over a single orbital revolution centered on the desired comparison epoch (*i.e.,* from half a revolution before to half a revolution after the evaluation point). This quantity (RMS over an orbital period of the magnitude of the vector differences) is referenced as "prediction error" hereafter for convenience.

For those satellites where we have an external reference orbit, the comparison is straightforward. For each orbit determination, the predicted state vector for 1, 2, and 3 days from epoch is computed and compared to the reference orbit at those times. For satellites without a reference orbit, however, there is no unassailable way to proceed to evaluate prediction error. If, however, it can be shown for those examples with reference orbits that the solution using both datasets is notably superior to either solution that uses only the EUSST or US data, then the solution arising from the combined dataset can be used as a *de facto* reference orbit point, with the difference between it and the EUSST- or US-only solutions serving as a proxy for the expected improvement wrought by combining data. To be sure, such results cannot be fully durable, and they are probably more meaningful for the larger prediction intervals (for which the prediction error is expected to be substantially larger than the inherent error in the reference point); but this approach should serve as a general index of the improvement that data sharing should yield.

Because the predicted ephemerides in this case share some of the measurement data with the constructed reference orbit that will be used to evaluate them, data correlation possibilities arise and should be minimized. Ideally one would use non-overlapping fit spans. But this would severely limit the number of samples available in our 60-day study period. As a compromise, to significantly reduce correlation possibilities the following procedure has been implemented. Suppose the fit-span for a particular satellite is 9 days and the first fit covers the closed time interval [1,9] days, with epoch placed at the end of day 9. To evaluate the vector arising from this fit, comparisons between the reference orbit and the propagated vector need to take place at the end of days 10, 11, and 12; the synthetic reference for comparison is the combined data fit shifted by one-half the fit-span. To be explicit, the day 10 prediction would be compared to the midpoint of the combined fit span [6,14], the day 11 prediction would be compared to the midpoint of the combined fit span [7,15], and the day 12 prediction would be compared to the midpoint of the combined fit span [8,16]. Fig. 1 below gives a schematic representation of this process. The green bar represents a nominal orbit determination using data in the closed interval [1,9]. The blue bars represent daily sliding windows of reference orbits using both the US and EUSST data. The Greek letter identifies the days used for the 1-, 2-, and 3-day prediction comparisons using the state vector with epoch at the end of day 9.

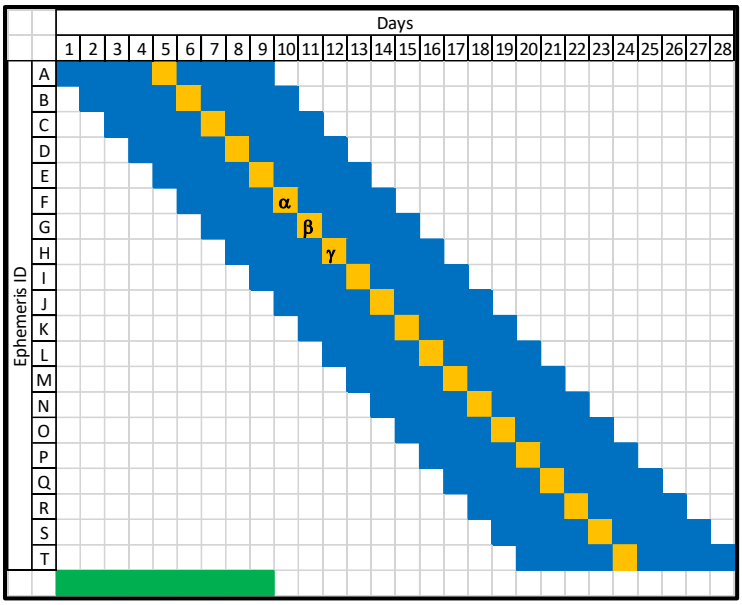

Fig. 1. Illustration of moving-window reference orbit construction and use

In an experiment such as this, in which both datasets and orbit determination capabilities are obtained from different entities, a natural question arises whether it is data disparity or OD fidelity that bears the responsibility for OD prediction error differences. The multiple graphical presentations presented and explained in detail in the following sections, while aiming principally to communicate the (salutary) effects of data combination, also are intended to help to answer this secondary question. One begins by examining data throughput to see whether notable differences are observed between the EUSST and the US data holdings. The left side of Fig. 2 below, which is an excerpt from Fig. 4 discussed in detail later, is, for the MEO orbit regime, a cumulative distribution function (CDF) plot of the percent of the tracks in each fit-span that come from one of the providers. Such a plot places the variable of interest, here the percent of data in each fit-span, on the x-axis and the cumulative percentage on the y-axis; it is essentially a continuous version of a histogram (because of the nature of the calculation, each of the two lines is an "odd" function, meaning that they are related by a 180-degree rotation through the graph center point; small differences between the two lines in the tails when making this rotation are due to using the Kaplan-Meier CDF generating function and are not significant). The (20, 80) point is both a good practice-point for reading the graph type and useful for the present analysis: 80% of the fit-spans for the satellites in this orbit regime have less than 20% of their data coming from the US, meaning that 80% or more of the data for those fit-spans come from the EUSST. At the same time, about 5% of the cases have 50% or more of their data coming from the US. So differences in performance due to differences in input tracking are thus expected.

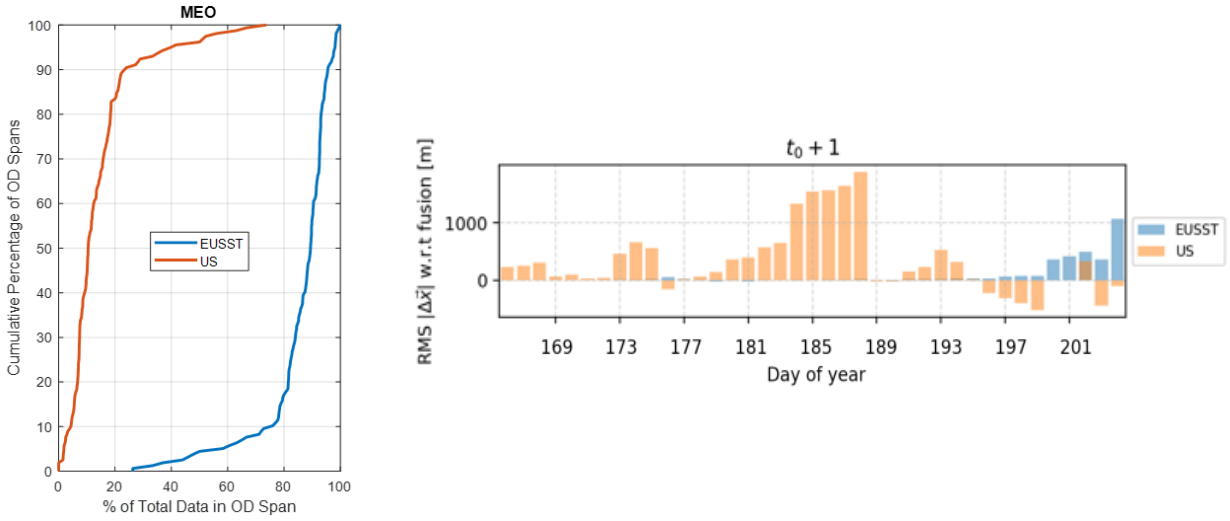

Fig. 2. Graph excerpts

The right side of Fig. 2, which is an excerpt from Fig. 6 discussed later, presents prediction error for a Galileo satellite (a MEO spacecraft) as a function of time (each bar is one day's results from 2022 spanning from day 166 to day 204). For this satellite, the OD results using EUSST data only, US data only, and both datasets combined were each compared to an external precision reference orbit; and the EUSST and US results are shown as a relative change over the results for the combined dataset, to wit: if the prediction errors for the updates using one supplier's data directly equal those for the combined dataset, then there is null relative change. The very high relative error observed for the US results for this satellite, at their worst in the middle section of this graph, represent a period in which there is extremely little US data and therefore very poor fits when using only that set of data. A similar effect is present at the very rightward portion of the graph, at which point the EUSST begins to experience a data gap for this satellite. When a similar graph is constructed using the US rather than the EUSST OD engine, nearly identical results are obtained; this builds confidence that the prediction error differences observed are due to differences in tracking throughput, in which each entity's tracking tends to cover for the other's gaps or dropouts.

4. TRACKING DENSITY AND ORBIT COVERAGE IMPROVEMENTS

It is generally the case that a more copious set of tracking data will produce a better orbit determination, but it is important to keep in mind the particular ways in which this general principle can miscarry. First, an infusion of additional "bad" data will not improve the OD and instead will usually erode accuracy, even if the fresh but marginal additional measurements are accompanied by realistic calibration data. Indeed, particular cases of this very phenomenon were observed during the experiment. Second, the addition of further quality measurements beyond a certain level of tracking is, by itself, unlikely to improve the situation markedly. As a general rule, the size of the least squares OD covariance varies inversely with the square root of the number of measurements in the correction; so beyond a certain data density the marginal improvement becomes small. Third, the OD improvement from additional tracking is governed most strongly by the improvement in dispersion of tracking about the orbit arc, to wit: additional tracking on a part of the orbit that is already heavily sampled will yield very little palpable improvement, but additional tracking on those parts that have little to no sampling usually results in notably improved OD.

Because the US space-based sensors were not involved in the study, and because the EUSST lacks a large number of radar sensors, interesting difference in orbit coverage from different sensor phenomenologies were introduced into this experiment, not directly by design but perhaps to the benefit of a more broadly applicable study. The performance statistics below bring out how data sharing addresses such differences; these statistics are presented using cumulative distribution function (CDF) plots, which is a convenient way of showing the relative prevalence of particular density and dispersion situations as a fraction of the total tracking results sets encountered. Looking at overall amounts of tracking data or data dispersion, even by individual satellite, is not necessarily particularly revealing; rather, it is the degree to which holdings are improved by combining datasets over each dataset standing alone, thus counselling the use of percentage or other comparative measurements. The degree to which the combined dataset improves prediction

accuracy is discussed in subsequent sections; the goal of the present section is to show improvements in data density and dispersion, which themselves are required for timeliness of update and improved orbit modeling. Finally, rather than examine overall levels for the entire experiment or even individual satellites, it is best to look at each orbit determination interval of each satellite to see what difference is made in the context of actually enabling OD fits. The y-axis in the subsequent plots indicates the "cumulative percentage" of all the fit-spans for all the satellites in the orbital regime summarized.

Fig. 3 below shows, for the case of combined measurements, the percentage of data in each fit-span from the two providers. In the GEO case, for almost 70% of the cases the US data holdings are very much in the minority—less than 20% of the data in the fit-span; and the situation is even more one-sided for the MEO orbit regime, in which 90% of the cases have a US contribution of less than 20% of the data in the fit-span. In LEO, the contributions appear more even; but it should be remembered that there is a large data-rate mismatch between the GRAVES radar and the US radars.

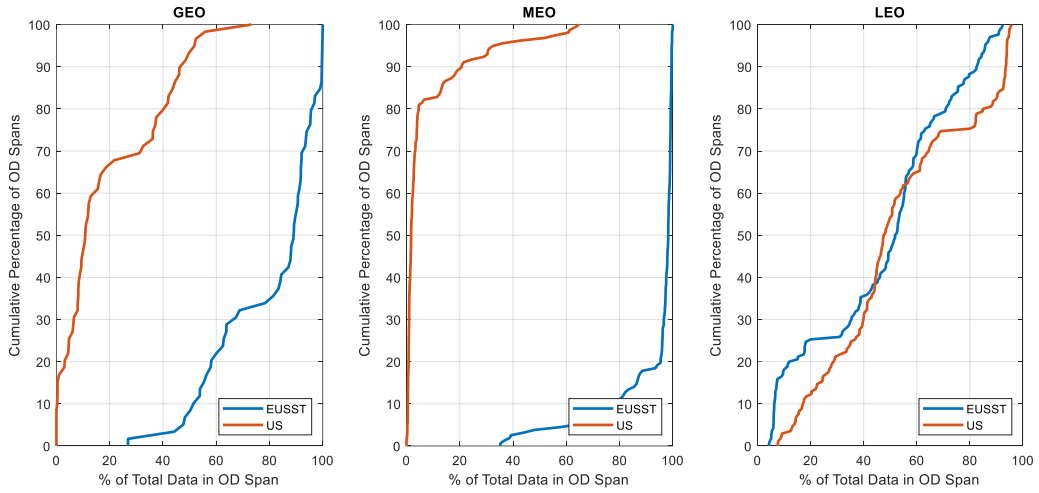

Fig. 3: Percent (by provider) of combined measurement data in each OD span

One way to neutralize this data density difference is to report "tracks," rather than individual measurements, contributed by both entities to each fit-span. A track is the collection of measurement data that a sensor takes during a single observing session on a particular satellite. For ground-based sensors tracking LEO objects, "passes" of a satellite over the sensor have well-defined durations (namely the periods of geometric visibility); so it is straightforward to determine whether a sensor took tracking data during a particular pass and identify that data burst as a single track of data. Such an assignment is more difficult in deep-space and is particularly problematic in GEO, since a satellite can dwell over a ground sensor indefinitely. In such cases, typically a 20-minute separation between measurements is used to define track boundaries; such a construct is admittedly somewhat artificial, but some stratagem must be employed.

Fig. 4 below shows the data density results for tracks rather than measurements themselves. The MEO case largely retains the same morphology, promoting the expectation for the combined dataset of substantial information gain over the US data taken alone; but the GEO and LEO track submission results are substantially different from the measurement results. While the EUSST concept of operations results in more individual measurements for GEO, if one applies the 20-minute criterion the amount of non-correlated information gain is much more equally shared; and for the LEO results, a reasonably equal sharing of overall measurements becomes a notable US advantage in supplying actual tracks: in 60% of the fit-spans, the EUSST provides less than 20% of the tracks.

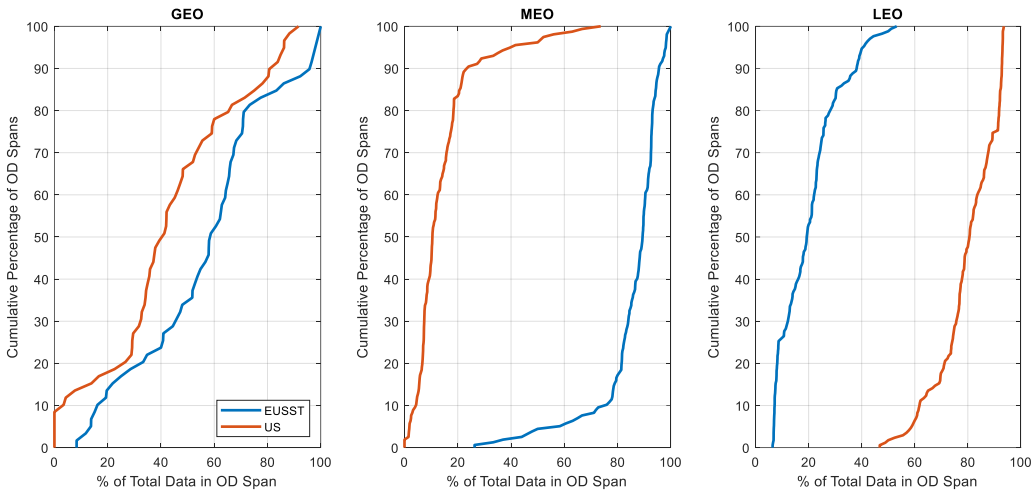

Fig. 4: Percent (by provider) of combined tracking data in each OD span

As remarked earlier, data density as a metric has limits; for it is the dispersion of the data about the orbit, rather than a simple count of measurements or tracks in a fit-span, that is more strongly correlated with a durable OD fit. There are many ways one might attempt to evaluate orbit dispersion; the approach that seemed to fit the experiment situation well was to divide the orbit up into a set of "bins" based on argument of latitude and determine what percentage of these bins were populated under the US, EUSST, and combined situations. For the present analysis, eight bins of 45-degree size in argument of latitude were selected (i.e., 0-45, 45-90, 90-135, &c.); and for each fit-span for each satellite, the number of such bins populated was turned into a total percentage of the orbit covered: for example, if five bins were populated, 225 degrees of the orbit was considered covered. Such an approach can be seen to suffer from a Nyquist sampling issue, namely that a measurement with an argument of latitude of 44 degrees and a measurement with an argument of latitude of 46 degrees would generate the conclusion that 90 degrees of the orbit was sampled, when in fact rather little of the orbit was actually sampled. This concern is lessened somewhat by reporting only comparative data: one is interested in the number of additional bins populated when the combined dataset is used over either entity taken alone.

Fig. 5 below provides these results in CDF form. Here, rather than working with 360 degrees of orbit coverage, the framework instead focuses on the percent of the orbit covered; since eight bins are being used, if a fit span populates all eight bins, then it provides 100% orbit coverage; if it populates seven bins, then it provides 7/8ths coverage, or 87.5%, &c. The reporting is therefore the number of percentage point improvement in orbit coverage of the pooled dataset over the individual dataset. For example, if a fit-span had 50% coverage with EUSST only and the combined dataset had 62.5% coverage, then the percentage point improvement in orbit coverage was 12.5 percentage points; and that is the datum that would be represented in the CDF plots.

In looking at the plots themselves, one sees the coverage improvements by orbit regime. In GEO, the improvements are quite modest; and this is not surprising given that geosynchronous objects hover over a particular point on the earth's surface and therefore are not well suited to coverage improvements from other ground-based sensors. In MEO, the situation is more sanguine; the US data benefit handsomely through combination with the EUSST data, with 60% of the cases having a greater than 20 percentage point improvement in orbit coverage; for the EUSST, the situation is more muted, but almost 15% of the cases have a 12.5-percentage point or greater improvement. In LEO, one sees that the US's large radar network is such that the EUSST radars almost never provide any improvement in orbit coverage. At the same time, because fit-spans in LEO are long enough to last many tens of revolutions, the natural earth motion ends up generating reasonable orbit coverage even for a small number of radars in a confined geographical location; the coverage improvements over the EUSST LEO dataset are thus rather small.

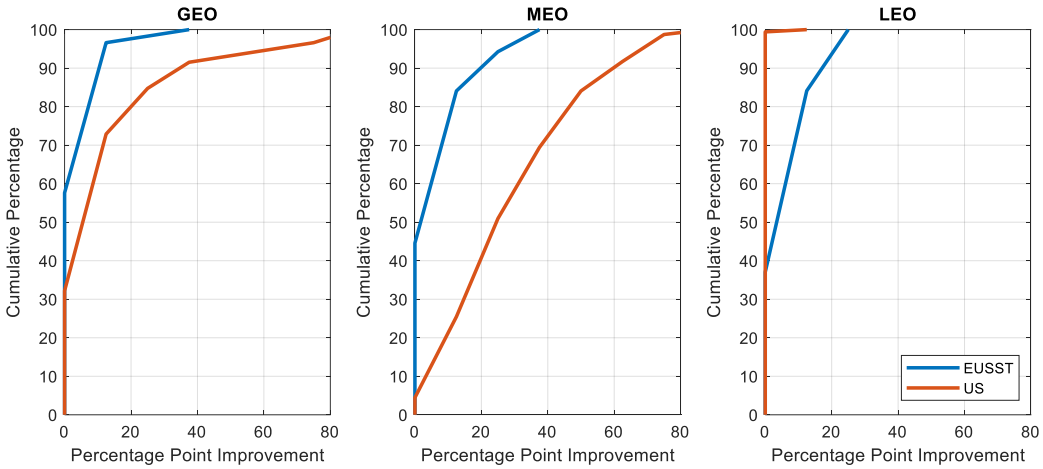

Fig. 5. Orbit coverage improvement for combined dataset over individual EUSST or US datasets

5. PREDICTION ERROR FOR SATELLITES WITH REFERENCE ORBITS

This section presents the analysis of orbit prediction accuracy using the different datasets for the chosen reference satellites, in this case, Galileo-21 and Sentinel-3A. These are operational satellites for which precise orbit determination products are available during the period of analysis, achieving accuracies to the centimeter level (and thus a negligible level of error for the comparisons in question). The main goal of this section is to use these precise ephemerides as ground truth to validate the orbit determination and orbit comparison methodology proposed in this work, as well as to assess the accuracy performance that can be achieved for the different datasets as compared to the real orbit of the objects.

The general methodology for orbit determination and prediction was described in Section 3. In the case that an independently developed reference orbit is available, we do not have to create a synthetic reference to evaluate prediction accuracy. Instead predictions using the state vector resulting from the orbit determination are simply predicted 1-, 2-, and 3-days from the epoch and compared directly to the reference position. It is still important to avoid time-phased periodic differences by computing the residuals RMS over a single orbital revolution centered on the desired comparison epoch.

In order to show the accuracy improvement achieved with data fusion, Fig. 6 and 7 below describe the relative accuracy improvement as compared to the data fusion for both reference satellites, as a function of the day of the year, for the three different propagation epochs of analysis. The improvement in accuracy is computed as follows:

$$\left|\overrightarrow{\Delta x}\right|_{RMS}^{wrt\ fusion} = \left|\overrightarrow{\Delta x}\right|_{RMS}^{dataset} - \left|\overrightarrow{\Delta x}\right|_{RMS}^{fusion}$$

Therefore, the improvement in accuracy is measured as the difference between the accuracy achieved with a single dataset versus the combined dataset, in relative terms. A positive value of the metric indicates that the data-fusion RMS metric is smaller than the single dataset RMS, or, in other words, that the data-fusion orbit predictions are more accurate.

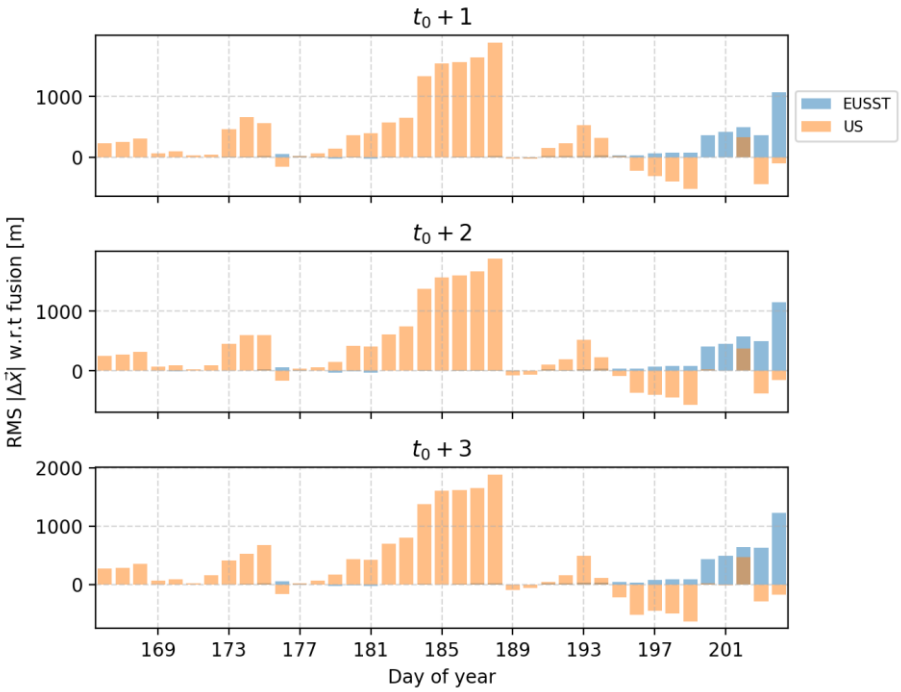

Fig. 6: Relative accuracy for the individual datasets as compared to data fusion for Galileo-21.

At first glance, Fig. 6 shows a generally positive metric of accuracy improvement when the combined dataset is used. This improvement reaches up to 2000 meters as compared to the US dataset in the middle of the analysis period, and up to 1000 meters as compared to the EUSST dataset at the end of the period. Both of these losses in precision are due to periods of data scarcity, leading to significant degradation of the orbit estimation and prediction accuracy. The US dataset has few measurements of Galileo-21 between days 177 and 188 and thus shows decreased accuracy as compared to the combined dataset, which benefits from the EUSST data. On the contrary, EUSST tracking of Galileo-21 stops around day 193, with US tracking filling the gap for the rest of the analysis period. This is seen as an increasing trend in the EUSST dataset error after some days without updates. Thus, generalized benefits of data fusion are shown for this MEO object for both datasets, filling the coverage gaps and improving the accuracy. Moreover, the combination of data not only provides a larger number of measurements and tracking coverage robustness, but also allows a reduction of the update intervals, significantly improving the accuracy.

Fig. 7, which describes the relative accuracy metrics for Sentinel-3A, shows a different scenario. EUSST dataset coverage and measurements distribution issues lead to an accuracy degradation between 100 and 200 meters as compared to the data fusion. In fact, the combined dataset solution is driven mostly by the US dataset contribution, seeing days where the US dataset alone provides even slightly better accuracy than the combined dataset. There are two main reasons for this. On the one hand, there is a clear network capability difference between the US radar network and the EUSST one, which leads the former one to find a small benefit on many occasions. On the other hand, as discussed previously in this paper, the GRAVES sensor is the one providing the highest pool of the EUSST data, outputting observations with a high rate and thus extremely time correlated. This could have been mitigated by the US Track Weighting method, but as already mentioned the US cataloguing system was unable to process bi-static range rate measurement. Moreover, GRAVES' performance was lessened by the fact that the data that could be shared within the frame of this experiment came with suboptimal ionospheric corrections pre-applied. This pulls the combined orbit trajectory towards it, deteriorating the accuracy. This is yet another example of the required effort for a cataloguing system when combining the data of multiple providers. When more data of high quality arrives, providing better geometrical coverage and reducing the weight of other sensors, unseen issues of specific sensors may be pointed out with this insight, where specific treatments such as measurements under sampling or de-weighting are to be applied to make the data combination effective.

Even though the US dataset does not benefit from the data fusion in this LEO scenario, the impact of data combination has a mild relevance on their solution, but significantly improves the EUSST dataset accuracy. This points out that

different cooperation procedures could be proposed for future collaboration, detecting the cases where a certain network is sufficiently accurate on its own and only apply a 1-sided data sharing when appropriate.

Therefore, several benefits of data fusion have been shown with this analysis of orbit prediction accuracy with respect to precise ephemerides. Gaps in the data are covered, improving the estimation and prediction of the orbits when network visibility or availability is reduced. A higher frequency of observations avoids the need for longer propagation arcs. A better diversity of the measurements in terms of geometry and observability allows improved orbit determination accuracy. And, as is shown in Fig. 6 and 7, effectively avoiding very inaccurate estimation at times of data gaps leads to an improvement in the robustness of the SSA system, whose overall benefit exceeds any minor accuracy degradation that may occur if any of the datasets presents undetected biases or inaccurate measurements.

Finally, it is worth remembering that this analysis is performed against precise orbits, known to represent the ground truth. However, in the general case of most SSA activities, the majority of objects are non-collaborative, and a precise ephemeris is not available. Nonetheless, it has been shown in this section that the accuracy and robustness of orbit determination and propagation is expected to be improved when fusing the data. This led to propose the combined dataset orbits as reference to analyze the performance of the individual datasets for the general case of satellites without reference orbit, as was explained previously in Section 3.

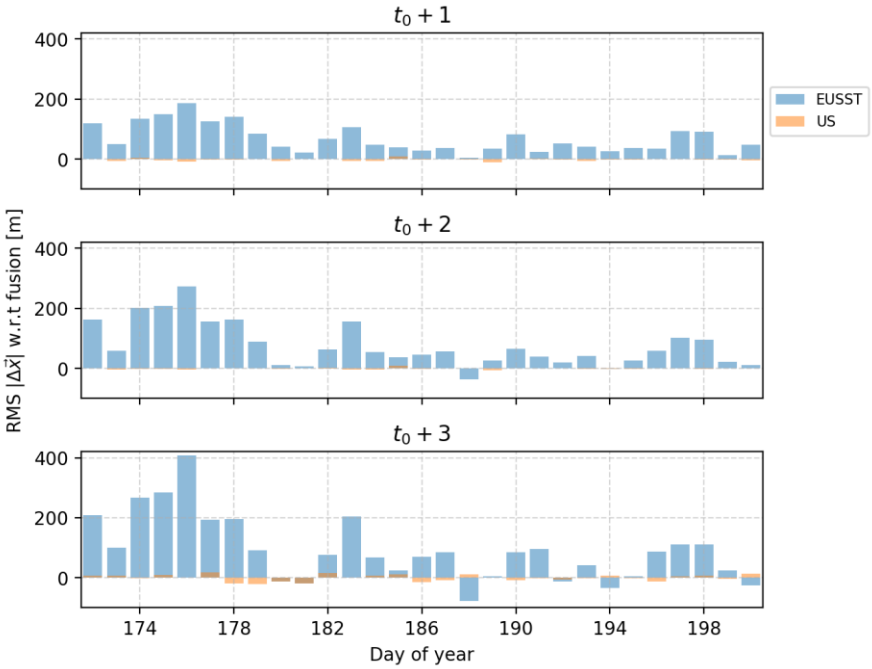

Fig. 7: Relative accuracy for the individual datasets as compared to data fusion for Sentinel-3A

6. PREDICTION ERROR FOR SATELLITES WITHOUT REFERENCE ORBITS

As previously mentioned, the study group chose satellites to cover the three main orbital regimes (GEO, MEO and LEO) with diverse altitudes, eccentricities and inclinations, in an effort to be representative of the global space population and thus of its cataloguing process. Consequently, objects that lack external precision reference orbits were included. This section focuses on the prediction error for these satellites, representative of the general RSO population. The particular process of using the combined dataset solution as a reference orbit has already been explained in Section 3 and it has been shown as appropriate after the analysis of Section 5. One further characteristic of the results presented in this section is the normalization of the prediction error outputs in order to facilitate their public release; they have thus been normalized to a reference accuracy and are unitless, a stratagem that still allows useful relative comparison, which is in fact the particular thrust of this section's analysis.

Two original test objects were, after further analysis, discarded for the following reasons:

- Object 1984-108B, having a mean altitude around 380 km, was originally chosen to represent low-altitude LEO. Unfortunately, the propagation times used for comparison were not suited to that particular regime, as it was found that after even one day of propagation the prediction errors due to atmospheric density uncertainties predominated the entire analysis and swamped any improvements due to more favorable tracking density. A specialized analysis would be required to examine this specific regime with adapted propagation periods, and this was beyond the scope of the present effort.
- Object 2020-018B did not have sufficient observations from either entity during the measurement campaign. Its cataloguing process was thus unable to produce enough prediction points using the US dataset, which made comparison often impossible and the results not meaningful.

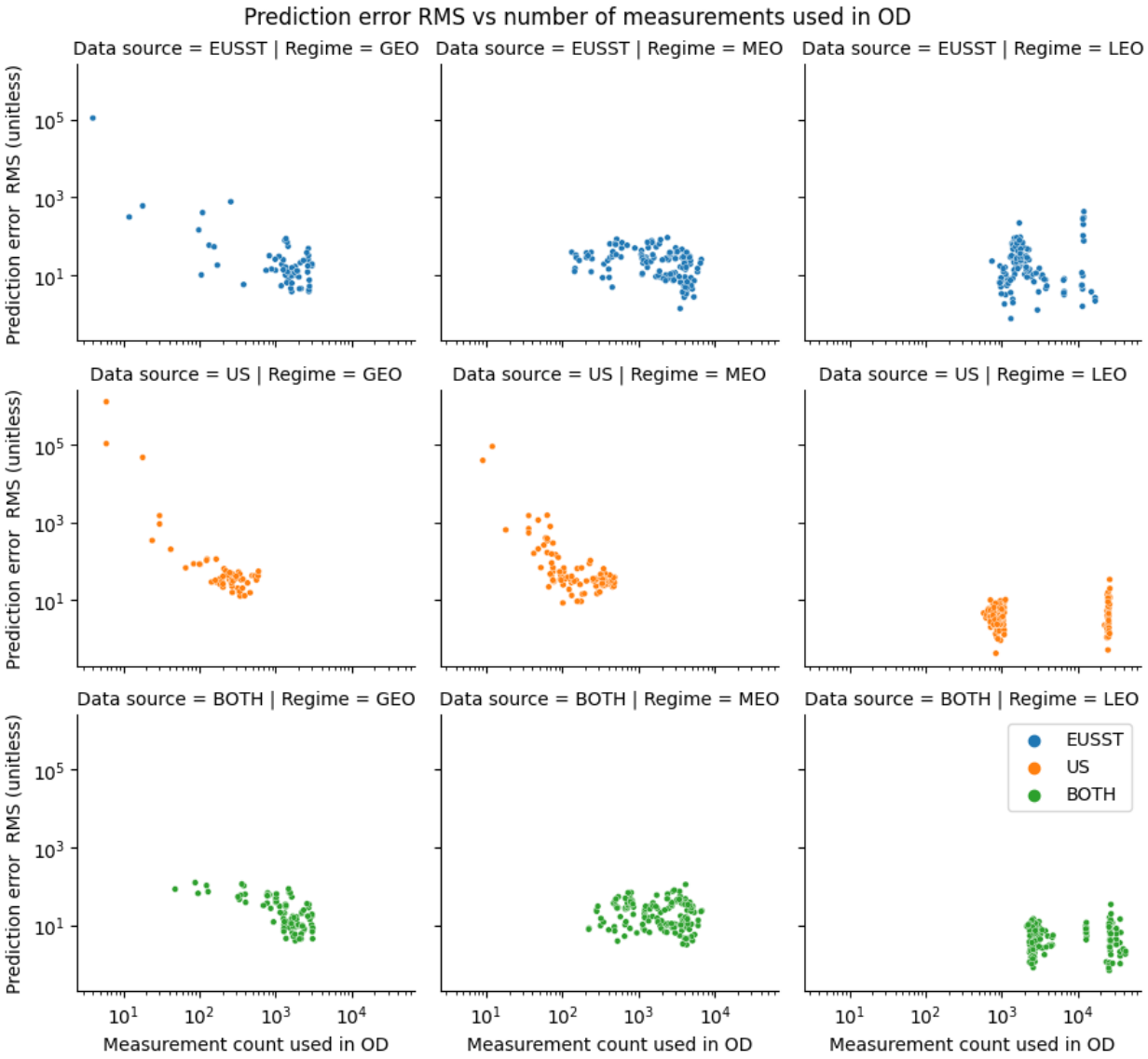

Fig. 8: Representation of prediction error vs. number of observations used in OD, both by regime and data provider

With the remaining objects, a first approach was to look at the prediction error as a function of the number of observations used to produce the corresponding orbit determination. Fig. 8 was produced by grouping all prediction errors by orbital regime and dataset used and plotting them against the number of measurements. Note that the graphs are using a log-log scale to accommodate the broad dynamic range of the data. Several inferences suggest themselves in comparing the data distribution from both providers in Fig. 8:

- For GEO objects, both data providers exhibit some predictions that encompass too few measurements to maintain consistency with respect to the reference orbit. Interestingly, the point at which this consistency disappears is different for each provider, confirming that both networks are not equivalent in the way that they provide information. The geographical dispersion of sensors, their respective performances, longitude revisit rate, and measurement frequency are all probably at play.
- For MEO objects, all objects seemed to be provisioned with sufficient observations in the EUSST dataset such that no clear point of inconsistency in the prediction error occurs. The number of observations in the US dataset, however, is an order of magnitude smaller; and a divergence in consistency does occur, roughly at the 100 measurement per fit-span point.
- For LEO, the number of measurements is not that appreciably different across both providers. However, the US dataset clearly performs better than the EUSST one, even with less data; this is a clear consequence of the network disparities discussed in the tracking density analysis.

In examining the results of the combined dataset now, it is already clear that some tracking synergy is taking place:

- The performance of the GEO combined dataset is better across the board. Clearly, the increased amount of measurement data improves the tracking distribution to a point of much better consistency. From this graph alone, it can be seen directly that it performs better in terms of prediction error than either of the provider's individual solutions.
- The MEO combined dataset performs better than the US one, which suffers from a lack of measurement for certain prediction points. Its overall performance relative to the EUSST dataset cannot be assessed in a conclusive manner from this graph alone.
- Inversely, the LEO combined dataset addresses the EUSST data density problem, with better prediction errors across the board; but it does not show clear signs of improvement with respect to the US dataset.

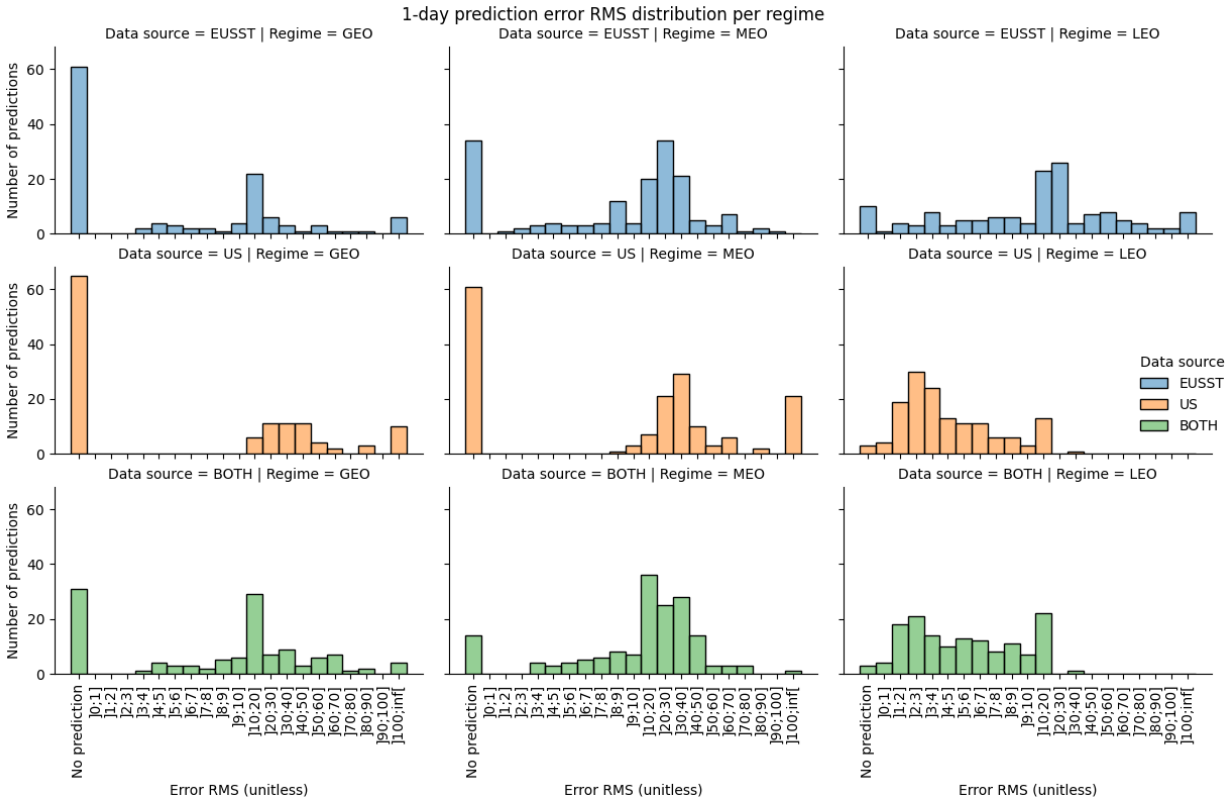

Fig. 9: Prediction error distribution across orbital regimes for each data provider

Fig. 9 presents the prediction error distribution, considering the days in the study period in which no prediction could be issued for each dataset. Results are shown only for a 1-day propagation period, as these comparative results are only marginally different for the 2-day and 3-day periods. The graphs again are categorized by data provider and orbital regime. The first x-axis bin accounts for days in which no prediction could be made; and the following prediction error bins are then sized logarithmically, using the same unitless error as in Fig. 8. Finally, the last x-axis bin gathers all the points with a prediction error above the unitless value of 100, at which point the prediction was judged unsuitable to be used for an orbital safety product.

These error distributions allow for additional assessment of the benefits (or lack thereof) gained from data fusion:

- In GEO, in addition to the global prediction error improvement, it is observed that both providers have around 60 instances in which the lack of data does not allow the production of a new orbit. This figure is roughly halved for the combined dataset; so the fused dataset has the combined benefit of better prediction accuracy as well as providing more consistent temporal coverage across the study period.
- In MEO, the fused dataset has a slightly better error distribution than the EUSST solution, but not of such a degree to advance this as an obvious improvement. Of course, this conclusion must retain as its context that the reference orbit is a synthetic one rather than a precise ephemeris; and therefore smallish differences in prediction error evaluated against a synthetic truth ephemeris remain inconclusive. Nonetheless, the temporal coverage is again much better, as the days without a prediction decrease from 34 to 14. When compared to the US dataset, it is clearly superior in both error distribution and temporal coverage.
- For LEO, however, it appears that no benefit is gained from the data fusion, relative to the US dataset. Some predictions appear worse, while others are slightly improved; and the error distribution is even slightly worse overall, with a more pronounced peak in the [10;20] bin. As can be seen from Fig. 8 and also from Fig. 3 and 4, very little extra information is provided by the two participating EUSST radars compared to the coverage already provided by the US network alone. As for MEO, it is not possible to confirm which of the two solutions (combined or US data alone) is objectively better. A larger statistical sample and a reference trajectory that is less correlated to the predicted one would probably help to resolve this question. As expected, the gain in prediction error is significant when compared to the EUSST dataset, yet the temporal coverage is only marginally better.

As one would expect, the benefits of data fusion are maximized when data scarcity comes into play. In some situations, like the GEO cases that were examined in this study, the combination is better than either of its parts; and the act of sharing data is a mutually beneficial process. In others, it is more of a one-sided approach, where one provider is filling in the data gaps of the other. In some cases, no improvement could be measured convincingly; but likewise there was no evidence of a statistically significant deterioration of the solution either.

It is important to note that all results in this section rely on the predictions made by the EUSST cataloguing system. A similar approach using predictions stemming from the US cataloguing system revealed that results were largely in line with the EUSST ones, save for the LEO case. As discussed previously, this was expected since one limitation that we encountered was that the US cataloguing system was not suited to readily process bi-static radar measurements. For the sake of simplicity and consistency, only the EUSST estimates were used to produce Fig. 8 and 9.

As a technical summary, Table 3 below gives a synthetic view of how the fused dataset performs with respect to the individual providers:

Table 3: Synthetic summary of the fused dataset's performance vis-à-vis those of the individual providers

		EUSST		US	
		Temporal coverage	Prediction error	Temporal coverage	Prediction error
Combined	GEO	Significant improvement	Significant improvement	Significant improvement	Significant improvement
	MEO	Significant improvement	Inconclusive	Significant improvement	Significant improvement
	LEO	Slight improvement	Significant improvement	No improvement	Inconclusive

7. CONCLUSIONS

The analysis of orbit accuracy using precise ephemerides as reference showed that that the precision and robustness of orbit determination and propagation is improved when combining the data, solving data gaps, coverage, network availability issues and reducing the update intervals. For instance, it has been seen how the US can significantly benefit from EUSST data in the case of Galileo-21 to cover the data gaps, whereas EUSST can improve the quality of their products for Sentinel-3A when adding data from the US radars network, which provides wider coverage and geometry. The combined dataset solution, once proven generally more accurate and robust than the individual datasets, has been used as reference to estimate the orbits accuracy of the rest of satellites under analysis.

The benefits of data fusion are maximized when data scarcity comes into play. In some situations the combination is better than either of its parts, and the act of sharing data is a mutually beneficial process. In others it is more of a one-sided approach, where one provider is filling in the data gaps of the other. Overall there are significant benefits to data sharing at the observation level both in temporal coverage and improved prediction accuracy. Also while some technical coordination is needed, this solution has the advantage of being readily applicable by any operational entities that have some level of proficiency in orbital cataloguing.

8. FUTURE WORK

The next step in assessing the benefits of data sharing might be to examine how the accuracy and realism of the covariance using combined data compares to that based on the covariances for the separate contributors. Assuming favorable results, a sharing process experiment could be applied to selected real world predictions of close approach events. Additionally, given the recent interest in the industry in the direct fusion of data products, a salutary future task would be for ephemerides to be developed first using just the US or EUSST data (representing the ephemeris that each would generate using merely their own data) and then to attempt direct fusion of these ephemerides into a single product (perhaps following the approach of [2]), which would then be compared to an OD that used the entire observational dataset. Such an experiment would represent the operational scenario in which ephemeris or Conjunction Data Message data fusion would actually take place, so it therefore would serve as a telling evaluation of direct data product fusion techniques.

9. ACKNOWLEDGEMENTS

The US authors gratefully acknowledge the sponsorship of this research by the Office of Space Commerce, US Dept of Commerce. In particular Mr. Richard DalBello and Mr. Scott Leonard realized the potential for this cooperative effort and were instrumental in bringing it about.

The European authors would also like to acknowledge Pascal Faucher for his involvement in the early discussions that allowed this project to come to fruition. The EUSST activities have received funding from the European Union programs, notably from the Horizon 2020 research and innovation program under grant agreements No 952852, No 785257, No 760459, No 713630 and No 713762.

10. REFERENCES

[1] Decision No 541/2014/EU of the European Parliament and of the Council of 16 April 2014 : Establishing a Framework for Space Surveillance and Tracking Support (L0158/227), 2014-05-27 https://eur-lex.europa.eu/legalcontent/en/ALL/?uri=CELEX%3A32014D0541

[2] Carpenter, J.R. "Estimate Fusion for Conjunction Assessment." *Journal of Guidance, Control, and Dynamics*, Vol. 43, No. 9, September 2020, pp. 1723-1726.