Space Catalogue Accuracy Modeling Simplifications

Matthew D. Hejduk[1]
SRA International, Woodway, TX, 76712

The most complex and forbidding aspect of Space Surveillance Network performance simulations is the calculation of state vector accuracy, as it requires the direct simulation and noising of both sensor observations and the orbit determination process. A procedure is here described for constructing expected state vector accuracy curves as a function of sensor tracking type and density, thus obviating the need for the explicit simulation of sensor observations and orbit determination. Building on the results of earlier studies, twenty-five million operational precision state vectors were grouped by orbit type and sensor tracking mix, accuracy values for each vector were determined for time periods of interest after epoch, and curve-fits of vector accuracy versus tracking density were executed and analyzed. The results identify a preferred "accuracy decay" functional form, develop a fitting methodology, and establish the general level of fidelity of the overall procedure, which is not of a level to recommend it for the evaluation of individual satellites but yet quite serviceable for establishing the aggregate accuracy of a catalogue of objects.

I. Introduction

All modeling and simulation efforts encounter a tension between fidelity and ease of use: increasing the explicit simulation improves fidelity but increases complexity and run-time; using a more parameterized approach gains simplicity and accessibility but reduces the predictive value of the simulation. The situation is certainly no different with efforts to simulate the operation of the Space Surveillance Network (SSN), the collection of sensors operated by the US military to detect and track artificial earth satellites. The longstanding Space Surveillance Network Analysis Model, operated by Air Force Space Command (AFSPC), opted for fidelity over simplicity: it executes full simulation of sensor observations (complete with sensor-specific noise models) and explicit orbit determination and updates. It collects the wages of high-fidelity in being certified as an operational model (the match to real-world operations is greater than 90%), but at the considerable complications of required parallel processing, huge datasets, and the instabilities and vicissitudes of the orbit determination process.

When the Space and Missile Center (SMC) commissioned the development of a simulation tool for the purposes of force structure planning and investment decision support, a broader range of the fidelity-simplicity trade-space was available. This type of decision aid required only a 70-80% match to real-world operations, both because of the tool's purpose and because of the imprecision with which the performance of proposed future systems is usually known. As execution speed and model flexibility are important features of planning tools, there was an impetus to simplify the operation of this model as much as possible. One area that appeared to bear investigation was whether a parameterized methodology for estimating state vector accuracy could adequately replace the explicit simulation of sensor observations and orbit updates.

II. Proof-of-Concept: "SSNO Study"

The parameterization of state vector accuracy determinations was initially seen as possible because of a similar, although more limited, study effort performed by the Omitron Corporation in 2002[2] and 2003[3]. The Space Surveillance Network Optimization (SSNO) study was commissioned to determine what level of satellite tracking would be required from the assembled SSN in order to obtain desired levels of vector accuracy; this was accomplished by dividing the space catalogue into appropriate orbit regimes and for each regime establishing a functional relationship between vector accuracy and tracking density. The manner in which each of these was accomplished (i.e., the division of the space catalogue and the form and success of the functional relationship) provided very helpful points of departure for the present investigation and are thus described in some detail here.

---

[1] Principal Scientist, SRA International, 10006 Willow Bend Drive, Woodway, TX 76712. Member AIAA.

A. Satellite Taxonomy

Several taxonomic schemes have been assembled over the years to divide the orbital population into groups that represented similar ease or difficulty *vis-à-vis* orbit maintenance; typically, these have focused on the apogee and perigee heights of the orbit; the "orbit class" scheme in use during much of the 1990's, for example, established 65 bins based on apogee and perigee height. This methodology appropriately captures the orbital maintenance differences due to atmospheric density, but it neglects the similarly important drag-related differences due to satellite frontal area properties. Since the actual physical properties of the satellite are rarely known *a priori*, a mechanism is needed to account for the total retarding force of the atmospheric drag. Omitron proposed using as a discriminator the instantaneous energy dissipation rate (EDR) of a particular satellite-orbit pair, given a formal definition of

$$EDR(t) = -\vec{A}_D \bullet \vec{V} \,, \qquad (1)$$

in which A_D is the inertial drag acceleration vector and V is the inertial velocity vector. This quantity (in watts/kg) represents the amount of energy being removed from the orbit due to atmospheric drag (or occasionally being brought to the orbit via the solar wind; objects with high surface-area-to-mass ratios can often gain significant orbital energy this way). A single numerical value for EDR for a given vector update is an averaged value calculated over the orbit determination interval. The SSNO study divided the orbital population into eleven EDR bins, and these divisions have proven a very useful categorical approach. Table 1 shows the precise definitions of the EDR bins and how the space catalogue (May 2007) divided itself among them. Because of the relatively small populations of EDR bins 2-10, it is often convenient to combine them into three supergroups: bins 2-4, 5-7, and 8-10. The definitions and catalogue composition of these three supergroups are given at the bottom of the table.

Table 1. Energy Dissipation Rate (EDR) bins.

EDR Bin	Lower W/kg Value	Upper W/kg Value	% of Catalogue (May 2007)
0	0	0	15.5
1	0+	0.0006	72.7
2	0.0006	0.0010	2.6
3	0.0010	0.0015	1.5
4	0.0015	0.0020	0.8
5	0.0020	0.0030	1.2
6	0.0030	0.0060	2.0
7	0.0060	0.0090	1.0
8	0.0090	0.0150	0.8
9	0.0150	0.0500	0.8
10	0.0500	0.0500+	1.1
2-4	0.0006	0.0020	4.9
5-7	0.0020	0.0090	4.2
8-10	0.0090	0.0500+	2.7

It is clear that the boundaries of these EDR bins in many places cross over the more traditional orbital region borders, such as those for low-Earth orbit (LEO), high-eccentricity orbit (HEO), and geosynchronous orbit (GEO). For example, objects in all three of those categories could belong to an EDR bin 0; or an EDR bin 6 satellites could be either LEO or HEO. Because tracking phenomenology and mission requirements differ among the traditional orbital regions, it is still useful to continue to divide satellites by these criteria as well. Table 2 outlines the combined division scheme used for the present analysis.

Table 2. Orbit regime and EDR division scheme.

EDR Bin	LEO Period < 225 min	HEO Period > 225 min e > 0.25	GEO Period > 1300 and < 1800 min Inc < 15 deg e < 0.25	MEO Not LEO, HEO, or GEO
0	✓	✓	✓	✓
1	✓	✓		
2-4	✓	✓		
5-7	✓	✓		
8-10	✓	✓		

B. Accuracy versus Tracking Density Functional Form

The Omitron SSNO experiment consisted of obtaining observation datasets for selected objects, creating a set of thinned observational datasets that represent a variety of tracking densities, executing batch differential corrections (DCs) from each of these variously thinned datasets, calculating the RMS error for each such DC, plotting the results, and attempting a curve fit of these results. A variety of different orbit regimes were investigated (although not precisely the same divisions that were used for the current project), allowing some general conclusions about this approach to be drawn. Figure 1, taken from the first of the Omitron study reports, is a typical example of the form and behavior of the data. On this graph is shown several different accuracy (error) versus tracking density curves, each generated using a different orbit determination interval.

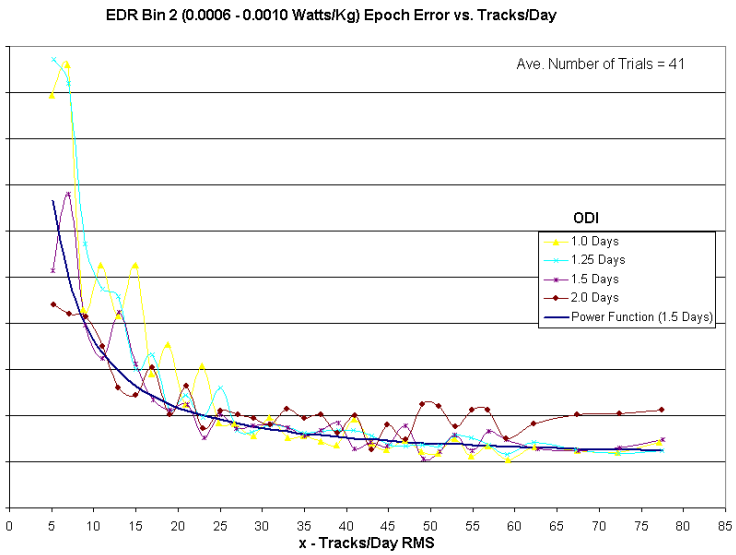

Figure 1. Example accuracy curve set and fit function (from Omitron report)

The first item of note is the overall functional form of the response, which conforms nicely to an exponential decay function. Omitron chose to fit such response to the form $y=a+bx^c$. Whether this is the preferred specific expression of the exponential decay functional form is a question explored in the present analysis, but one can see that the fit (at least in the present example) is at least visually satisfying. The second noteworthy item is the non-monotonic nature of the accuracy error response. The exponential decay seems to bisect the oscillatory behavior, but the fact remains that there is significant movement about this apparent decaying center. Imperfect response is, of course, a necessary consequence of using real rather than fabricated datasets, and one should be prepared to see this type of behavior all the more grandly in a framework that includes even more empirical elements. Finally, it should be noted that the range of tracking densities for the group of satellites shown in the above graph, and in fact for all of the groups examined, is not particularly broad; while it may not be common to exceed an average density of 100 tracks/day or so, it is certainly common to encounter situations in which the density falls well below 1 track/day. For the purposes of studying the effects of reducing the number of tracks collected and thus pushing more and more satellites into such a low-tracking-density region, it is very important that this portion of the curve be accurately represented. Perhaps one of the principal benefits of revisiting the SSNO study would be to extend these curves to significantly lower tracking densities.

C. SSNO Study Limitations for Present Application

The SSNO study proof-of-concept certainly suggests a simple parameterized approach: if one can determine the amount of expected tracking that a family of sensor systems can provide to a given satellite, then using the accuracy-versus-density relationship one can determine immediately the expected accuracy of the vector that would have resulted had simulated observations been generated and an orbit update executed. Such an approach seems highly desirable for the present simulation application, so the temptation is to use the SSNO study results directly as the basis for simulation accuracy determination. Is there any reason why the present SSNO results are not suitable for this purpose?

One reason has already been mentioned: the SSNO study datasets did not explore the region of low tracking densities, a region that is required for simulation applications that propose to reduce the number of sensors and thus the amount of available tracking. There are other limitations as well, principal among them the fact that the SSNO study prescinded from considering the full range of tracking-type mixes that the present SSN supplies. Range-enabled tracks are provided by most SSN radars, angles-only tracks are provided by the optical sensors, and angles-only tracks with a notably degraded quality (attributed in part to a necessary but unfortunate co-ordinate system transformation) from the SSN's single interferometer, the NAVSPASUR fence. Obviously, it is not sufficient to consider tracking density as the only independent variable; the particular type of track (radar, optical, interferometer) and/or mix of tracks will influence the result, at times only subtly, at other times substantially. The SSNO study did

not consider the interferometer observations in a separable way (an appropriate decision, given the study's original purpose, but problematic for the present simulation application) and addressed only a subset of the mix possibilities with angles-only observations. Additionally, the SSNO study did not address all eleven EDR bins within the HEO orbit regime. Finally, the SSNO study calculated its accuracy results in terms of fit RMS, an approach that is meaningful and internally consistent but unfortunately at variance with the approach used to define official accuracy requirements; while the RMS methodology can be used to make relative distinctions among different families of sensor systems, it is not helpful in comparing response to official requirements thresholds. For all of these reasons, a re-execution of the SSNO curve-set is desirable.

III. "Empirical" Vector Accuracy Data using SuperCODAC

The SSNO study executed controlled differential corrections from carefully-selected datasets; this was certainly a reasonable experimental procedure, but it naturally limited the number of satellites that could be considered (around 100) and, therefore, any broad sampling of some of the more obscure orbit types. A natural counterproposal to the SSNO approach is the data mining of actual operational vectors, the data densities that produced them, and their accuracies; while this introduces the messiness of real operational data (and the human and processing errors that it invariably contains), it greatly increases the sample size and does reflect the more typical operational, rather than the merely experimental, situation. The following paragraphs describe the operational data available for capture and the methods to be used to transform it into a dataset appropriate to regenerating a set of SSNO-like curves.

For the last decade or so, special perturbations (precision) differential corrections have been routinely performed on the entire space catalogue; and while the original emphasis was to improve orbit accuracy for objects in the manned-flight regime, the mandate has broadened to improve the vector accuracy for all orbit types. The update method of choice has remained a batch least-squares approach, which has included both the High-Accuracy Satellite Drag Model (HASDM) atmospheric drag modeling[4] and segmented drag solutions, in which instead of solving for a global ballistic coefficient for the entire fit-span, the interval is segmented and a drag coefficient determined for each segment, thus improving the overall solution. Additionally, the operational corrections employ a dynamically-adjusting orbit determination interval (ODI), which is the period over which sensor observation data will be chosen. The approach, called the Dynamic LUPI Algorithm (DLA), attempts to adjudicate the tension between desiring a longer orbit determination interval to reduce errors in the orbit determination process (due to low data rates) and desiring a shorter orbit determination interval to reduce errors due to inadequate drag modeling. This particular approach is described in the 2003 SSNO report, previously cited. What this does mean is that the orbit determination interval cannot be calculated, as it is with the GP theories, as a simple function of satellite orbital parameters; one must know both the orbital parameters and the tracking history and replay the DLA iterative algorithm.

Fortunately, the automated record-keeping that is part of the DC process allows one to determine which sensor observations were included in the orbit determination process that produced each vector, so the tracking mix (amounts of range-enabled, angles-only, and interferometer tracking) that produced each vector is known. Rather than tabulate individual sensor observations, which are often products of essentially arbitrary filter settings and heavily correlated one with another, it makes more sense to consider the individual sensor track—the set of observations that results from one data-taking session—as the individual unit, the quantity and quality of which drives vector accuracy. For NE satellites tracked by ground sensors, a track consists of all of the observations taken during the viable "pass" of the satellite over the sensor; for orbits where sensor coverage of a satellite can endure for long periods, such as for the tracking of deep-space satellites and space-based tracking of many satellites, a track has been defined as the satellite position data taken with less than twenty minutes' separation from any other data points. From these definitions, it can be determined the number of tracks of each tracking type (range, angles-only, interferometry) that produced each vector.

The missing datum, then, is the accuracy (error) of each produced vector; if that were known, then there would be a large dataset that linked, by satellite orbit type, the input tracking density and the resultant accuracy. Fortunately, a feature of the current operational system is an accuracy evaluation of every SP vector produced, with a methodology similar to that used for calibration reference orbit construction. The approach is described more completely in a separate conference paper,[5] but an abbreviated summary can be given here.

A common method of reference orbit construction involves the linking together of pieces of ephemeris taken from the fit-spans of moving-window DCs. If the construction process is set up properly, these fit-spans do not share observational data and are therefore (essentially) statistically independent; and it can be shown that in this case the variance of the reference ephemeris abutment errors (which can be measured) is twice the variance of the actual ephemeris error. When building a small number of reference orbits, the situation can be massaged to get the DC fit-

spans to align properly so that the ephemeris-pieces can be stitched together easily. However, when reference orbits are attempted on an entire catalogue of satellites, in which operational requirements require updates when new observations are available (in eight-hour batch intervals presently), arranging for the ideal fit-span definitions simply becomes impossible; but if some overlap of fit-spans is allowed, a reasonable reference orbit can still be constructed despite the introduction of a certain amount of correlation between ephemeris pieces. In fact, if one can achieve a 50% overlap between fit-spans, the factor of two difference between the variance of the abutment errors and the variance of the actual ephemeris error disappears, allowing a direct relationship. The "adjudicated settlement" for the operational implementation was to take ephemeris pieces from as close to the middle of the fit-span as possible, try for situations that give approximately 50% overlap, and use pieces from intervening updates if necessary. This will allow for the efficient creation of a reference orbit for every object. While clearly not as reliable as satellite laser-ranging (SLR) reference orbits, by some measures they compare favorably: for a collection of eight calibration satellites, epoch accuracies of produced vectors differed, on average, by only 8.3 m when assessed against the constructed reference orbit versus the SLR reference orbit; at the 18-hour prediction point, the mean difference was 5.5 m.

If each satellite has in place an acceptable reference orbit, then accuracy (error) information at time points of interest are calculated easily: the vector under evaluation is propagated to the closest ephemeris point to the time of interest (for the current analysis epoch, 18 hours after epoch, and 72 hours after epoch) and a position comparison made; the vector magnitude of this position difference is the accuracy datum used for the present analysis. At this point, it would seem, all the pieces are in place: each vector update for each satellite has associated with it the number of tracks that produced it and the type of each track (range, angles-only, interferometry) and an accuracy evaluation; this is a very large dataset (about 25 million vectors for the two-year period of 2006 and 2007) that should be able to be mined in a straightforward way to produce a set of SSNO-like accuracy curves.

III. Accuracy Curve-Fitting

An initial approach to establish accuracy-tracking relationships attempted two-dimensional curve-fits, in which the resultant accuracy was presumed to be a double-exponential decay as a function of the range-enabled and angles-only tracks per LUPI. Unfortunately, these surface fits proved unwieldy to evaluate and visualize, and in any case do not save difficulty when interferometry tracks are included, as four-dimensional fitting becomes truly unsustainable. From this experience, it was decided that the fit-space should be constrained to a single independent variable, which would be the number of tracks per ODI. Because of the different track-mix possibilities, certain canonical mix-types would need to be defined. Table 3 outlines the tracking mix types used for the current analysis.

Table 3. Tracking mix types and definitions

Tracking Mix Type	Description
(1) Radar only	Most satellites in LEO orbits receive a mixture of radar and interferometer tracks; most deep-space satellites receive a mixture of radar and optical tracks. This category addresses the case in which only radar tracks are obtained.
(2) Radar and others (LEO only)	As stated above, most LEO orbits receive radar and interferometer tracks; so this category is intended to represent the typical LEO situation. There is a small number of optical tracks that also appear here (generally from space-based assets), and these are included as well.
(3) Optical only (DS only)	Only angles-only data from optical sensors is included.
(4) Interferometer (NAVSPASUR) only	The SSN contains currently one interferometer radar (the NAVSPASUR fence). There are cases in which satellites are maintained only with data from this asset; and these cases are considered here.
(5) Radar and optical only (DS only)	DS objects maintained with both radar and optical (but not NAVSPASUR) data—the typical DS case
(6) Radar, optical, and NAVSPASUR only	DS objects that receive all three tracking types

This is a good time to consolidate and expand upon all the divisions of data discussed so far in this paper, as it is the superimposition of all of these divisions that will create the individual datasets to be fit. The first division is by orbital regime, as outlined in Table 2. There are five EDR bin divisions for LEO, five for HEO, one for GEO, and one for MEO, giving a total of twelve groupings for orbit regime. Next is the division by tracking mixes, as given in

Table 3. Three tracking mixes apply for LEO (mixes 1, 2, and 4); five mixes apply to the other orbit regimes (mixes, 1, 3, 4, 5, and 6; GEO has only 1, 3 and 5). The nested application of these two divisions produces 49 dataset divisions. Each of these needs to be further subdivided into time-since-epoch groups, as the accuracy data were analyzed for epoch, 18 hours after epoch, and 72 hours after epoch (making 147 datasets total). Finally, one must determine what statistical accuracy value should be selected to represent the vector accuracy for a given tracking density. Because Vmag accuracy is an asymmetric, right-skewed distribution that does not precisely follow any of the canonical distribution types (largely due to the fact that the individual component errors do not in fact follow a Gaussian distribution), it is preferable to use a non-parametric characterization technique, such as percentiles. Each set is thus evaluated with 50^{th} and 95^{th} percentile accuracy values. One thus reaches a grand total of 294 accuracy curves to attempt to fit.

It was additionally required that any given dataset have at least fifteen data points in order to attempt a meaningful fit. This criterion excluded sixteen datasets from the fitting process, leaving a grand total of 278 to be fit and subsequently evaluated.

IV. Results

A. Motivation for and details of exponential-decay fitting

The Omitron Corporation SSNO study, discussed previously in this paper, established as proof-of-concept the exponential decay functional form of the relationship between vector error and tracking density. The precise functional form that Omitron proposed is the following:

$$y = a + bx^c; \tag{2}$$

and this seems a logical form with which to begin, although it seems also appropriate to try the similar (the first few terms of the Taylor series expansion are the same) but competing form

$$y = a + be^{cx}, \tag{3}$$

as it may have preferable curvature and asymptotic properties.

One approach to determining quickly whether the Omitron situation is indeed a *comparandum* to the accuracy-density relationship in the empirical data is to investigate the Pearson correlation coefficient in log space; if a large negative correlation is observed, one can expect the two situations to be comparable. A CDF of this value is given in Fig. 2, and it is clear that the inverse correlation is quite strong, as expected: 50% of the values are smaller than -0.8 and 95% are smaller than -0.4. So an exponential functional form seems quite appropriate to the empirical data.

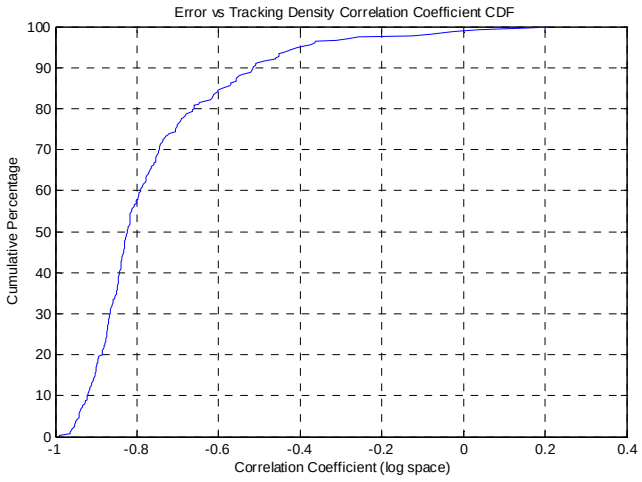

Figure 2. CDF plot of Pearson correlation coefficient (in log space).

With datasets of this form, one faces a difficulty in selecting the particular curve-fitting approach to use. Both proposed functional forms are non-linear in the independent variable, so a non-iterative solution will not exist. Nonlinear least-squares certainly suggests itself as the logical candidate, as it will properly represent the expected peak at the left side of the curve (the high accuracy errors corresponding to low tracking density). However, if the data are at all noisy, the miscarriages that least-squares fitting can exact are all too well known. One is tempted, therefore, to attempt a robust fitting technique in order to reduce the effects of outliers and other data noise; what was not known is whether such a technique would deweight the curve's left side that it will no longer be representative of the accuracy function, but an attempt was certainly warranted. The robust technique selected is the iteratively-reweighted bisquare approach, in which the data are iteratively reweighted according to the weighting function

$$w_B(e) = \left[1 - \left(\frac{e}{k}\right)^2\right]^2, |e| \le k; \quad 0, |e| > k, \tag{4}$$

in which w_B is the vector of weights, e is the residual error, and k is a tuning constant. The functional forms in Eqs. (2) and (3) were thus both fitted with a straightforward nonlinear least squares approach (Levenberg-Marquant) in addition to this robust technique.

After an initial trial run with this approach, it became obvious that many datasets possessed long tails that decayed slowly; and the fits to the functional forms given in Eqs. (2) and (3) produced asymptotes that settled at too large a value. A linear term was thus added to the functional forms to allow the fitted response to follow the actual decay more closely; the functional form is thus

$$y = a + bx + ce^{dx}; \quad y = a + bx + cx^d \tag{5}$$

Of course, allowing this kind of decay creates the possibility that the fitted form might eventually assume negative accuracy values; and this potential problem was addressed in the following way. Predicted accuracy values for tracking densities larger than any used in the fit-span were assigned the fitted value for the end of the fit-span; this eliminates the possibility of a negative value arising from a predictive use of the fitted curve. If negative fitted values were encountered within the fit-span itself, the root of the fitted equation was determined, the median value of the ten data points preceding the zero was calculated, and the fitted curve was forced to this constant value at all points after which it naturally occurred. This latter approach is somewhat arbitrary in its construction but nonetheless has proven quite serviceable.

B. Fit Quality

It is somewhat premature to introduce fit quality at this point, since it would be more naturally discussed after all the details of the fit-execution process have been addressed. However, this issue is best treated by introducing sample graphs that demonstrate different fit qualities; and showing these graphs at this point will help to anchor in something concrete all of the preceding discussion, which by taking place apart from specific examples has remained somewhat abstract.

The fitting software produced plots that show the actual data as circles and the four fitted functions as different-colored lines. If all of the fit approaches used a least-squares minimization as the cost function, then a single parameter, such as the standard error of the estimate or an RMS summary, could be used as at least a preliminary indication of the relative success of the four fit types. Of course, the virtue of the robust techniques is that they do not use a least-squares minimization, so it is certain that their resultant RMS will be worse than that for a least-squares approach. As a replacement, two percentile points—the 50th and the 90th—of the absolute value of the residual set have been chosen. The 50th percentile level gives a sense of the overall behavior of the fit and the 90th percentile insight into what range the grosser residuals might occupy. With larger datasets, to be sure, one could encounter several very large residuals not reflected by either of these statistics, and an overall favorable residual picture does not guarantee a desirable fit. As such, the assignment of the best fit for each dataset was ultimately made subjectively, by eye.

This visual examination process assigned a single best functional form, a runner-up functional form, and an overall quality-of-fit ranking. The best fit was selected on the basis of how well it was able to represent the peaked, early part of the response (the more important one for the present application, as the error increase caused by any tracking decrease below the visible "knee" in the curve needed to be accurately reflected) and not badly misrepresent the tail of the dataset, where there was comparatively little error variation. That said, there were situations

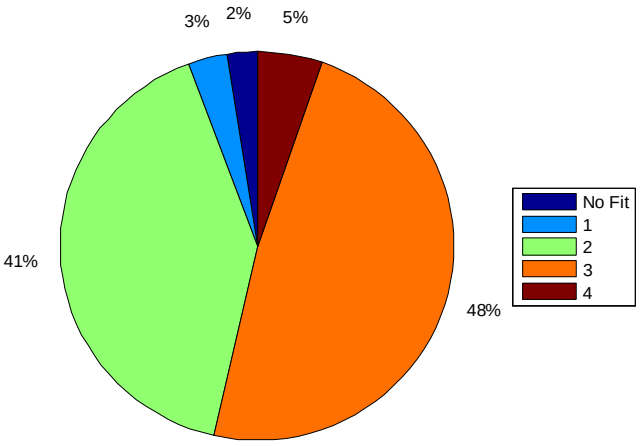

Figure 4. Distribution of fit quality. *Most of the fits merit either a 3 or 2, with the 1 and 4 ratings composing less than 10% of all of fits.*

in which much of the "peak" of the early part of the response was caused by a single point, and it was difficult to have an abiding confidence in that single point; so fits that did not reach or represent that point completely but provided a compelling fit to the rest of the data could emerge as the preferred choice. The 50th and 95th percentile residual statistics were consulted as part of the selection but were not determinative; the one role that they fulfilled consistently was to guide the choice between two equally visually satisfying fits.

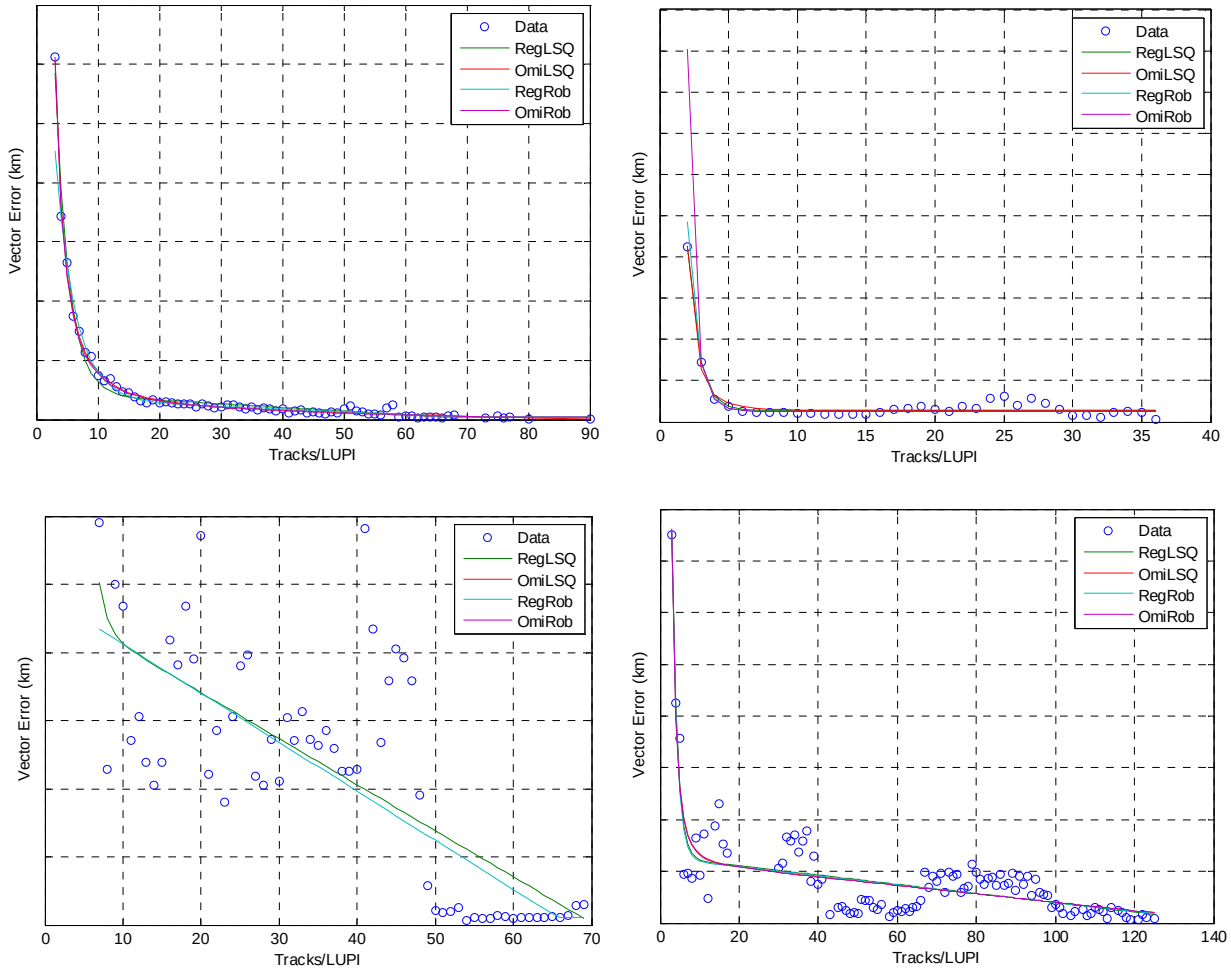

Figure 3. Examples of four different fit qualities. *Clockwise from top left, the assigned quality values are 4, 3, 2, and 1, respectively.*

Fit quality was evaluated on a scale of 1 to 4, 4 being high. The assessment is naturally subjective, although it of course correlates largely with the absolute residual percentile figures; and it means that at least one of the four fit types provided a fit of the specified quality. Figure 3 gives an example of a dataset corresponding to each of the four fit qualities. A quality level of 4 indicates a fit that is good essentially throughout. A level of 3 indicates acceptable deviation in either the peak section or the tail. In the Fig. 3 example (upper right), the aqua line is an acceptable deviation in the peak section, whereas the purple one is not. A level of 2 indicates a noisy but tolerable fit in both the peak and the tail regions. A level of 1 indicates a situation that really cannot be said to fit either of the functional forms. Figure 4 (previous page) provides the overall composition of different fit qualities. As can be seen, qualities of 3 and 2 constitute essentially 90% of the cases, with the more extreme 4 and 1 evaluations relegated to the remainder.

C. Least-squares versus Robust Regression Performance

The principal advantage of least-squares fitting is its reliability, and the principal advantage of robust regression is its ability to present a more convincing fit in the presence of noisy data. Figure 5 is a combination graph (and a potentially somewhat misleading one, in that it actually combines two conceptually different results) that shows both of these virtues in play in the present dataset. One observes the much greater reliability of the nonlinear least-squares fitting: these fits succeed (converge without errors) nearly all of the time, whereas the robust regression cases succeed only about two-thirds of the time. However, when the latter do succeed, they tend to be the preferred fit; they are selected about 40% of the time as the fit of choice, whereas the least-squares fits are selected only about 25% of the time. These results encourage two distinct actions: first, to try to improve the success rates of the robust fits; and second, to keep both fit-types in the mix so that there is a fallback for the expected failures of the robust fits.

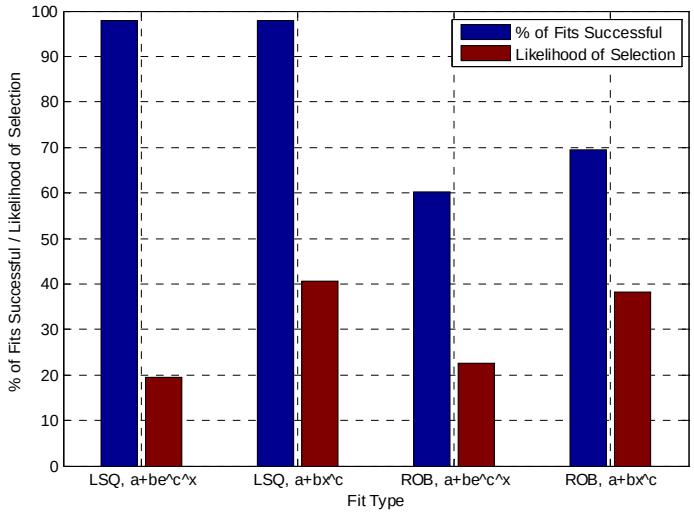

Figure 5. Fit success and likelihood of selection by fit type.

D. Preferred Functional Form

In choosing the best fit for each of the 278 datasets, both the best and the runner-up of the four fits was identified. The purpose of choosing both a first and second place was to help to determine whether fit functional form or fit approach (least-squares or robust) was the greater contributor to ultimate success. Figure 6 gives two pie

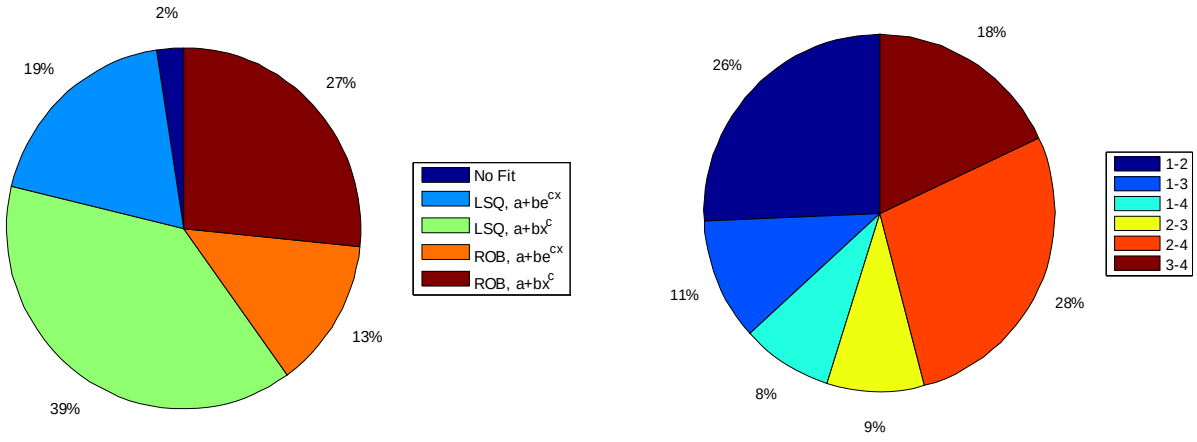

Figure 6. Best fit composition and first-second place combination frequency. *The left graph shows the portion of best-fit cases represented by the four fit-types (and absolute failure, a fifth type). The right graph shows the different combinations of the four fit types and the frequency that these combinations (and their reciprocals) occurred.*

composition graphs to help make this decision. The leftmost graph is a simple composition graph of the winning fit-type. From this chart, it is clear that the Omitron-recommended functional form, $a+bx^c$, prevails as the form of choice, accounting for about two-thirds of the winning assignments (if both the least-squares and robust forms are considered). Staying within the least-squares or robust venues, the Omitron form outperforms the alternative by almost a factor of two: 19 to 39% and 13 to 27%. The rightmost graph draws the runner-up form into the mix to see whether a particular functional form prevails frequently as the top two assignments. The functional form / fit type pairings follow the legend for the left graph (minus the "No Fit" category): 1 is LSQ, $a+be^{cx}$, 2 is LSQ, $a+bx^c$, &c.;

the 1-2 designation means either type 1 was first and type 2 was the runner up, or the reverse (type 2 was first and type 1 was the runner-up). The data are somewhat difficult to interpret because robust fit failures often occurred in pairs (meaning both types 3 and 4 failed), which forces the assignment of that case to the 1-2 category even if there is no intrinsic pairing between those two categories. Be that as it may, it is clear that the 2-4 category dominates, which aligns with the overall superior performance of the Omitron functional form. The 3-4 category similarly outstrips the 1-3, 1-4, and 2-3 categories, which speaks to the superior performance of the robust fitting. If one were to select a single functional form and fit type, it is clear that it would be the Omitron form and robust fitting.

E. Residual Properties

The method of residual evaluation appropriate to a mix of least-squares and robust fit methods was earlier described; it is to take and examine the 50th and 90th percentile points of the absolute values of the residuals for a given fit-type. These percentile points (hereafter abbreviated as 50ile and 90ile) can be collected for each fit-type for each dataset and then further summarized by taking the 50th, 68th, and 95th percentile of the entire collected set for each fit-type. These results, given in graphical form in Fig. 7, establish a level of fidelity for the overall project of representing state vector accuracy from the fitted functions.

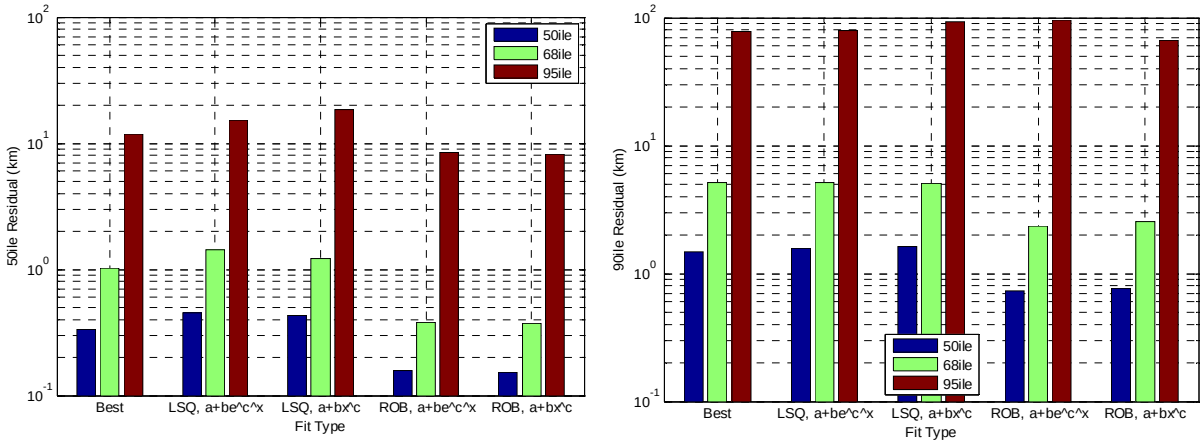

Figure 7. Residual properties versus fit-type. *The 50ile case is given in the left-hand graph; the 90ile case on the right.*

It is clear immediately that the robust fit results significantly outperform the least-squares approach, but one is drawn instead to the values of the residuals themselves. The "best fit" results range from a median value of 300m to over 10km at the 95th percentile, and that is just for the 50ile division; if one moves to the 90ile level, then even at the median level the "best value" results exceed 1km, and the 95th percentile value is close to 80km. This is not absolute vector error, of course; it is merely the ability of the fitted curve to reflect the actual datasets. As one is generally used to smaller-valued residual sets, this set stands out simply for its largeness.

There are three items that must be kept in mind. First, the accuracy curves have been deliberately extended into low-tracking density areas, in which the vector errors are much larger and performance much more erratic than for the datasets that are generally studied. Many of the large residuals come from these portions of the accuracy fits. In a percentile sense, by providing more upper-end values, they both produce unusually high values for the higher percentile points and similarly increase generally more damped percentile points (such as the median) by pushing what are generally higher values down to that percentile level. Second, the goal of the present approach is not infallibly to represent the state vector accuracy of a given object; rather, it is to allow the reasonably accurate characterization of an entire catalogue of objects. While the approach is generally not reliable, or is reliable only by accident, for small groups or individual objects, for a large catalogue in the context of a simulation run over a reasonably long simulation period the statistical properties of the fits can exercise themselves; and large residuals thus do not affect the overall ability to predict accuracy in the aggregate. Third, the results in Fig. 7 presume that each orbit regime is equally important for the typical space catalogue distributions encountered in simulations; but the percentages given in Table 1 make clear that this is not in fact the case. In fact, nearly 70% of the catalogued objects are contained within just one of the fifteen orbital regimes used for this study. A more appropriate statement of the residual error would be a representation weighted by the sample population of each of the orbital regime bins. Because non-parametric statistics have been used up to this point, it makes sense to provide weighted statistics in

this same framework: a given percentile measurement will be duplicated the same number of times as there are object samples in the dataset, all of these duplicated measurements will be placed into one long vector, and the vector will be sorted and percentile points extracted from it. This will give weighted percentile points that can be plotted in the same framework as the equally-weighted results of Fig. 7.

These plots are provided in Fig. 8, and the results are much more sanguine. For the 50ile case, two-thirds of the datasets have residual results less than 200m, and the 95th percentile value is only 500m. The 90ile case is of course less satisfactory, but one must remember that this is a compilation of the worst results for each case; in this light, the fact that two-thirds of the cases are better than 600m is not nearly so disappointing; and the overall "worst of the worst" situation is within the accuracy of a typical GP element set in GEO.

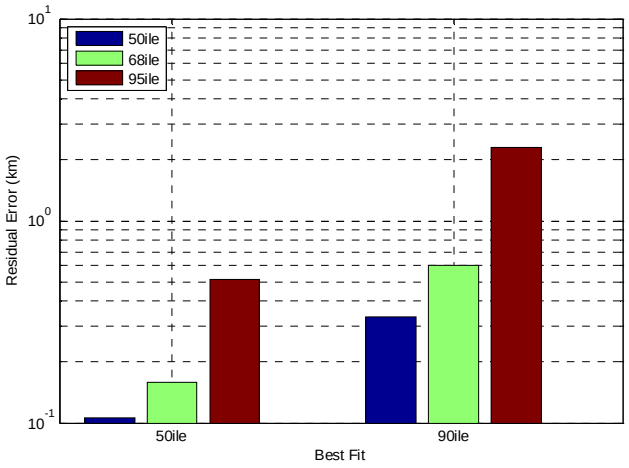

Figure 8. Weighted residual values for the "best fit" situation.

V. Conclusions and Future Work

The present analysis has outlined a viable technique for greatly reducing the complexity of the state vector accuracy calculation while still maintaining acceptable fidelity. The preferred functional form for the recommended accuracy decay curves has been developed and the preferred fit approach identified. The divisions of orbital regimes and tracking mixes produced datasets that, for the most part, generated visually satisfying fits. The residual sets, when taken in an unweighted sense, were somewhat larger than one would like; but when presented as weighted statistics they were much more palatable.

Two activities suggest themselves to improve and buttress this approach for broader application in SSN simulations. The first is a rigorous bias analysis of the residual sets to ensure that they do in fact have unbiased statistical properties. From the visual inspection of the graphs, this appears to be true; but it should be confirmed formally. Through such an investigation the approach can be recommended for large-catalogue simulations without any reservations, as any larger-than-desired residuals will not affect aggregate assessments of catalogue state vector accuracy. The second activity is to improve the stability of the robust regression fits so that they can contribute to more of the fit datasets. A data conditioning routine may need to be written to detect underflow or other situations in which residual weights, once inverted, produce boundary violations for the fitting routine. The degree to which this fit technique reduces the overall residual error is clear, and in most cases it does not jeopardize the "left peak" of the curves; so additional effort to broaden its applicability for this particular problem would pay sizeable dividends.

Acknowledgments

Special thanks are due to AFMC/SMC/SY and Northrop Grumman Mission Systems for sponsoring the research described in the present paper. Members of the AFSPC Catalogue Accuracy Assessment Working Group are thanked both for their encouragement and review of intermediate products. Finally, the author would like to express his appreciation Drs. Joe Raquepas and Dan Snow, both of AFSPC/A9A, for their valuable suggestions in response to earlier versions of this study effort.

References

[2]"Space Surveillance Network Optimization (SSNO) Study," ESC/NDS Technical Report B001-3.1.4-SSNO-RPT-1, Colorado Springs, CO, August 2002.

[3]"Space Situation Awareness Command and Control 2003-03 Support and Development Project (Final)," ESD/NDC Technical Report B002-5.7.4-SAC2-1b_1e_1f-RPT-02, Colorado Springs, CO, October 2003.

[4]Storz, Mark F.; Bowman, B. R., and Branson, J. I.; "High Accuracy Satellite Drag Model (HASDM)," AIAA-2002-4886, AIAA/AAS Astrodynamics Specialist Conference (Monterey, California), Aug 2002.

[5]Hejduk, M.D., Ericson, N.L., and Casali, S.J. "Beyond Covariance: A New Accuracy Assessment Approach for the 1SPCS Precision Satellite Catalogue." 2006 MIT / Lincoln Laboratory Space Control Conference, Bedford, MA. May 2006.