Covariance Manipulation for Conjunction Assessment

M.D. Hejduk[1]

Astrorum Consulting LLC Woodway, TX   76712 USA

The manipulation of space object covariances to try to provide additional or improved information to conjunction risk assessment is not an uncommon practice. Types of manipulation include fabricating a covariance when it is missing or unreliable to force the probability of collision (Pc) to a maximum value ("PcMax"), scaling a covariance to try to improve its realism or see the effect of covariance volatility on the calculated Pc, and constructing the equivalent of an epoch covariance at a convenient future point in the event ("covariance forecasting"). In bringing these methods to bear for Conjunction Assessment (CA) operations, however, some do not remain fully consistent with best practices for conducting risk management, some seem to be of relatively low utility, and some require additional information before they can contribute fully to risk analysis. This study describes some basic principles of modern risk management (following the Kaplan construct) and then examines the PcMax and covariance forecasting paradigms for alignment with these principles; it then further examines the expected utility of these methods in the modern CA framework. Both paradigms are found to be not without utility, but only in situations that are somewhat carefully circumscribed.

Nomenclature

CA	=	Conjunction assessment
CDF	=	Cumulative distribution function
OCM	=	Conjunction data message
GP	=	General perturbations
HBR	=	Hard body radius
OD	=	Orbit determination
Pc	=	Probability of collision
PcMax	=	Maximum probability of collision
SP	=	Special perturbations
TCA	=	Time of closest approach

I.   Introduction

SATELLITE conjunction risk analysis relies heavily on the calculated Probability of Collision (Pc) between two space objects. To assess the risk of any event involves determining the likelihood of its occurrence, and the only practicable way to represent an event's likelihood is through a probability. This heavily reliance on the Pc emerged from the experience of trying to assess collision risk with satellite miss distances alone. Since satellite position uncertainty information is not present in miss-distance point estimates, it is extremely difficult to determine if a particular miss distance estimate is significant; likelihood estimates based on this parameter alone have thus been seen as not sufficient to enable collision risk analysis, directing the industry's attention to the Pc. Because the orbit determination (OD) covariance matrices for both the primary and secondary objects are necessary to calculate the Pc, a substantial interest has developed in securing and ensuring the propriety of these covariances.

The importance of the OD covariance to the conjunction assessment (CA) community has further spurred the development of methodologies to manipulate it in order to attempt to improve risk assessment. Such approaches are generally of two types, and the first of these is to try to compensate for known or suspected deficiencies in the covariance. Formulae have been developed to determine the maximum Pc possible when one or both of the objects' covariances are missing or so poor as to be rendered completely invalid. Such is the case, for example, when one is

---

attempting CA with the published two-line general perturbations (GP) element catalogue, which lacks covariance information; or an owner-operator-generated ephemeris that does not include covariance. In these situations, this maximum value, called PcMax, provides an upper-bound Pc for the present conjunction and thus could allow a worst-case risk assessment and potentially enable remedial action. Alternatively, compensatory covariance manipulation could be performed by recognizing that there is a certain inherent error or uncertainty in covariance formation and exploring whether (relatively) minor changes to the covariance size and/or aspect ratio could have a notably deleterious effect on the Pc. This can be done through what is called "k-scaling," in which the combined covariance in the conjunction plane is multiplied by a set of scale factors and the maximum Pc determined, or by scaling the primary and secondary covariances separately through nested application of laddered scale factors. Rather than seeking an overall maximum, such approaches are formulated to remain within a bounded set of scale factors that it is believed represent the uncertainty range of the covariances.

A second purpose for which covariances are altered is to try to forecast the size and orientation of a satellite's epoch covariance at some future time. This activity is different in kind from the usual propagation of covariance into the future. After a conjunction between two satellites has been identified through a conjunction screening process, it is routine to propagate the two satellites' states and covariances forward to the objects' time of closest approach (TCA) and then calculate the Pc from this set of propagated states and covariances. The covariance-altering prognostication technique, sometimes called Pc Forecasting, is to attempt to construct an expected *epoch* covariance at some future point (by attempting to predict the amounts, sources, and times of satellite tracking that will be received between the present time and this future point, which are the main input to covariance formation) and then propagate that covariance forward to TCA and calculate the Pc. It can be used as an attempt to predict the risk assessment situation as it might occur at a critical decision point (such as two days before TCA, a common maneuver commitment point) as a way of guessing at the criticality of the situation when it will most matter. Some practitioners even try to prognosticate the epoch covariance at TCA to produce a quasi-definitive risk statement.

Both of these covariance manipulation approaches will be examined in the present analysis, with an eye to determining their appropriate utility given the modern practice of CA conducted with a precision catalogue. The treatment will begin with a philosophical examination of risk analysis and the standard framework that has emerged over the last few decades for performing risk analysis reliably. After this, the two different covariance manipulation types will be examined individually. Pc maximization and covariance scaling will be given both a conceptual and quantitative treatment. Covariance forecasting will be given a conceptual and numerical treatment by proxy; while this will not serve as a full evaluation, it will allow certain conclusions to be drawn. A planned second part to this paper will address covariance scaling activities and give a full quantitative treatment of covariance forecasting.

II. Risk Analysis Overview

There is certainly a large number of different risk management methodologies in use presently in industry, and these span from extremely loose, heuristic structured thinking models to extremely rigorous quantitative methods. Most of these quantitative methods, however, draw their foundational principles from the 1981 seminal article of Kaplan and Garrick[1]. There are two major, framework-establishing concepts outlined in Kaplan's treatment, one of which is reasonably commonly known and a second that is less frequently encountered; both of these will be discussed here in the context of CA.

The first of these concepts is what has come to be known as the "Kaplan Triplet." In trying to understand the risk posture for a certain enterprise, all of the possible outcome scenarios (S) are enumerated. Each one of these scenarios has some certain likelihood or probability of taking place (P) and some particular consequence (X) should it take place. The enumeration of these items for *n* scenarios thus produces a set of "triplets" as shown below:

$< S_1, P_1, X_1 >$
$< S_2, P_2, X_2 >$
$\dots$
$< S_n, P_n, X_n >$

If for all of these triplets the P and X points are plotted on a two-dimensional Cartesian grid (usually with Consequence (X) on the x-axis and Probability (P) on the y-axis, the generated curve is called a "risk curve," an example of which (from the Kaplan paper) is given in Fig. 1.

One can see the implications of formulating risk this way: risk is defined as the interplay between the likelihood of unfavorable events and the consequence of such events should they happen. Some practitioners go further and define risk not just as the relationship but as the *product* of likelihood and consequence; such an arrangement can be suitable for many situations, but it flows from a risk-neutral posture that is willing to accept essentially any consequence level, whereas some consequences may be judged to be so high that they push the risk much higher than the simple product would dictate.[2]

In conjunction risk analysis, the move to the use of Pc rather than miss distance has brought the discipline more into conformity with this paradigm by articulating, or at least attempting to articulate, the actual likelihood of a collision. However, the accompanying part—consequence—has been addressed hardly at all; most approaches consider all potential collisions to be the same in terms of consequence, regardless of whether the satellite could potentially survive the collision, the amount of debris that a collision could be expected to generate, and the orbit corridors that would be polluted with this debris, both shortly after generation and well into the future. The set of Kaplan triplets really becomes just a single triplet, with "collision" as the single consequence and the calculated Pc the likelihood. There is clearly room for an expanded construct here that considers the consequence aspect of each potential collision scenario, but the important point for the present discussion is that a true likelihood assessment is a constitutive element of each Kaplan Triplet. The Pc, when properly formulated, provides the estimate of likelihood.

The second foundational concept in Kaplan's framework is the understanding that any given Kaplan Triplet is only one instantiation of an estimate of the likelihood and consequence of a particular scenario. The inputs to the Pc calculation (covariances and hard-body radius) have uncertainties associated with them; so when a single Pc value is generated, this point estimate is just one instantiation of an entire probability density of Pc values. Estimates of collision consequence, should they be attempted, present the same situation: one would expect a probability density of consequences rather than a single estimate. Kaplan represents these probability densities as a series of contours, each with some likelihood of instantiation; his graph is reproduced here as Fig. 2.

CA risk assessment has in general not incorporated such a view in its operational analysis methods. A single Pc value is calculated and used as the full representation of likelihood; and given that a static understanding of consequence is applied to every conjunction, there is certainly no effort expended to determine the uncertainty of that consequence actually arising should a collision occur. The NASA Conjunction Assessment Risk Anlaysis (CARA) team has begun a formal incorporation of input parameter uncertainty into its Pc calculations, thus producing probability densities of Pc for each conjunction rather than a single value[3]; other practitioners are pursuing this type of capability as well. The important point here, however, is that statements of likelihood and consequence themselves have uncertainties and that these uncertainties need to be considered in formulating robust risk curves.

In evaluating different CA covariance manipulation methods, one must examine in what manner and to what degree they improve CA risk analysis. These two key foundational concepts from Kaplan are proposed here as a framework for making this evaluation. While it is certainly possible to lend aid to the risk analysis enterprise in a
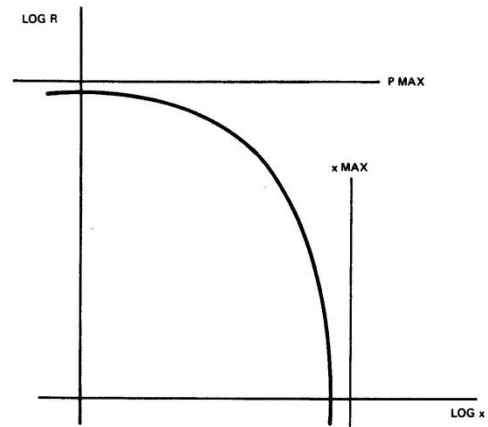

Figure 1. Kaplan likelihood vs consequence risk curve. *Likelihood is given on the y-axis and consequence on the x-axis.*
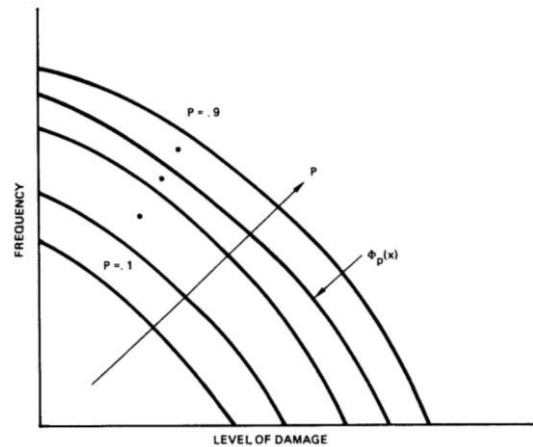

Figure 2. Kaplan risk probability density. *Contours are lines of constant probability.*

manner not directly addressed by these two concepts, proposed improvements that are at variance with or fail structurally to make room for them will in all likelihood detract from rather than enhance CA operations.

III. Absolute and Relative Pc Maxima

The two-dimensional Pc calculation, in which the short conjunction duration allows the assumptions of rectilinear motion and a static covariance in the vicinity of the encounter and thus the reduction of the dimensionality of the problem from three to two, is the standard technique used by most CA risk assessment toolsets.   The full details of the calculation are thoroughly explained in Chan (2008),[4] but the conjunction geometry is usually visualized in the conjunction plane (plane normal to the relative velocity vector), and the Pc is the portion of the combined position uncertainty density that falls within the objects' combined hard-body radius (HBR) area, as shown in Fig. 3.
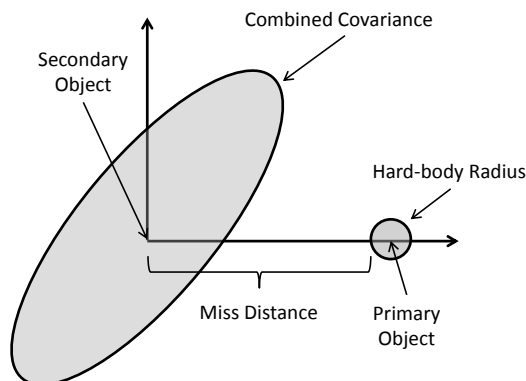
Some years ago, when general perturbations (GP) element sets, which lacked a covariance, were the only available set of state estimates for conjunction secondary objects, it was recognized that the full possible range of sizes, aspect ratios, and orientations of the combined covariance in the conjunction plane could be explored to find the particular arrangement of these three attributes that produced a maximum Pc value. This concept appears to have first entered the literature in Alfriend *et al.* (1999),[5] who recommended it for situations in which the reliability of the two objects' covairances was suspect. The approach was studied thoroughly and methodically by Alfano (2005),[6] who developed a number of useful approximations for calculating the maximum Pc (hereafter abbreviated "PcMax"). He provides several different approximations for PcMax when the combined covariance aspect ratio is known or can be reasonably guessed, and he evaluates these approximations' accuracy for aspect ratios from 1 to 50. He subsequently provides a derived formula and simplifying approximation for cases in which the aspect ratio can be considered essentially infinite (> 1000)—the maximum Pc is reached when this degenerate ellipse lies along the relative miss vector in the conjunction plane. Frisbee expanded this latter concept to consider the case in which one of the two conjuncting object's covariances is known but the other is not.[7] A procedure similar to Alfano's is applied to the object without covariance, but the combination with the known covariance pulls uncertainty density away from the degeneracy along the miss vector and thus lowers the maximum Pc value. In some circumstances, this tempered maximum may be significantly smaller than that for the unencumbered PcMax.

Figure 3. Pc calculation elements projected into conjunction plane.

Neither of the researchers cited above believed that the calculated PcMax, in either the constrained or unconstrained form, was a substitute for an actual calculated Pc; and in their work they both are quite careful to articulate limitations for the use of this construct. While very much endorsing their statements of limitations, the present author believes the actual utility of PcMax calculations may be even more constrained. The following commentary is an attempt to examine the many different ways in which the PcMax construct can be misused.

A. PcMax is not a Probability of Collision in the True Sense

A "regular" Pc represents the likelihood that two space objects, given their state uncertainties, will pass each other with a distance less than the combined hard-body radius. The PcMax construct, in contrast, postulates a set of conditions on the objects' covariances that will push this calculated Pc value to a maximum. What it does not do, however, is provide an adjoining statement of the likelihood that those particular conditions will in fact arise. Because it is merely a maximum without any sense of the likelihood of this maximum's actually occurring, it does not represent the probability of a particular risk scenario—it cannot be used in the context of a Kaplan Triplet. To make PcMax into a probability that perhaps could be used in this way, it would need to be combined with a calculation of the likelihood of its occurrence. Given that the PcMax construct is being pursued in the first place because of significantly reduced knowledge about the conjunction situation, it is extremely unlikely that one could calculate a reasonable likelihood of occurrence of the antecedent conditions that make PcMax possible. The PcMax does of course describe a scenario in which a particular type of consequence—the highest Pc possible— is realized;

but because the likelihood of this scenario's arising is unknown, this information does not have a role in a risk analysis context.

B. PcMax is not Practically Useful as a Screening Parameter for the Present

While there have been only occasional attempts to deploy PcMax as an actual risk assessment metric, its use as a conjunction screening threshold has been much more common. While it cannot give the actual risk of a conjunction, as its name implies it does give the maximum Pc that could be obtained for the particular conjunction; so if the PcMax value is below one's screening threshold for conjunctions of interest, one can safely ignore this particular conjunction for the present (whether it can safely be ignored for the future is treated in the next section). As an abstract principle this statement is certainly cogent; what remains to be established is whether PcMax, in either its unconstrained or constrained form, is useful practically as a conjunction screening filter.

To examine this question, a subset of the CARA Orbital Conjunction Message (OCM) database was selected for analysis. An OCM is generated whenever, as part of a conjunction screening activity that extends some number of days into the future, a catalogued object penetrates a screening volume ellipsoid centered on a protected primary asset. OCMs were collected for eleven of the satellites that CARA protects (all taken from a near-circular ~700km orbit, which is the orbit band that tends to generate the most conjunctions) for a one-year period

Figure 4. Miss distance CDFs for analysis dataset, separated by color.

(May 2015 to May 2016)—a dataset of approximately 72,000 OCMs. One useful way to organize such data is by the severity of the calculated Pc, and a "green-yellow-red" color scheme is used operationally for such categorization. Green events are those with a Pc so low that the event does not merit further examination; such events have Pc values that range from 0 to 1E-07. Yellow events are those that, while not meriting remediation at present, have the propensity to become serious; the range of Pc values for this category is 1E-07 to around 1E-04. Red events are those that are considered serious and in some instances will merit remediation; such events have a Pc greater than 4E-04 (according to current CARA practice). In order to make the present analysis somewhat more ecumenical and embrace round numbers, for the purposes of this analysis "red" events are those with a Pc greater than 1E-04. For the dataset under investigation, about 60,000 OCMs were categorized as green, 9000 as yellow, and 2000 as red, corresponding to percentages of 85%, 12.2%, and 2.8%, respectively.

The screening volume used for this particular orbit regime is an ellipsoid centered on the protected primary with semi-major axes of 0.5 km in the radial direction, 17 km in the in-track direction, and 20 km in the cross-track direction. Screening volume sizes are set based on desired capture percentages and processing limitations, but this volume size has been used for some years with satisfactory results. A

Figure 5. Maximum Pc as a function of miss distance. *Presumes a constant HBR of 20m.*

volume of this size will, of course, restrict the range of conjunction miss distances observed, but miss vectors of some consequence are still encountered. A cumulative distribution function (CDF) plot of the miss distances for the conjunction dataset, organized by color, is given in Fig. 4. While the miss distances for the red OCMs are, as
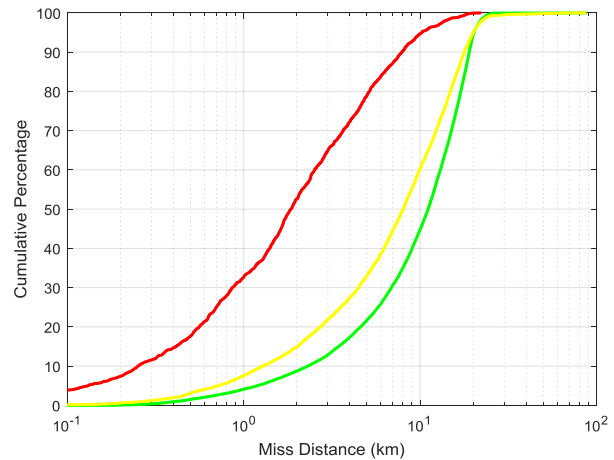
expected, notably smaller than those for yellow and green OCMs, the median miss distances for both yellow and green are close to 10km and therefore not by any means small.

Alfano's formula for PcMax is given in Equation 34 of his 2005 paper, for which he then gives a very good linearized approximation in Equations 35 and 36; over the miss distances in play for the present analyses, the two expressions produce results that are are essentially indistinguishable. Both formulae render PcMax as a function of the ratio of the HBR to the miss distance; but by choosing to operate at a fixed HBR of 20m (the value used by CARA for most of its protected assets), the expressions become a function of miss distance alone; and this linear relationship (in log-space) is shown in Fig. 5. One can see immediately that, for routine levels of the red threshold, only conjunctions with extremely large miss distances—much larger than any miss distances permitted by the screening volume used for the present dataset—will produce PcMax values smaller than the nominal red Pc threshold of 1E-04. The actual distributions of PcMax values for the present dataset, given as CDF plots, are provided in Fig. 6. Even the more restrictive CARA red threshold of 4E-04 would produce essentially no screening exclusions

Figure 6. Unconstrained maximum Pc CDFs for analysis dataset, separated by color.

by using the PcMax formulation. Larger screening volumes are, of course, used for certain orbit regimes, and for these a PcMax-based pre-screening activity could perhaps identify more candidates for elimination. However, screenings in these orbit regimes also tend to identify far fewer conjunctions; so the absolute benefit of this pre-screening in such situations would be relatively small, especially given the ease of calculation of the 2-D Pc.

Alfano also developed a more bounded methodology for the maximum Pc calculation for those cases in which the aspect ratio of the combined covariance projected into the conjunction plane is known or could be reasonably guessed. He further profiled a series of GP-based CA screenings for which he approximated the covariances based on a GP error growth evaluation and formulated histograms that could be used to make informed choices of possible aspect ratio values. A similar profiling was conducted for the aspect ratios of the combined covariances for the current dataset, and the results are given in Fig. 7. Alfano's profiling showed 99% of conjunctions to have an aspect ratio of 40 or less and 99.9% a value of 70 or less, with 140 as essentially an upper bound. The current dataset, for which the black line represents all conjunctions, gives a very different set of results: a median aspect ratio of about 20, a 90th percentile of about 130, a 95th percentile of approximately 250, and an absolute maximum that exceeds 12,000. It is clear that the characteristics of a 2015-16 precision catalogue, which includes the results of two large debris-producing events, are very

Figure 7. Conjunction plane covariance ellipse aspect ratio CDFs for analysis dataset, separated by color. *Black line shows results for all data.*

different from those arising from his GP-based investigations with a 2005 catalogue. The increased dynamic range of the projected combined covariance aspect ratio present in the special perturbations (SP) catalogue makes it much more difficult to use a technique that requires an *a priori* guess at this value. Given the information vacuum that would drive the user to use a PcMax calculation approach in the first place, and given the increased difficulty in
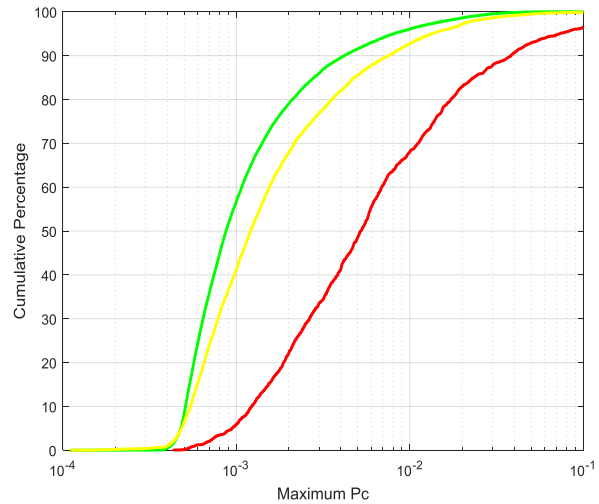
choosing a representative aspect ratio in the SP situation, it is unlikely that one could justify using a technique that requires an assumption of the aspect ratio value over the PcMax technique, which does not require this.

The more bounded PcMax technique developed by Frisbee—which addresses the case in which one of the object's covariances is known but the other is not— would seem to hold more promise as a screening threshold, as one should expect the Pc values to be not as large and therefore perhaps smaller than a typical "red" threshold. Figure 8 gives CDF results for the Frisbee technique against the analysis set of OCMs; the solid lines represent cases in which the covariance of the secondary object is presumed to be unknown and the dashed lines those cases in which the primary covariance is presumed unknown. The former situation (secondary covariances unknown) would be encountered by a satellite owner who posssesses a precision ephemeris and covariance for his own satellite but is perhaps performing CA against a GP catalogue of secondaries, which lacks a covariance. As can be seen in the figure, screening against a red threshold of 1E-04 would produce essentially no eliminations, and a more aggressive figure of 5E-04 would result in elimination of about 20% of the cases. Perhaps this amount of elimination would be seen as somewhat helpful; but given that this testset produced about 200 OCMs per day,

Figure 8. Constrained maximum Pc CDFs for analysis dataset, separated by color.

eliminating 40 still leaves 160 to process—a large enough number that methods for investigating large numbers of potentially serious conjunctions will still be needed. The situation in which the primary covariance is unknown, which can arise when using owner/operator ephemerides (that sometimes lack a covariance) to screen against a precision catalogue of secondaries that contains covariances, is represented by the dashed lines in Fig. 8. Here one can see that the situation is more sanguine; against a red threshold of 1E-04, about 20% of the green OCMs can be eliminated and over 60% of the yellow OCMs. The use of the higher 5E-04 threshold eliminates over 50% of the red and green OCMs and 90% of the yellow OCMs. Such numbers document a real utility for the technique. However, if the screening agency has access to an entire space catalogue with reliable covariance data, then presumably it also has access to the catalogue entry—with covariance—for the primary object. Simply using this covariance along with the owner/operator ephemeris, while introducing perhaps some theory incompatibility, would still be a much better estimate of the primary object's covariance than presuming that nothing at all is known and employing a PcMax technique. If there is concern about theory incompatibility, one could respond by making the screening threshold somewhat lower to compensate.
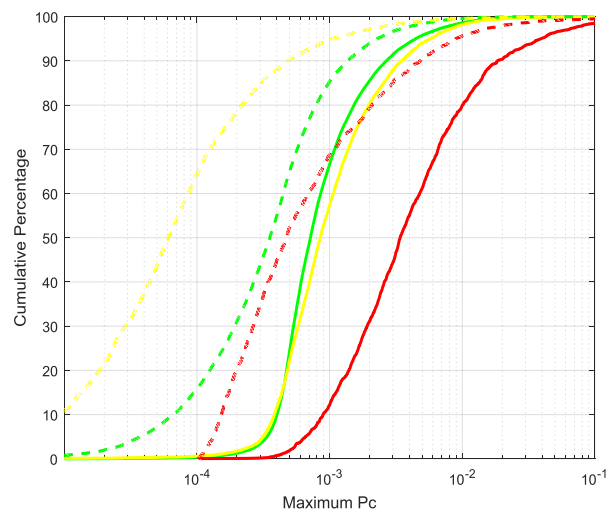
C. PcMax is not Practically Useful as a Screening Parameter for the Future

There is virtue in being able to dismiss an event for the present because the PcMax falls below the chosen threshold for considering a conjunction to be worrisome and therefore worth pursuing further. There is much greater virtue, however, in being able to dismiss the event for the rest of its multi-day history based on a PcMax calculation from a single OCM. If one could reasonably conclude from a single report that the event had no reasonable chance of ever becoming serious, this would eliminate the need for any further consideration or calculations for this event, even if it continued to appear in screenings.

In order for such decisive dismissals to be possible on the basis of a single OCM, the miss distance between the two objects at TCA must remain relatively constant from update to update. After all, the unconstrained PcMax formula is driven only by miss distance (actually, by the ratio of the HBR to the miss distance, but with a fixed HBR it is the same thing), so an enduring miss-distance estimate would give similar endurance to the PcMax formula's results. Frisbee's constrained PcMax formulae are more problematic candidates for enduring estimates; for essentially invariant enduring estimates for miss distance and covariance would be required, and propagated covariances are expected to be smaller in size with shorter propagation intervals. But to enable either case, a miss distance estimate that can be expected to change rather little over the remaining event history is necessary.

If a PcMax approach is being pursued because of questions regarding the propriety of the covariance, it could be further wondered whether the same issues that affect the realism of the covariance similarly affect the quality of the
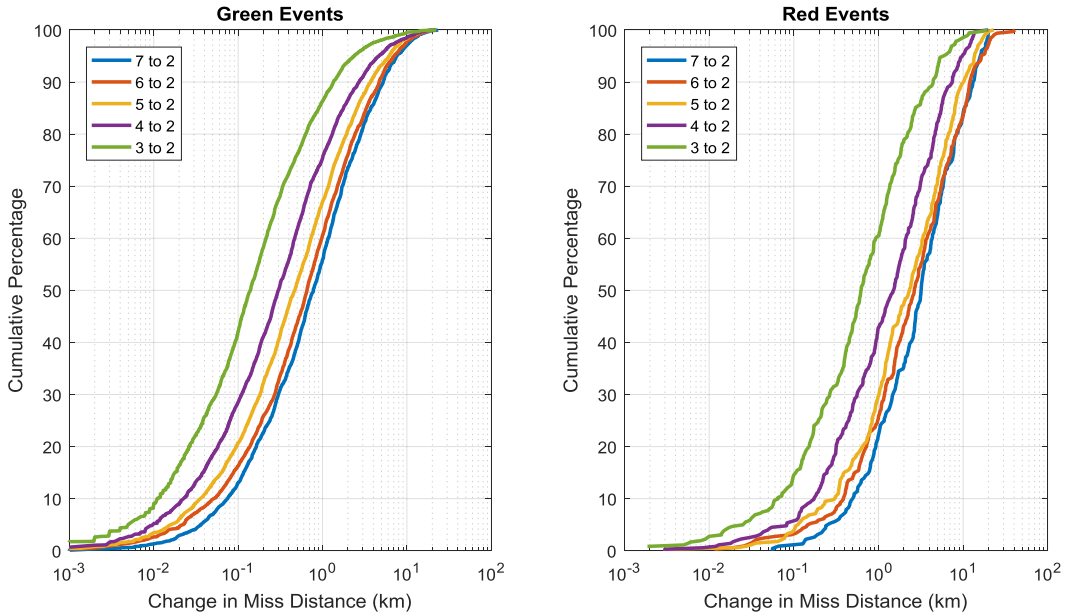
Figure 9. CDF plot of change in miss distance, by time to TCA. *Green and red events are plotted separately.*

state estimate and therefore the calculated miss distance, at least to enough of a degree to challenge the expectation of a fairly constant miss distance. In any case, Fig. 9 provides data on the change in relative miss distance as conjunction events develop. For each event (of which there were ~7300 in the dataset under analysis), the miss distance at two days to TCA was used as the standard against which changes were measured; this point was chosen as a typical maneuver-commit time and therefore the point at which conjunction analysis would essentially cease. The set of CDFs in the left graph summarizes the situation for green events and the set on the right for red events. For this chart, the color definitions are modified slightly: an OCM is assigned a color based on the highest Pc that the event achieved during its history; for example, an OCM with a Pc of 1E-10 would be given red status if at some point in the event history an OCM were received with a Pc greater than 1E-04. Each CDF line describes the set of miss distance changes over a certain number of days; for example, the blue lines summarize miss-distance changes from 7 days to 2 days to TCA. One can see that, even in some cases over short periods, there is a non-discountable number of rather large changes in miss distance, especially given that the screening volume itself (0.5km x 17 x 20) will allow a maximum miss distance of only ~26km, with observed values (from Fig. 4) tailing off at about 22km. Twenty-five percent of green cases change by more than 1km from 4 to 2 days, and 10% of red events change by more than 10km from 5 to 2 days. While these are not large percentages of the dataset, the fact that one out of ten red events experiences such a large miss-distance change is nonetheless a sufficiently large frequency to prevent the PcMax calculation to be used as an event dismissal criterion—such changes can impose too large a volatility on the Pc for the miss distances obtained using a screening volume of the present size.

It must be emphasized that the present study finds no theoretical problem at all with the PcMax formulae developed by Alfano and Frisbee; they quite correctly produce the envisioned outcomes, and the CA industry has been done a notable service by their development. Given the current available data products and CA practices, however, the added value of these approximations as screening filters appears to be much lower than was or would be encountered in a GP environment.

IV. "Forecasting" Epoch Covariances

A. Definitions and Basic Theory

The previous section has discussed manipulating the combined covariance in order to compensate for either a direct or selective absence of actual covariances for primary and secondary objects. In this section, a very different type of manipulation will be addressed: the attempted construction of a future epoch covariance in an attempt to

give better-contexted risk assessment information. Because the concept and techniques are somewhat more complicated, in order to allow clearer explanations it is important to establish some terminology.

An orbit determination (OD) activity produces a state estimate of a satellite at a particular point in time, and this time point is called the epoch time. There are conventions for setting OD epoch times, but the epoch can in fact be at any time chosen for convenience. The OD process also produces an epoch covariance, which represents the estimation uncertainty at this epoch time. If the covariance at some future time is desired, the epoch covariance is propagated forward in time; there are a number of different methods for doing this, ranging from a simple linearized propagation to a full Monte Carlo approach. This forward propagation of covariance (usually through simple linearized means) is a common practice in CA: a screening exercise finds a conjunction expected to take place several days in the future, so the objects' states and covariances are propagated forward to TCA and the Pc calculated. Therefore the typical situation is to perform a risk analysis some number of days before TCA, with the objects' states and covariances propagated to TCA—the objects' epoch times are some amount of time in the past, prior to the time the risk assessment is being performed, and the propagation interval for each runs from their epoch times to TCA.

Covariance Forecasting also attempts to construct a future covariance but pursues it in a different way; its objective is to try to determine what an object's *epoch* covariance will be at some future point. To understand how this is possible, it is necessary to understand how the batch estimation technique used at the JSpOC conducts covariance formation. Except in those relatively unusual instances where initial orbit determination is required, OD at the JSpOC consists of using sensor observations taken on the object to differentially correct the orbit in order to minimize the observation residuals. The minimization equation and variable definitions are given below:

$$dx = (A^T W A)^{-1} A^T W b \qquad (1)$$

- dx is the vector of corrections to the state estimate
- A is the so-called "normal" matrix of partial derivatives, and it consists of two conceptual parts. The first is a set of partial derivatives to map changes in observation space to changes in state space, and the second is to map state changes at a particular time to state changes at the chosen epoch time. A matrix of partial derivatives for each of these two transformations is developed (the observation-to-state space matrix is usually called the H-matrix and the time-modification matrix called the state transition matrix), and the product of these produces the A-matrix.
- W is the matrix of relative weights of the observables from different sensors. It is acknowledged that different sensors take observations with different levels of accuracies. Typically the weighting factors are the reciprocals of the variances of a sensor's expected errors in each observable, determined by a sensor calibration process.
- b is the vector of observation residuals against the trajectory given by the previous state estimate.

The term $(A^T W A)^{-1}$ is the covariance matrix; it gives the expected variances in each of the state variables and the covariances of these variables with each other.

Given this covariance production mechanism, it is important to note what information is and is not required in order to construct the covariance matrix. To determine observation-space-to-state-space partial derivatives, time transformations from observation time to epoch time, and relative weighting of observables, all that is required is to know the amount of expected tracking, the reporting sensor for each track, and the time at which the track is taken. Even though the covariance gives an estimate of the error in the state estimate produced by the OD, it is not a summary of the actual residual errors; rather, it is an *a priori* estimate of what those errors are expected to be given the amount of tracking, the sensors that produced it, and the times at which the tracks were taken. There is sometimes some disappointment about the more abstract nature of this covariance, but the corresponding advantage is that it can be constructed in the absence of actual tracking data. If one believes that it is possible to make durable predictions of the tracking that an object will receive and the times and sources of this tracking, then the covariance matrix can be assembled and even used based on these predictions only.

While the approach described above is the most rigorous for constructing a covariance (as it is the method actually used in the OD process), there are more simplified alternatives. One could presume that, most of the time, the epoch covariance for any given object is more or less the same; so one could simply take an epoch covariance from the past and deploy it as a proxy for a future epoch covariance. A second approach could be to characterize the "shrinkage" in the covariance propagated to TCA as one moves from solutions from 7 to 6 to 5 days to TCA and extrapolate from this trend what the propagated-to-TCA covariance will look like at, say, a decision point two days

to TCA. These simplified methods assume that the tracking levels, sources, and relative times are similar enough that such proxies or trends are durable and predictive with relatively small errors; this issue will be addressed in more depth in a subsequent section.

B. Forecasting Covariances to TCA

The above outlines how one may go about generating a future epoch covariance (or, using simplifying methods, a propagated-to-TCA covariance from a different temporal starting point) without the possession of actual tracking data. Given that this is possible, what should one do with this capability? One possibility that comes to mind immediately is to forecast the primary and secondary covariances to TCA. This certainly seems *prima facie* to be a reasonable application: propagating the covariance to TCA increases the covariance size because of the growing effects of state estimate uncertainties once projected into the future, so avoiding the need to propagate by forecasting the epoch covariance at TCA should give the most reasonable and bounded estimate of the covariance and thus yield a better risk assessment result.

While this approach is at many levels attractive, it encounters some philosophical difficulties that leave the forecasting of the epoch covariance to TCA a problematic concept. The issues are best explained by stepping through the different possibilities of the temporal placement of the epoch covariance, either real or projected, in the typical life cycle of a conjunction event.

Conjunction risk analysis is performed by taking information actually in possession of the evaluators and projecting it into the future to illuminate the circumstances of the expected future conjunction. State estimates and covariances for the primary and secondary object, which have an epoch time some period of time in the recent past, are propagated forward to TCA; and the Pc is calculated, along with other information of interest such as collision approach angle, combined covariance orientation in the conjunction plane, etc. Quantified uncertainty in the satellites' state estimates is captured in the epoch covariances, and these uncertainties grow as they are propagated forward to TCA, often presenting the analyst with more uncertainty about the situation than he or she would like. But everything that is calculated about the conjunction at TCA is grounded in information that is actually known; while state estimates and covariances are propagated, they are based on OD results using real observations from real sensors.

One possible application of covariance forecasting is to prognosticate an epoch covariance at a future decision point. For example, suppose a conjunction has been discovered through a screening activity and is 7 days from TCA. A typical decision point for conjunction remediation is 2 days from TCA; this is close to the "maneuver commitment" point at which the maneuver commands are authorized and will be uploaded to the spacecraft, and it may not be possible to call off the maneuver past this point. One could use covariance forecasting to try to construct an epoch covariance at the 2-days-from-TCA point; by propagating this forecasted 2-day epoch covariance to TCA, one can construct what the decision-maker is likely to be faced with when he or she actually has to render a remediation decision at the 2-day point. The decision will not, of course, be made 7 days from TCA even if a 2-day covariance forecast has been executed and is in hand, but what can be done is to determine how the availability of other information might improve or erode the decision-maker's trade-space at the 2-day point. For example, one could construct several different epoch covariances for the 2-day point, each of which is assembled presuming different levels of sensor tracking, to determine whether an increase in tracking would substantially facilitate the remediation decision at the 2-day point. Some conjunctions show substantial reduction in the Pc given additional tracking for one of the objects (usually by shrinking the size of the secondary object's covariance); other conjunctions are relatively insensitive to tracking increases. If, based on a forecast covariance, it appears that increased tracking could considerably improve or clarify the decision situation at the 2-day point, then one would have a reasonable justification for asking for increased sensor tasking on the object. This covariance forecast approach provides a more rigorous method for determining to which conjunctions relatively scare sensor tracking resources should be allocated—it is superior to examining simply recent tracking levels or the epoch age of the OD, for there is no necessary connection between an object's appearing to be in need of tracking and additional tracking actually aiding an associated conjunction remediation decision, and vice versa.

While one can indeed attempt to predict what the risk assessment parameters will look like at some future decision point (here by constructing epoch covariances at the 2-day-to-TCA point, propagating them to TCA, and calculating risk assessment parameters from these data), it is not proper to make the actual remediation decision days in advance of the 2-day point even with such information in hand. First of all, the entire utility of the covariance forecasting enterprise rests upon the presumption that the conjunction geometry and miss distance will not change from the current estimate—only the covariances are prognosticated, not the actual states; so that part of the forecasting activity is taken from the 7-day propagation and left unaltered. Subsequent tracking is likely to change the miss distance and geometry at least to some degree, and potentially quite significantly (as was shown

American Institute of Aeronautics and Astronautics

previously in Fig. 4); so it would be extremely ill-advised to render an actual remediation decision in the face of a presumption that in many cases does not inhere well at all. Second, the actual amount of tracking may be quite different from what was predicted, including potentially the receipt of no new tracking at all. This latter case throws the precariousness of the forecast covariance nicely into relief. If no additional tracking is received between the discovery of the conjunction and first analysis (7 days to TCA) and the decision point (2 days to TCA), then the reasonable practice is to make the remediation decision at the 2-day point using the data that were in fact available at the 7-day point—these data represent the most recent information, stale though they may be. It is difficult to contemplate at the 2-day point using instead results from a forecast epoch covariance for the 2-day point— assembled five days previously—based on expected tracking that never in fact materialized. The forecast covariance is useful as a direct decision aid only to the degree that its guesses at the expected tracking levels actually align with the observed tracking levels. Otherwise, if a forecast epoch covariance analysis shows that more tracking would drive the risk below an action threshold, why even bother to ask for the additional tracking? One could simply dismiss the event based on the forecast covariance, regardless of whether the sensors ever followed through with the requests. A forecast covariance can reveal whether additional data would or would not be helpful, but it cannot serve as a substitute for the actual data. If this were not true, why not postulate an infinite amount of extremely accurate data, which would drive the risk down to 0 every time, so long as the nominal miss distance is greater than the HBR, which it almost always is?

With this background discussion in place, it is time to address the forecasting of the epoch covariance to TCA. TCA, as the moment at which a collision will occur (in the hyperkinetic collision case at least, which appropriately describes most conjunctions), is not a remediation decision point; and no assembly of the decision aids as one believes they might appear at TCA would ever have the benefit of the "actual" data that these constructions suppose will be received. So a decision based on anticipated constructions of the epoch covariances at TCA can never be a decision grounded in real data.

This fact, in the end, renders the forecast covariance at TCA without any real risk assessment utility. Forecasting covariances at TCA is presently performed by guessing at the amount of tracking that will be taken on the primary and secondary objects and constructing a pair of covariances from this information, in one of the manners described in the previous section. However, because there is no intention to tie these predictions to the actual receipt of such tracking data, there is no reason to limit the hypothesized tracking to what the space surveillance network (SSN) is actually likely to do or, for that matter, even to remain within its present capabilities. Once the point for actually making a decision (*e.g.,* the 2-day point) has passed, why is it important anymore what subsequent tracking data are actually taken? Tracking data after the decision point cannot change the assessment of the collision likelihood in any useful way—the decision has already been rendered, and such data were not available to participate in the decision. Suppose that after the remediation decision point all of the SSN sensors devoted all of their tracking energy to both objects whenever they were in view, and in light of all of this tracking, the calculated Pc dropped substantially; while that is interesting information, it is not helpful because the decision point has already passed. Now suppose that fictitious sensors are added to the SSN and turned on right after the 2-day decision point, and they produce so much excellent tracking that the conjunction risk now drops essentially to zero. Finally presume that no tracking is expected on the objects at all after the 2-day point and thus the Pc value calculated using the forecast covariances is exactly the same as the Pc at the 2-day point. All four of these outcomes (regular tracking, high-volume SSN tracking, fictitious almost infinite tracking, and no tracking) in this context render information that, at the time the remediation decision is to be made, cannot be verified or supported by actual tracking levels. Therefore, and this is the important point, there is in the end no substantive difference among any of these four options, even though some of them are pure fantasy. The default position has been, of course, to try to guess what tracking the SSN would actually perform between the remediation decision point and TCA, but why is this inherently a more useful position to hold than, say, the absurd position of presuming that the SSN would between now and TCA dedicate itself entirely to tracking the two objects in question? In both cases one is guessing at how the perception of the risk might change if that amount of tracking were obtained, but in neither case will any of the actual tracking data make it back to the decision-maker as actionable information. The understanding of the likelihood of a collision is affected by the actual amount of information the decision-maker has at the decision point; it is not affected at all by the amount of fictional information that one hypothesizes could be taken after the remediation decision is made. The statement "if the SSN tracks as it usually does, an OD were performed at TCA, and the conjunction miss distance and geometry do not change from right now until then, the Pc would be X" is no more or less useful than the statement "if we were to get an infinite amount of error-free tracking between now and TCA, an OD were performed at TCA, and the conjunction miss distance and geometry do not change from right now until then, the Pc would be 0."

A different way to look at the situation is the following. Seven days from TCA, a screening activity calculates a miss distance and a Pc, but there is not much certainty about the situation because the tracking is less than desirable and seven days of propagation to TCA is required. Forecast covariances for the primary and secondary at TCA are assembled, presuming increased tracking (because increased sensor tasking would be requested). A Pc is calculated from these forecast covariances (but the unchanged current miss distance), and the Pc is much lower because the covariances are better determined (more tracking) and no covariance propagation is required since they are determined for TCA. One now brings this new Pc to the risk analysis activity and concluded that this event's Pc is likely to decrease.

Why is this same miss distance, which 7 days prior to TCA is very uncertain given the large covariances, suddenly very certain in the forecast at TCA when no new data have actually been brought to the evaluation? Why would it be presumed that all of this hypothesized tracking would not result in a different miss distance? If the situation is so uncertain when this (unchanged) miss distance is first calculated, should that not almost guarantee that the miss distance will be different should all of the hypothesized tracking actually get taken and incorporated into an OD? Uncertain situations cannot be made more certain through the postulation and incorporation of fictitious data.

C. Covariance Forecasting Uncertainties

While there may be disagreement regarding the rectitude of using an epoch covariance forecast to TCA, there clearly are uses for covariances forecast to time-points up to and including the final decision point. The question now becomes how one may go about determining these forecast epoch covariances and the amount of uncertainty latent in these methods.

The NASA CARA project presently has an open task to create a large number of synthesized covariances as part of a study on the expected effects on the CA mission of the US Air Force's S-Band Fence radar deployment; after the completion of that study, a large number of numerical examples will be available to help determine the ability to predict actual tracking levels and therefore the actual differences between hypothesized covariances and the real covariances that eventually were formed by the OD process at those same time points. For the present, suffice it to offer some general comments on what must be done in order for a forecast covariance to be a reasonable contributor to an assessment of risk.

The question of the best way to produce such a covariance appears to remain a problem without a definitive solution. Duncan[8] addresses it at only the highest level, stating that "Forecast covariance can come from time history files or propagated VCM covariances"; presumably the former refers to empirical studies that determine a canonical covariance growth history. Vincent[9] suggests that quadratic error growth curves could be used to "shrink" covariances appropriately but comments on the sensitivity of covariance formation and ultimately Pc to the observation history available for and used in the OD. Cerven[10] conducts his study within a GP framework and therefore turns to GP error growth curves to synthesize a covariance as appropriate. None of these three practitioners appears to be attempting to construct covariances from the actual predicted tracking.

A good reason for this hesitancy can perhaps be found in examining tracking level histories for secondary objects, as the secondaries tend to have the much larger covariances that thus dominate the Pc calculation. The tracking histories for the ~2100 secondaries involved in the ~7300 events in the present analysis dataset were obtained and analyzed. The one-year span of this dataset was divided into one-day, two-day, three-day, etc. up to seven-day time windows, and the number of tracks each object received in each of those windows was tabulated; descriptive statistics were then generated by object for this dataset, with the goal of trying to determine the amount of variability in the tracking data. The best statistic to represent this is the coefficient of variation, which is the quotient of the sample standard deviation and the sample mean of the dataset. Figure 10 gives the results of this analysis.
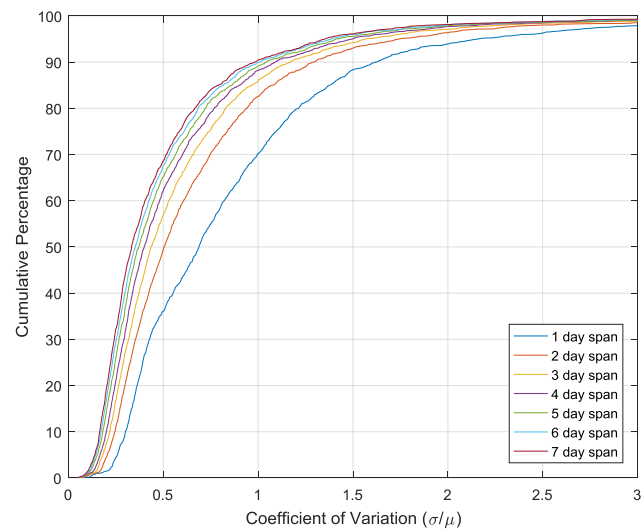

Figure 10. **Coefficient of variation of secondary tracking (σ and μ for # of tracks received in given time span).**

To be sure it is clear how the descriptive statistics were calculated, an example will be explained here in prose. The one-year period was divided up into five-day time windows. The number of tracks that each secondary received in each five-day time window was tabulated and stored—73 (365/5) tracking counts for each object. The mean and standard deviation of these 73 counts for each object was calculated, producing one standard deviation and mean per secondary object. The quotient of these two values was taken to produce one coefficient of variation per secondary object. A CDF curve of these coefficients of variation was then produced. This procedure was repeated for the other time-bin sizes; and the results for all seven time-bin sizes are provided in Fig. 10.

A coefficient of variation of unity means that the standard deviation and the mean value are the same, so the 1-sigma span runs from zero to twice the mean value. A distribution with these characteristics is unlikely to be Gaussian and is almost certainly right-skewed, so it does not make sense to apply Gaussian sigma percentiles to the situation; but what is clear is that the variation is quite large—the mean value is hardly a good proxy for the number of tracks that any given time window might contain. Coefficients of variation of 0.5 and larger would in this context be at the beginnings of a large value, and the CDF plots indicate that for a seven-day prediction (the most stable of the seven bins), 30% of the secondaries exceed this value of 0.5. Shorter predictions show even more volatility because there is less damping of smaller variations by the larger window size. So for a significant number of the secondaries, it will be difficult to claim that the amount of expected tracking can be predicted merely on the basis of past tracking rates; the amount of variation in these rates is too large to make this a credible claim. Furthermore, in order to construct the covariance, one must further specify the reporting sensors and the times of tracking; if it is difficult to anticipate simply the amount of expected tracking, it will be that much more difficult to predict the precise stations and times of such tracks.

This situation serves as an excellent illuminator of the second concept from Kaplan's risk assessment framework described earlier: estimates of likelihood contain uncertainty, and this uncertainty needs to be represented in the risk analysis trade-space. If one wishes to try to forecast covariances—and useful situations for doing this have been outlined previously—it will be necessary to give some statement of the uncertainty associated with the forecast covariance. The variation in tracking rates shows that tracking estimation errors are unlikely to be either negligible or static, so either an accompanying error function will need to be developed and flowed through the calculation or a Monte Carlo approach will need to be deployed that considers the actual likelihood of tracking by any given station during any given pass, considering tasking levels and previous tracking at that station. Only such an explicit recognition of the uncertainties inherent in this difficult estimation process can enable an estimate of collision likelihood that is sufficiently nuanced to be useful in the risk analysis process.

V. Conclusions and Future Work

Foundational principles from the Kaplan risk analysis treatment are helpful in evaluating two different types of covariance manipulation approaches. Maximum Pc calculation techniques do not produce an actual likelihood estimate and therefore cannot be included in a Kaplan Triplet as part of risk assessment. They are legitimately deployed as screening thresholds, but with the high-precision SP catalogue and typical CA screening volumes they have only very limited utility. Covariance forecasting techniques certainly do have a reasonable technical basis, but deploying them to forecast epoch covariances to TCA introduces philosophical difficulties that are not easy to resolve and, in the opinion of this author, ultimately leave them ill-suited for CA risk assessment. The forecasting of covariances to actual future decision points avoids these philosophical problems; but estimation uncertainties arising from the difficulties of predicting the sources, amounts, and times of future tracking are large enough that they simply must be accounted for explicitly in any covariance forecasting methodology. This latter requirement is in alignment with Kaplan's dictum that statements of event likelihood are never single point-estimates but must be given as probability densities in order to represent the full range of possibilities given the uncertainties in the inputs.

A planned follow-on to this study will consider the additional covariance manipulation approach of scaling known covariances within bounded ranges to determine a maximum Pc value (a common practice to address relatively minor but known covariance realism deficiencies) and attempt to provide actual uncertainty methodologies and values to the covariance forecasting activity. The hope is to provide a single, collected set of study results that examine all of the CA covariance manipulation approaches in the context of the Kaplan risk analysis foundational principles.

References

[1]Kaplan, S. and Garrick, B. "On the Quantitative Definition of Risk." *Risk Analysis*, Vol. 1 No. 1, pp. 11-27.
[2]Hubbard, D.W. *The Failure of Risk Management*. New York: John Wiley & Sons, Inc., 2009.

[3]Hejduk, M.D. and Johnson, L.C. "Approaches to Evaluating Probability of Collision Uncertainty." AAS Space Flight Mechanics Meeting (paper 16-241), Napa CA, February 2016.

[4]Chan, F.C. *Spacecraft Collision Probability*. El Segundo, CA: The Aerospace Press, 2008.

[5]Alfriend, K. *et al.* "Probability of Collision Error Analysis." *Space Debris*, Vol.1 No. 1 (1999), pp. 21-35.

[6]Alfano, S. "Relating Position Uncertainty to Maximum Conjunction Probability." *Journal of the Astronautical Sciences*, Vol. 53 No. 2 (April-June 2005), pp. 193-205.

[7]Frisbee, J.H. "An Upper-Bound on High-Speed Satellite Collision Probability when only one Object has Position Uncertainty Information." AAS/AIAA Astrodynamics Specialist Conference (paper # 15-717), Vail CO, August 2015

[8]Duncan, M., Wysack, J. and Frisbee, J. "Collision Probability Forecasting using a Monte Carlo Simulation." AMOS Technical Conference, Wailea HI, September 2014.

[9]Vincent, M.A. "Bridging the Gap between Academia and Operations for Orbital Debris Risk Mitigation." AMOS Technical Conference, Wailea HI, September 2015.

[10]Cerven, W.T. "Bounding Collision Probability Updates." AAS/AIAA Astrodynamics Specialist Conference (paper # 15-571), Vail CO, August 2015.