AGU100 ADVANCING EARTH AND SPACE SCIENCE

Space Weather

RESEARCH ARTICLE
10.1029/2017SW001720

Special Section:
Low Earth Orbit Satellite Drag: Science and Operational Impact

The Effect of Neutral Density Estimation Errors on Satellite Conjunction Serious Event Rates

M. D. Hejduk[1] and D. E. Snow[2]

[1]Astrorum Consulting LLC, in Support of NASA Goddard Space Flight Center, Waco, TX, USA, [2]Divide, CO, USA

Abstract While past studies have investigated the effect of neutral atmospheric density mismodeling on satellite conjunction assessment (CA), none has focused their investigation specifically on serious (high-risk) conjunction events, which are the event types that drive both risk and workload for CA operations. The present study seeks to do this by reprocessing chosen groups of archived actual conjunction events, artificially introducing atmospheric density error to these events, and then examining the effect of these introduced errors on the probability of collision (Pc) calculation, which is the principal parameter used to assess collision risk. These reprocessed calculations are executed both with the satellites' covariances unaltered and with a covariance modification that accounts for the induced atmospheric density error. The results indicate that the situation is greatly aided by an a priori knowledge of the approximate density estimation error, even if the model itself is unaltered—missed detections due to density estimation uncertainty are notably reduced when the density model prediction error is characterized and can be included in the satellite covariance and thus Pc calculation. Overall improvements in density model predictive performance, in situations of both low and high solar activity, substantially benefit the CA enterprise, especially for false alarm reduction; but model enhancements that include a robust, in-model error analysis offer the most significant improvements overall.

Plain Language Summary Space debris growth may render the near-Earth space environment unusable, and the best way to contain this growth is to predict potential collisions between satellites and then maneuver satellites in order to avoid situations of serious risk. The greatest source of error in satellite collision event prediction is the estimate of the atmosphere's drag effect (essentially "air resistance") on the satellite, and this error stems from difficulties in estimating the often-rapidly changing atmospheric density in space. The present study shows how even relatively small errors in this density estimate can lead to the missed detections of serious satellite collision risks. But if the errors in the models used to estimate atmospheric density are known, this uncertainty can be included in the "probability of collision" estimate between two satellites and the collision risk much more correctly predicted. It is thus important not just to improve the accuracy of the models themselves but also to develop accompanying algorithms that can statistically estimate their prediction error.

Key Points:
- Atmospheric modeling effects on satellite conjunction assessment (CA) are best assessed by the effect on high-risk conjunction event rates
- Including density model errors in conjunction risk calculations greatly improves the stability and durability of the results
- Density model accuracy improvements will improve CA notably, but even more strongly if model errors are characterized

Correspondence to:
M. D. Hejduk,
mdhejduk@astrorum.us

Citation:
Hejduk, M. D., & Snow, D. E. (2018). The effect of neutral density estimation errors on satellite conjunction serious event rates. *Space Weather*, *16*, 849–869. https://doi.org/10.1029/2017SW001720

Received 29 AUG 2017
Accepted 5 JUN 2018
Accepted article online 19 JUN 2018
Published online 25 JUL 2018

1. Introduction

Conjunction assessment (CA), or the determination of the possibility and likelihood of collision between space objects, has become an area of greatly increased interest and research within space situational awareness. The increase in satellite population, the deployment of sensors that can track smaller and smaller objects, and U.S. Strategic Command's more expansive data-release policies to commercial and foreign entities have all combined to create a significantly expanded CA solution space: both an increased number of known conjunctions between space objects and the distribution of relevant conjunction data to the concerned entities so that further analysis and potential conjunction remediation can be performed. Considerable academic research has accompanied this increased operational interest and been directed to all aspects of the problem, from the basic astrodynamics of identifying conjunctions, to the determination of collision likelihood, to efficient methods for calculating conjunction remediation options. In analyzing the entire chain of CA-related calculations for potential sources of error, drag acceleration error suggests itself as the single greatest contributor to satellite state propagation uncertainty (and thus CA calculation error) for satellites with altitudes less than about 1,000 km, and one of the principal ingredients to the drag acceleration calculation is neutral atmospheric density estimation. While substantial literature exists on comparative performance among different atmospheric density models when presented with different space weather

phenomena, two studies focused specifically on the relationship between density mismodeling and resultant satellite position errors (Emmert et al., 2017; Vallado & Finkleman, 2014) have shown that mismodeling effects can be substantial—potentially large enough to influence CA-related operational conclusions. Emmert et al. (2014) has worked out some bounding calculations to try to quantify this effect, using an idealized CA framework.

The purpose of the present work is to enhance the efforts referenced above by quantifying this effect within the CA calculation, risk assessment, and decision support methodologies that follow the best practices of the leading CA institutions. The National Aeronautics and Space Administration (NASA) Robotic Conjunction Assessment Risk Analysis (CARA) project is the largest single governmental organization to conduct routine CA operations. With approximately 70 protected primary spacecraft spanning a wide variety of orbit regimes and a full-time research staff, CARA has been active nationally and internationally in developing, operationally deploying, and evaluating CA best practices. It will be against these practices that atmospheric density modeling error will be examined to determine its effect on the number of serious conjunctions and therefore actual imputed CA workload. The result of such an investigation can be used to construct a more definitive statement of the operational impact of atmospheric density error and therefore make evident the operational benefits of improved atmospheric modeling.

This study thus follows the following organizational schema. First, basics of the CA enterprise will be discussed in order to establish the proper framework for subsequent discussions. Next, the drag acceleration equation will be examined and the uncertainties presented by each term discussed so that atmospheric density error can be properly substantiated as the principal contributor to drag acceleration error. After this, the relationship between errors in atmospheric drag acceleration and the calculation of the probability of collision (Pc), the basic parameter for performing conjunction risk assessment, are delineated in order to establish the link between atmospheric density estimation errors and changes in the assessed risk of a particular conjunction and thus formulate the study's first key question: how strongly are Pc values affected by atmospheric density modeling errors? Finally, methods for attempting to compensate for these density estimation errors within a satellite's state covariance matrix are discussed, and with this the investigation's second key question is defined: by how much are CA risk assessment errors that arise from atmospheric density estimation errors attenuated by properly adjusting the satellite's covariance matrix to account for these errors? With these background items presented, one is then ready to describe the data set and conduct of the experiment itself and its results.

2. The CA Enterprise and Associated Processes

The CA enterprise is typically divided into three parts. While some of these divisions are not intrinsic to the calculations themselves, they are nonetheless both conceptual and practical divisions guiding the way that conjunction-related data are presently generated and distributed to the agencies responsible for risk assessment. They thus serve as useful points of demarcation of the process.

2.1. CA Screenings

The purpose of CA screenings is to discover potential conjunctions between space objects some time in advance of the time of closest approach (TCA) so that risk assessment activities can be performed, orbit determination (OD) refinements can be executed, and, if necessary, conjunction remediation actions taken. A screening is an evaluation of a single protected asset's future positions in comparison to the future positions of all other objects in the space catalogue. Typically performed for a look-ahead period of at least seven days, the ephemeris of the protected object (called the "primary") is compared to ephemerides for all the other catalogued objects (called "secondaries"), and any secondaries that come within a specified (componentized) distance of the primary are identified as possible conjunctions, which can then be sent on to the risk assessment portion of the process. To be specific, a particular volumetric region is constructed about the primary object and propagated along its trajectory; any penetration of that volume by a secondary object constitutes a conjunction. These screening volumes are rarely spherical but usually ellipsoidal in order to align most strongly with the expected satellite state error distribution in the particular orbit regime. The current screening volume sizes used operationally were derived from analyses based on histories of satellites' state error covariance matrices (Narvet et al., 2011), although a reprisal based on the data mining of screening test data using very large screening volumes is presently underway, with the first stage of the study completed

(Hejduk & Pachura, 2017). A number of efficient filtering mechanisms have been developed in order to decrease the computational burden of this process (e.g., Alfano, 2013; George & Chan, 2012; Hoots et al., 1984).

2.2. CA Risk Assessment

The identification of a conjunction between a protected primary and a secondary object is not a datum without utility, but from a satellite protection point of view, it alone is not actionable information. In an earlier period of CA operations, risk assessments were attempted based solely on the predicted closest miss distance between the primary and secondary. However, as the discipline matured, this methodology was found to be less than fully adequate because it did not consider the state uncertainties between the two objects. If the state uncertainties are large, a small miss distance usually does not indicate a high likelihood of a collision because the actual satellites' positions could well be far from the mean values. Similarly, a miss distance that does not seem particularly small, if it aligns properly with the actual state uncertainty values, can produce a situation in which the collision likelihood is larger than one might expect. It thus became clear that a collision likelihood, or probability of collision, calculation should be developed in order to provide a statement of actual collision risk. The initial assembly of a Pc calculation methodology was performed for the Space Shuttle program (Foster & Estes, 1992), and since that time a number of calculation approaches have been developed and are capably summarized by Chan (2008). The availability of the Department of Defense (DoD) precision catalogue and accompanying covariance data has allowed these probabilistic calculations to be performed by nearly all CA practitioners. A full risk assessment evaluation includes not only the Pc but also an assessment of the adequacy of the primary and secondary OD and prognostications of what is likely to take place regarding additional secondary tracking and therefore at what particular time a final risk assessment calculation might be best made (Newman et al., 2014). A recent expansion has considered the inherent uncertainty of the Pc itself by profiling and using expected errors in the two objects' covariances and uncertainties in the hard-body radius (HBR; Alfano & Oltrogge, 2016; Hejduk & Johnson, 2016). The covariance uncertainty portion, which examines the correspondence between predicted covariances and actual state errors and from this generates probability density functions of possible covariance realizations and thus probability density functions of possible Pc values, is a global uncertainty correction that will also implicitly account for neutral density prediction errors (Newman et al., 2016). It is, however, an imprecise correction method that should not be considered a replacement for proper modeling and cognizance of atmospheric density model errors.

2.3. CA Serious Event Remediation

If the collision risk is considered high, usually because the Pc value exceeds a particular threshold and the OD results appear credible, the focus of the process turns to conjunction remediation. This is typically accomplished through the execution of a specially assembled satellite maneuver, called a Risk Mitigation Maneuver, to change the primary satellite's trajectory in order to reduce the risk of the conjunction to an acceptable level. Sometimes a similar outcome can be achieved by changing the time and/or intensity of a maneuver already scheduled for a different purpose (such as a drag makeup maneuver) in a manner that both accomplishes the maneuver's original objectives and also mitigates the conjunction risk. The planning process typically is a nested analysis of the resultant Pc for the main conjunction versus potential instantaneous burn intensity and burn time. The basic trade-off is that burns conducted earlier (i.e., longer before TCA) can be smaller yet produce the same remediative value, but waiting longer increases the likelihood that additional satellite tracking will refine the orbits such that the collision risk drops considerably, thus obviating the need for a maneuver at all. Maneuver timing depends on mission maneuver planning capability: some older missions with a large manual component to this planning need to make an execution decision almost three days before TCA; more modern missions with automated and streamlined procedures can wait until ~12 hr before TCA before making a final decision. About 10% of all conjunction events that reach the level of detailed maneuver planning actually result in executed maneuvers; reasons for failing to perform a remediation maneuver include the risk falling off to a tolerable level through natural event development (the most common), the required maneuver being so large as to propel the satellite out of its control box and/or jeopardize its mission performance too strongly, or the OD quality for the secondary object being so poor as not to serve as a basis for conjunction remediation. Satellites that use longer burns, electric propulsion, or

other approaches for trajectory control such as differential drag, require modified remediation methods, but the overall concept and procedural outline is identical.

The focus of the present analysis is on the risk assessment stage of the process, that is, to determine the effect of atmospheric density mismodeling on the number of CA serious events, for it is these events that drive nearly all of the CA analytical and communications workload. The previously referenced works by Emmert et al., in which he develops analytical expressions that link density mismodeling error and satellite position error and in which he calculates expected differences in the number of conjunctions found by screening processes, are both excellent studies, but they do not formulate conclusions within the framework of current CA risk assessment terminology and practices and therefore cannot be immediately employed to assess the expected effect of mismodeling on CA operations. To examine this issue so that operational conclusions can be drawn, two activities are necessary: first, actual conjunctions with typically encountered conjunction geometries and Pc calculations must be analyzed, and second, the effects of density mismodeling on state estimate covariances, including common approaches to try to account for such modeling errors, must be included. The next three sections provide a more extended treatment of the relationship of density mismodeling to Pc calculations and discuss atmospheric error compensation strategies often applied to state covariances.

3. Drag Acceleration Sources of Uncertainty

Using Montenbruck and Gill (2005) notation, the satellite acceleration due to drag is given as

$$\ddot{r} = -\frac{1}{2} C_D \frac{A}{M} \rho v_r^2 e_v, \tag{1}$$

in which r-double-dot is the antivelocity acceleration, C_D is the drag coefficient (dimensionless), A is the spacecraft frontal area (normal to the velocity vector), M is the spacecraft mass, ρ is the atmospheric density, v_r is the magnitude of the velocity relative to the atmosphere, and e_v is the unit vector in the direction of the spacecraft velocity. Because the spacecraft mass and frontal area are usually not known independently, the group quantity of C_DA/M is solved for as a unit as part of the routine OD update and called the ballistic coefficient. Uncertainty exists in all of the components that combine to produce the drag acceleration; the following discussion examines each of these components to characterize its imputed uncertainty and demonstrate that in the great majority of cases, it is the atmospheric density error that predominates. Because the ballistic coefficient is solved for as an ensemble parameter with each OD, what is important to establish is that its constituent elements are either largely to entirely invariant over the OD prediction interval (rarely more than 10 days) or if they do vary significantly during a time interval shorter than this, this variation is minor compared to that of the atmospheric density error.

3.1. Satellite/Atmosphere Velocity Terms ($v_r^2 e_v$)

A satellite's inertial velocity is a quite well-determined parameter: in examining all of the CARA CA data for secondary objects (with perigee heights less than 1,000 km) for 2017, in fewer than 2% of the cases did the 1-sigma velocity uncertainty exceed 1% of the velocity vector magnitude. However, the velocity parameter of interest is not its inertial value but the magnitude of the velocity relative to the atmosphere. The usual assumption is that the atmosphere corotates with the Earth, and under this assumption, the calculation of this relative velocity is straightforward. The potential presence of upper-atmospheric winds unfortunately can complicate the situation. Under quiescent conditions, such winds are in the 150 m/s range and therefore much smaller ($<$2%) of a typical satellite velocity of 8,000 m/s, but during geomagnetic storms, such winds can increase substantially, to measured values of 650 m/s at 200 km (Zhang & Shepherd, 2000) and to modeled values of up to 900 m/s at 300 km (Wang et al., 2008), and such winds may be even stronger at higher altitudes. In the midst of such geomagnetic activity, this source of uncertainty would need to be considered. As it is, such situations are both relatively infrequent and present large uncertainties in the more fundamental atmospheric modeling. Until predictive atmospheric modeling for solar storm perturbations becomes much more sophisticated, it is probably acceptable for the purposes of CA to attempt to account for upper-atmospheric wind uncertainties in the general accounting for atmospheric density uncertainty during solar storms.

3.2. Satellite Mass (*M*)

For active payloads with propulsion systems, satellite mass obviously changes with time as propellant is expended for orbit maintenance burns. Fortunately, within the CA enterprise, such objects are nearly always the primary satellite, for which an owner-operator-produced predicted ephemeris is available; this calculated ephemeris takes into account changes in satellite mass due to propulsive burns. In the absence of a priori information, the mass values of secondary objects (typically space debris) are usually difficult to estimate; recent attempts employing solved-for ballistic coefficients and estimates of the satellite C_D and frontal area from signature data (radar cross section or satellite visual magnitude) do not expect results better than an order of magnitude (Hejduk et al., 2017). With such techniques, it is therefore not possible to obtain enough fidelity of estimation to detect relatively minor but astrodynamically significant changes in secondary object mass. Fortunately, with perhaps the exception of a few satellite classes that have been known to shed or leak contents over time (of which the best known is perhaps the NaK coolant spheres leaked from retired RORSAT reactor cores; Wiedemann et al., 2005) and satellites in the final states of decay, there is no proposed physical mechanism to alter an inactive satellite's mass substantially over short time intervals (except perhaps an undetected collision!). As such, it is reasonable to consider satellite mass as an invariant parameter over short periods.

3.3. Satellite Drag Coefficient (C_D)

Early theoretical work on drag coefficient computation, including both closed-form solutions and iterative models with at least some comparison to empirical measurements, was accomplished by Schamberg (1959) and Cook (1965). A reasonably full picture of the phenomenology, along with more expansive in situ measurements and consideration of empirical issues, was set out in a series of articles by M. M. Moe and Wallace (1993, 1995), K. Moe and Wallace (1998), and K. Moe and Bowman (2005). It is this body of work that informs the following explication and sensitivity analysis, as most research efforts subsequent to this have tended to focus on more elaborate C_D modeling techniques for known, complex objects (e.g., Mehta et al., 2013, 2014; Pilinski & Palo, 2011).

As the satellite passes through the atmosphere along its trajectory, if there were no random thermal motion and all of the encountered molecules stuck to the incident surface on the satellite, without reemission, then by Newton's second law, the drag coefficient would take on a value of 2. If one now allows reemission but no random thermal motion, then the drag will increase to a degree governed by three parameters of the satellite-atmosphere encounter:

1. The accommodation coefficient, which is the ratio of the energy difference between incident and reemitted molecules ($E_i - E_r$) and the energy difference that would be observed if the reemitted molecule had merely the energy content governed by the temperature of the satellite surface (or "wall," $E_i - E_w$). A movement of the accommodation coefficient closer to unity usually indicates a transfer of the energy into other radiometric phenomena (such as satellite "glow") rather than more energetic redirection of incident molecules and as a result tends to reduce the satellite drag and thus the drag coefficient.
2. Diffuse versus specular reemission. Diffuse reemission requires that the incident molecules be appropriated to some degree by the encountered surface, and thus, any reemission tends to occur at a random angle from that surface. Specular reflection, on the other hand, is mirror-like, in which the incident molecules are reemitted at an angle equivalent to the angle of incidence. Quasi-specular reemission, which is the real-life alternative to diffuse reflection, produces a lobe of varying thickness centered on the vector with a surface angle equal to that of the angle of incidence. While in principle increased specular reflection could either increase or decrease the satellite drag, in actuality (at least for spheres), it has the effect of reducing drag and thus the drag coefficient.
3. Satellite shape. The angles of the incident surfaces to the satellite-atmosphere relative velocity vector change the reflection patterns and also govern flow about the satellite and thus affect reemission phenomena.

Finally, if random thermal motion is allowed, drag typically increases because the satellite surfaces parallel to the satellite-atmosphere relative velocity vector now have an increased opportunity for molecule incidence and momentum exchange.

In the regions of the atmosphere of interest to the present problem, the physical mechanisms of drag described above are influenced by an additional phenomenon: the satellite surface adsorbtion of molecular oxygen. A significant presence even at altitudes as high as 700 km (K. Moe & Bowman, 2005), adsorbed oxygen tends to push the accommodation coefficient toward unity and the reemission modality toward diffuse reemission, both of which then tend to reduce the drag coefficient. In addition to (and because of) these mechanisms, adsorbed oxygen has a homogenizing effect on satellite-atmosphere surface interactions, lessening the typical (or, more importantly, laboratory-observed) behaviors of the particular satellite material properties.

While the determination of a satellite's drag coefficient is certainly a complex process, what is necessary for the present purpose is to establish whether, as stated previously, it is reasonable to consider it to be essentially invariant over short periods (on the order of the OD fit-span), such as that of days or weeks. The mechanisms that determine this coefficient are affected by the gas composition of the portion of the atmosphere in which the satellite resides and the surface properties of the satellite (both dampened to a large degree by oxygen adsorbtion). The surface properties of the satellites are what they are, and the gas composition is a function of satellite altitude and solar activity, so it remains to determine whether either of these changes sufficiently over short time periods to influence the drag coefficient determination process.

3.3.1. Altitude

Pardini et al. (2009) assembled calculated C_D data for spherical satellites at a series of different altitudes during periods of high solar activity. Performing a linear fit to the data (which conformed well to such a representation, with an r^2 value of 0.98), a slope of 0.00081 C_D/km was obtained; this implies that in order to produce a 1% change in the C_D at 500 km, a 28 km change in orbit altitude is required—an extremely large change over a short period for a satellite not in the final throes of decay. So it is unlikely that the small-altitude changes sustained by a typical satellite over the period of at most a few weeks could affect a meaningful change in the C_D.

3.3.2. Solar Activity

Pardini's study does not also provide C_D data on test objects for periods of low solar activity, but it does give measurements of the accommodation coefficient under those conditions. If one assumes that the C_D is heavily correlated to this parameter, it is clear that C_D during low solar activity will be higher, and if one presumes that the two are perfectly and linearly correlated (which in reality they of course are not), the C_D difference for these spheres can be reconstructed from these data, is greater than 10% at 325 km, and should only get larger at higher altitudes. Because geomagnetic storms can increase the atmospheric density during solar minima to average levels typically observed during solar maxima, in principle short-term atmospheric density variations could well have a nontrivial effect on the C_D itself.

It must be remembered, however, that short-term density increases are brought about by a different mechanism (Joule heating) from general density increases during solar maximum (extreme ultraviolet [EUV] radiative heating), and Joule heating density increases may not bring the same mix of gases as those associated with the cyclical EUV heating. Furthermore, if surface oxygen adsorbtion, as the evidence indicates, governs the process heavily, then time must be allowed for the oxygen adsorbtion profile of the surface to change based on the short-term density increase due to the solar storm. While no known studies have been performed to determine the temporal aspect of oxygen adsorbtion of spacecraft materials in space-similar conditions, it is unlikely that the short-term density increases caused by Joule heating (and presumed similar increases in oxygen density), in which storm onset and recovery is typically complete after at most a few days, are long enough to alter the surface oxygen adsorbtion sufficiently to affect the C_D profile notably. In short, it seems reasonable to treat the C_D as invariant over short periods, even with changes in solar activity. The success of the High Accuracy Satellite Drag Model (HASDM) density model debiasing effort, in which extreme stability of ballistic coefficients for certain identified objects is presumed and this fact used to back out historical atmospheric density values through ensemble precision C_D fits for such satellites, is a further testament to the reasonableness of the presumption of C_D stability over short periods (Casali & Barker, 2002; Storz et al., 2002).

3.4. Satellite Frontal Area Variation

Because of the myriad forces to which they are subjected, most unstablized satellites undergo precession and thus present a varying frontal area profile along their velocity vector, and if the satellite is of irregular shape,

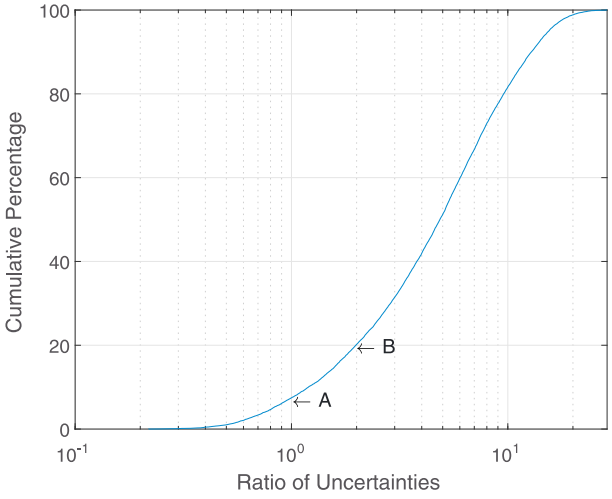

Figure 1. Ratio of atmospheric density uncertainty to frontal area uncertainty.

this area variation can be considerable and thus change the drag calculation substantially. The present question is the frequency with which the drag uncertainty caused by this variation is of comparable or greater significance than that caused by error in neutral density forecasting.

Fortunately, the JSpOC possesses and operates functionality to enable this assessment. The details of the algorithms are not publicly accessible, but a short description can suffice for the present purpose. For each satellite, historical ballistic coefficient values (generated far enough in the past to be calculated from fully issued space weather indices and durable HASDM debiasing coefficients) are temporally regularized and thinned (to produce only one ballistic coefficient term per OD update interval) and the standard deviation of the percent deviation from the global mean calculated. A similar calculation can be effected for the one-sigma percent error for the atmospheric density forecasting, as a function of predicated solar activity and satellite altitude; this has been done by comparing the Jacchia-Bowman-HASDM 2009 atmospheric model prediction performance to actual density values determined ex post facto by the HADSM functionality. These two calculations, one of which is relatively stable (the frontal area uncertainty, as it is built from ideally one year's worth of historical ballistic coefficient data) and the other dynamic (as it is altered daily in response to the forecasted solar indices), can be compared and thus a relative statement of these two uncertainties made for each object.

Figure 1 shows the ratio of atmospheric density forecasting drag error to frontal area uncertainty drag error for all satellites for which this calculation was possible during the month of November 2017. One can see that the frontal area uncertainty is equal to or greater than the atmospheric density forecast uncertainty in only 8% of the cases (point A), and in 80% of the cases, it is less than 50% of the density forecast uncertainty (point B). The cases in which frontal area variation is the principal contributor to drag uncertainty are thus very much in the minority.

Having examined all of the sources in uncertainty in the drag acceleration equation, one can conclude that other than atmospheric density modeling uncertainty, these other sources are either essentially invariant over short periods, potentially manifest variation only during solar storms (when all aspects of the CA problem encounter difficulties), or will exhibit variation at a level significantly subordinate to that of the density modeling error in the great majority of cases. One thus properly looks to the density model error as the principal error source to be understood and accommodated.

4. Relationship of Drag Acceleration Error to Resultant Pc

In examining the drag acceleration equation (equation (1) from the previous section), two conclusions can be drawn immediately. First, the atmospheric density is multiplicatively linked to all other terms in the drag calculation, so any errors in the density estimate flow directly to the calculated acceleration: if the density estimate is incorrect by a factor of 50%, the resultant drag acceleration calculation will be misrepresented by that same amount. Second, because all of the terms in the expression are multiplicatively combined, the ballistic coefficient and the density estimate can be used as aliases for each other; if one wishes to increase or decrease the atmospheric density by a certain percentage, this can be accomplished by varying the ballistic coefficient by that same percentage.

Because it acts in the antivelocity direction, the immediate and largest manifestation of drag acceleration, and similarly drag acceleration error, is in the satellite in-track velocity and therefore in-track position. This change in in-track satellite velocity also has a secondary effect on the orbit's semimajor axis and thus the satellite radial position and velocity. For most conjunctions, radial separation between the two orbits tends to govern the Pc calculation the most strongly, so errors that affect the radial component are quite likely to affect the satellites' relative position at TCA and therefore the collision probability.

However, it must be remembered that the covariance contributes substantially to the Pc calculation, and adjustments to the covariance to account for atmospheric density mismodeling can allow a correct and

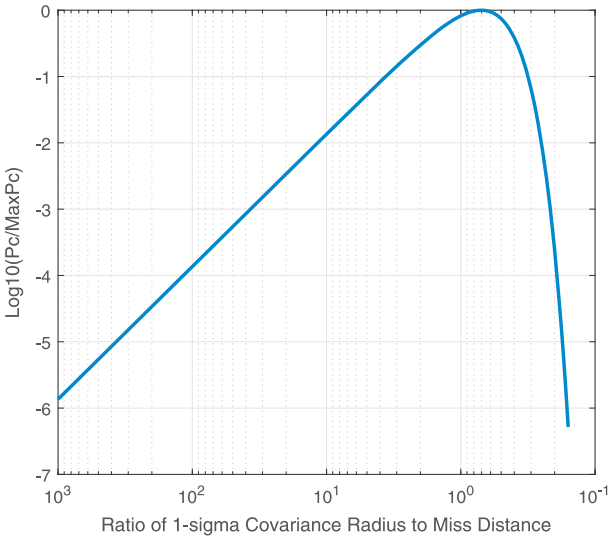

Figure 2. Relative Pc plot as a function of ratio of covariance size to miss distance.

useful Pc to be calculated even if the induced position estimation errors themselves cannot be remediated. The Pc represents the likelihood that given the uncertainties in the two satellites' positions at TCA, their actual miss distance will be smaller than a specified tolerance, called the HBR; this would be considered the equivalent of a collision (two satellite's flying closer than the HBR means that the "hard bodies" of the two satellites could contact each other, although it is possible that due to fortunate alignment of the two vehicles, an actual collision may not take place). If the uncertainty in the density estimates were characterized and known, there are techniques to include this uncertainty in the covariances and therefore the Pc calculation. These techniques, which will be discussed in section 5, allow a probability calculation to be executed that accurately represents the risks, given all of the known uncertainties of the situation.

The effect of increasing the state estimate uncertainty does not have a monotonic effect on the resultant Pc; rather, it is a function of the ratio of the size of the joint uncertainty of the two objects' state estimates to the miss distance. This is a subtle but important point and thus bears some additional discussion, aided by Figure 2, which provides a curve that gives the Pc value (as a ratio to its maximum value) as a function of the ratio of joint covariance size to miss distance (for simplicity, a spherical Gaussian covariance is used here, with its size represented by the sphere's radius). So long as the analysis is limited to situations with reasonable miss distances (i.e., not ridiculously large or vanishingly small), one can see that very large and very small combined covariances (very large or very small in relation to the miss distance, that is) will usually push the Pc to very small values, with a peak in the middle. In thinking about the implications of these conditions, this result makes physical sense. It is important to recall that OD processes produce an estimate of a mean state and a covariance that indicates the expected position (and velocity) dispersion about that state. When the uncertainties are very large, an estimate of the mean state is still the expected value, but it is not a very strong expression of central tendency; in the limit as the variance moves to infinity, the mean becomes merely the center point in a uniform distribution of vanishingly small density (at least for error in a single position component, which can be expected to follow a Gaussian distribution). So while the mean is still the most likely value, the uncertainty is so large that the likelihood of the two objects being at the mean position indicated by the OD and propagation processes is not great and the Pc is therefore small. Thus, if the two satellites' positions are poorly determined, then on the basis of the information available, the strength of a conclusion that they will actually pass in close proximity of each other will have to be weak. Conversely, when the uncertainties are very small, if the miss distance at TCA is notably greater than the HBR, one can conclude that the likelihood of an actual collision is low, thus also producing a small Pc. One can thus achieve a small Pc for two different reasons: because the knowledge of the satellites' states is poor enough that a definitive conclusion of a dangerously close approach cannot be credibly extracted from the data, or because the knowledge of the satellites' states is so good that one can state with certainty that a dangerously close approach will not occur. Alfano (2005) pointed out the difference between these two ways of achieving a low Pc and argued preferentially for the latter, as it is a conclusion stemming from good data rather than poorer data; Frisbee (2009) noted in response that in either case, the Pc is an appropriate assessment of the collision risk and that, while one would always prefer better to poorer data to enable decision-making, there is nothing improper or illegitimate in using the Pc from either "side" of the curve in Figure 2 for operational CA decisions.

Adjusting the covariance to account for atmospheric density estimation error aligns with the concept of reflecting all known state estimation errors into the accompanying state covariance matrix. While including this additional error source in the covariance can be expected to make the covariance larger, this increase, as discussed above, will not in all cases decrease the Pc value. The effect will depend on the ratio of the covariance size to the miss distance before the atmospheric density estimation error compensation is added. If one is well on the right side of the above curve, a modest increase in covariance size can have the effect of increasing the Pc substantially and thus changing the characterization of the event from "dismissible" to "serious." Extremely large density errors will probably decrease the resultant Pc in most cases, but that outcome is

not necessarily undesirable: if the uncertainty is properly represented, then the Pc reflects the actual ability to state that a collision is likely given the quality of the data available at the time an operational decision is to be rendered. The next section outlines some of the proposed and deployed methods for adjusting the covariance to incorporate known sources of error.

5. Accounting for Neutral Density Mismodeling Within State Covariances

Reflecting neutral density estimation error within state covariance matrices is a subset of the general problem of covariance or uncertainty realism, a topic that has been treated at length in a recent report of the Air Force Space Command Astrodynamics Innovation Committee (Poore, 2016). There are a number of techniques that are frequently employed to improve the realism of covariances—that is, their ability realistically and reliably to represent the actual state errors. Each of the major techniques is discussed briefly below.

5.1. Covariance Scaling

The governing presumption in single-factor covariance scaling is that the covariance is oversized or undersized by a scalar amount, so the covariance can simply be multiplied by a factor (actually typically the square of a factor) to make it more representative. Such factors are determined by analyzing past precision orbit data on the satellite in question (or ensemble sets of satellites with similar orbit maintenance properties), examining the relationship between the actual state errors and the statistical error summary represented by the covariance, and calculating a scale factor that will size the covariance so that it may accurately represent the statistical distribution of the actual empirical state errors (Hejduk et al., 2013, Poore, 2016). A natural extension of this approach is to calculate a scale factor for each of the three position components; if S represents a vector of three scale factors, scaling is accomplished by the matrix multiplication $S * C * S^T$. This approach is generally used to perform omnibus covariance realism corrections rather than to respond to a particular source of error (such as that from neutral density mismodeling).

5.2. Covariance Correction Matrices

An additional level of sophistication to simple covariance scaling was introduced through the work of Cerven (2011, 2013). Here entire correction matrices, rather than simple scaling vectors, are produced in order to attempt to correct the matrix for both size and orientation. Premultiplying and postmultiplying the covariance by the correction matrix produces a properly sized and oriented covariance. Again, applications of Cerven's method to this point have attempted only an ensemble correction for all covariance realism errors rather than those arising from a single source.

5.3. Physically Connected Process Noise

When using one of the many strains of Kalman filters to propagate covariances, provision is usually made to include process noise, or characterized acceleration uncertainties, which grow over time and can be propagated along with the covariance in order to increase the covariance size appropriately. In astrodynamics this was originally deployed to reflect characterized uncertainties in the applied geopotential model, for which processing requirements imposed practical limitations on the order of the model that could be applied, but in more recent times it has been used as an omnibus covariance correction methodology (Vallado et al., 2010) and as part of dedicated covariance realism efforts (Duncan & Long, 2006; Zaidi & Hejduk, 2016).

5.4. Consider Covariance Parameters

A traditional method to adjust covariances to make them more representative is through the use of consider parameters (Tapley et al., 2004). Such adjustments are determined from a priori error information external to the OD and are thus not solved-for but rather "considered" as part of the estimation process. This methodology has been embraced by U.S. Strategic Command, the DoD entity that supervises the production and distribution of the space catalogue; its purpose is to compensate for expected state propagation errors due principally to atmospheric density mismodeling. The covariance as formulated does not contain an atmospheric density uncertainty term directly, but because atmospheric density and the ballistic coefficient are multiplicatively coupled, one can alter the ballistic coefficient variance, which does appear in the covariance, and achieve an equivalent effect. If the percent error of the atmospheric density estimation is known, then the square of this amount can be added directly to the normalized ballistic coefficient variance. As the covariance is propagated (the usual method is through premultiplication and postmultiplication by a state

transition matrix), this increased variance will manifest an effect on the propagated covariance's position and velocity variances, as well as cross-correlation terms. In the present study, it is this approach (because it aligns with current DoD operational practice) that will be used to attempt to account for characterized atmospheric density modeling error. It is called a "dynamic" consider parameter (DCP) because its value is a function of the atmospheric density forecast and thus changes day to day.

All of these methods do improve the situation because they allow the covariance to accommodate density errors when they are known. Of course, the actual state errors due to density mismodeling are still present, and for the CA problem, in most cases, this will cause the miss distance between the two satellites to be incorrectly calculated, affecting the rectitude of the Pc. The only complete solution is to develop and deploy an atmospheric model with very low modeling errors. Nonetheless, the ability to incorporate error into the covariance is a preferable position over not considering the error at all, as it properly folds this characterized uncertainty into the probabilistic calculation.

6. Experiment to Determine Effects of Neutral Density Mismodeling on CA Calculations

6.1. Categorization Method for Severity of Conjunction Events

As stated previously, while investigations of effects on CA screening results or satellite propagated position errors are of interest, it is changes to the number of serious conjunction events that actually affect CA operational decisions and workload. Therefore, it is this parameter that must be determined in order to ascertain the actual effect of atmospheric mismodeling on the CA enterprise. It is straightforward to reprocess historical conjunctions using different atmospheric density values in order to see the effect on the resultant Pc, and in fact, such an approach is the basis for the present study, but simple comparison of ranges of Pc values gives data but not information that is properly operationally contextualized and therefore not particularly easy to interpret definitively. Instead, it has been shown to be helpful operationally to view Pc data within a "color coding" framework similar to one used by the CARA project as its method to communicate conjunction severity to its supported missions (Newman et al., 2014), and this approach is explained below.

Green conjunctions are conjunctions with a calculated Pc $<$ 1E−07. Such conjunctions will rarely rise to a level at which they would be considered serious, so no additional analysis effort is directed to them. Should the Pc increase as the event develops, the event can be recategorized. However, this particular threshold value was chosen so that the number of green events that eventually become serious remains below 0.1%.

Yellow conjunctions are conjunctions with a calculated Pc between 1E−07 and ~1E−04. Conjunctions in this range do have a reasonable likelihood of becoming serious as each event develops, so they are given additional monitoring, which typically includes manual inspection and massaging of the object ODs and requests for additional tracking, if it is felt that supplemental tracking will improve the OD and therefore the Pc calculation. The yellow category is typically a way-station for conjunctions on a path to either a red or green status (usually the latter), but some number of events reach their TCA still in the yellow category.

Red events are conjunctions whose calculated Pc is 1E−04 to 4E−04 and higher (the precise value depends to some degree on the particular protected spacecraft; for the purposes of the present study, a threshold value of 1E−04 has been used). Such events are considered serious, and their presence engenders a considerable amount of additional analysis, including in most cases the production of a high interest event presentation to be delivered in person to the satellite owner/operator. Events still in red status at about three days to TCA will often require remediation. At that point, a parallel analytical effort also begins to examine conjunction remediation options based on the methods available to the particular spacecraft and to make a series of recommendations to the owner/operator.

Given this approach to event categorization, the present analysis can usefully frame results in terms of the color changes sustained by events as atmospheric density estimation error is added to the Pc calculation; this will reveal the sensitivity of the calculation to such density errors. For example, if a particular event's color, with density unaltered, is green, the density can be increased/decreased incrementally up to a predetermined scale factor and the color change of the event, if any, noted for each increase/decrease level. This change can be accomplished both with and without attempted modifications to the covariance to account for expected atmospheric density mismodeling error. These color changes can be thought of more

Table 1
Nine Data Extraction Periods of CARA CA Events

Period Start	Period End	Median F10	Median Ap	Peak Ap	Peak Dst	Numbers of events			Solar Level
						Green	Yellow	Red	
20/6/15 20:24	29/6/15 2:03	108.4	15	240	−170	514	43	9	High
6/9/15 7:55	15/9/15 6:33	85.6	15	140	−94	556	18	8	High
31/8/16 21:19	9/9/16 11:27	95	15	80	−56	1,755	67	17	High
24/4/16 19:11	2/5/16 21:23	92.1	7	80	−41	1,435	54	13	High
13/11/15 12:57	19/11/15 20:18	106.1	7	48	−30	294	6	2	Moderate
15/5/16 0:39	21/5/16 8:55	102.1	7	48	−34	947	13	8	Moderate
17/6/16 13:21	26/6/16 0:22	79.7	7	48	−21	2,008	51	18	Moderate
28/4/15 17:02	4/5/15 15:04	105.25	5	15	−19	711	30	4	Low
14/8/16 2:07	21/8/16 1:58	82.8	3	15	−23	1,107	18	12	Low

classically as both false alarm and missed detection errors. Events that with perfect density knowledge would be red or yellow but with density error become yellow or green are missed detections and are certainly the more worrisome—these are either serious or concerning conjunctions that, due to density modeling errors, are not being afforded the requisite amount of analysis and attention and not being considered for potential remediation. However, given the workload associated with nongreen events, false alarms (here events that are yellow or green when perfect atmospheric density values are used but instead emerge as red or yellow when errors are introduced) are also of concern. Significant amounts of analysis and workload, as well as risk mitigation orbit alterations and all of the inherent risks that those processes contain, can be applied to events that in actuality are not risky. So while of course missed detections are the more serious error condition, it is operationally important to try to reduce both types of errors as much as possible.

6.2. Proposed Experimental Data Set

As stated above, the proposed experiment is to reprocess conjunction event data with atmospheric density model error injected into the situation; to do this requires choosing particular groups of conjunction events for this reprocessing. It seems appropriate to select results from routine CARA CA screenings that took place at different levels of solar activity. It may turn out that data segmented this way do not produce results that behave appreciatively differently, but nonetheless, it is good practice to allow for control of this variable in experiment construction. In order to add density model error to the satellites' propagated results, one must obtain their state estimates at epoch; this type of data is not contained in the usual distribution products and must be obtained independently, and it is readily available only for the last few years. This has limited the spectrum of space weather situations that can be investigated to only those available near solar minimum: the solar radio flux value will remain rather low, but solar storms will cause the geomagnetic indices to spike, allowing representation over a much greater portion of their natural range.

Nine two- to three-day periods of CARA conjunction results, taken from periods of different levels of solar activity, were selected for reprocessing. Table 1 gives vital statistics for each of these nine periods.

A number of items related to this table require discussion. First, one may wonder why such long periods of time are indicated for each of the nine data captures, especially when it is previously stated that each capture contains only two to three days of actual event data. The indicated period is intended to give the entire range of time during which satellite state and covariance propagation might take place. When a CA screening is performed, it looks a number of days into the future and catalogues close approaches that take place over that interval, so the TCA can take place several days subsequent to the screening time (for this experiment capped at four days from the time of screening). Additionally, while the primary object is typically updated regularly and frequently, the secondary object, especially if is a small debris object that is trackable by few sensors (often only one sensor), may be several days "old" at the time of screening: the object's state and covariance will have to be propagated for several days just to get to the current time and then be further propagated to TCA. These periods thus represent the entire swath of time that encapsulates all of the propagation intervals for all of the events in the capture, and the space weather statistics provided for each capture are summary statistics over that period. Space weather indices summarized include F10, the solar radio flux at 10.7 cm that is a good proxy for EUV heating (and is low in all nine periods due to their occurring near solar minimum); Ap,

Table 2
Atmospheric Density Scale Factors as a Function of F10 and Ap

F10 Level	Ap level								
	10	20	30	40	50	60	70	80	90
80	0.30	0.47	0.60	0.69	0.76	0.82	0.88	0.95	1.01
90	0.35	0.55	0.69	0.79	0.87	0.94	1.01	1.08	1.16
100	0.41	0.63	0.79	0.91	1.00	1.08	1.16	1.23	1.31
110	0.48	0.73	0.91	1.04	1.14	1.23	1.32	1.40	1.49
120	0.56	0.84	1.05	1.19	1.30	1.40	1.49	1.59	1.68
130	0.64	0.97	1.19	1.35	1.47	1.58	1.68	1.79	1.89
140	0.74	1.11	1.36	1.53	1.66	1.78	1.89	2.01	2.12

which indicates geomagnetic disturbances and is thus an indication of Joule heating; and disturbance storm time (Dst), which is often used to predict and respond to solar storms. Major storms are typically indicated by a Dst value smaller than −75; one observes two such storms here, as well as two other periods with significant Ap spikes; these are all indicated as high solar activity periods in the table. Elevated Ap for shorter periods characterizes three other captures, allowing them to be categorized as representing moderate solar activity. Two final periods show a very quiescent state and are thus billed as low solar activity. There is, to be sure, a certain arbitrariness in the boundaries among these three levels. Fortunately, because (as will be seen) the final experiment results do not differ substantially among the three levels, one need not be unduly detained here establishing rigid criteria for such boundaries. Finally, the event count, divided by color, for each capture is given. There is some natural day-to-day variation in event density, but the major differences in event counts tend to fall between 2015 and 2016 captures, for in 2016 there was a movement from one to three CA screenings per day.

6.3. Proposed Experimental Methodology

As stated previously, the basic experimental methodology is to reexecute historical CA events with additional atmospheric density model error added to determine the effect of such additional error on the conjunction severity. While this approach is insensitive to the precise manner in which density error would actually be manifested (i.e., with geospatial, diurnal, and other variations), it is a reasonable and expeditious way to impose a global average density error. To do this, one must first establish a reasonable range of density errors by which these events are to be perturbed: this error is a combination of the error in the prediction of the solar indices that drive the atmospheric models and the inherent error in the models themselves. Vallado and Finkleman (2014) examined historical data sets of predicted and actual F10 and Ap values and determined that while under quiescent conditions the errors were usually reasonably bounded (e.g., ±1–5 solar flux units for F10), under perturbed conditions, these errors could easily become as large as ±40 solar flux units for F10 and ±40 gammas (units for Ap) for Ap. To determine the possible density errors that could arise from such variations, the matrix given in Table 2 was generated, in which perturbations up to this magnitude were introduced in laddered fashion into the Jacchia 70 atmospheric model and the resulting density differences calculated (a floor value of 80 for F10 was established as values below the mid-70s rarely occur, except during solar minima).

A base solar activity level of 100 in F10 (similar to the levels observed in the nine captured data sets) and of 50 in Ap (to allow dynamic range when adding or subtracting Ap error) was chosen, and the table entries represent ratios between the density at the new level and that produced for the base value of 100/50. For the curious, a latitude/longitude location of (40, 270) degrees and a geodetic altitude of 600 km were used, as well as a seasonal date of 1 January 2016. In examining the table entries, one observes a minimum value of 0.3 and a maximum value of 2.1; so if one wishes to work from the base value of 100/50 as a center point, a span of scale factors to apply to the density roughly equal to this range (i.e., ~0.3 to ~2.1) should be appropriate. Under solar storm conditions, it would be possible, in cases in which such storms were either completely errantly predicted or entirely missed in prediction, to achieve as a scale factor something close to the ratio of the full extrema in the table: 2.12 (bottom right of table)/0.3 (top left of table), or ~7.1; Bruinsma and Forbes (2007), for example, have shown that severe solar storms can increase neutral density values by up

to 800%. Inherent errors in the neutral density models, which by rule of thumb have been historically considered to be about 20% (Alfriend et al., 1999) but have been improved in recent times to the 10% to 15% range (Vallado & Finkleman, 2014), are very much overshadowed by the atmospheric density forecast error, but such errors would act only to increase the effective scale factors and give them an even greater range. However, as will be seen in section 7 (Experimental Results), substantial differences in CA results, in some cases to the point of complete recategorization of all events in a risk group, are observed with the more modest range of scale factors indicated in Table 2. While a broader span of scale factors could be justified and would in all likelihood produce even more extreme results, both durable conclusions and a certain ease of presentation are achieved by using the midpoint-based scale factor range summarized in Table 2 and approximated by boundary values of 0.5 and 2. This is thus the approach that has been chosen, more specifically by implementing the following graduated set of scale factors: [0.5, 0.67, 0.77, 0.83, 0.90, 1, 1.1, 1.2, 1.3, 1.5, 2]. This set of factors also gives insight into the effects of relatively small-scale factors, which can help to ascertain the benefits of relatively minor improvements to intrinsic model and solar index forecast errors (which would confer maximum benefit during solar minima, in the absence of difficult-to-characterize space weather perturbations).

Because the ballistic coefficient and atmospheric density are multiplicatively coupled (equation (1)), density error was "added" by multiplying the ballistic coefficient for the primary and secondary objects by the desired scale factor. It is important to recognize that such an approach, while straightforward, does impose certain limitations. For example, all atmospheric density models of any consequence include latitude dependencies, so an Ap increase (or error) will not affect the model's density field uniformly, but imposing a density scaling factor by altering the ballistic coefficient essentially scales density values equally in all atmospheric regions. Results would be more or less equivalent between the two approaches for satellites with similar orbital parameters but would diverge for conjunctions between notably different orbit types. Because this study employed the U.S. Air Force (USAF) Astrodynamics Standards precision propagator in order to ensure compatibility between the OD theory that generated the satellite state vectors and the propagation engine, it was not possible to modify the density field directly, as this propagator is distributed only as a compiled product with the density model fully embedded and thus inaccessible. However, the expectation is that the sample size used in this study—nine multiday periods of screenings against an entire satellite catalogue—is large enough that differences introduced by this simplification will tend to level out statistically: in some cases, the multiplicative alteration of the ballistic coefficient will result in conjunctions that are more risky than would have been produced by a nuanced altering of the density field, and in other cases it will result in conjunctions that are less risky, so it is believed that the overall result would be similar under either approach. A study performed recently by Bussy-Virat et al. (2018), which was similar in intent to the present analysis but instead performed a detailed investigation of a small number of conjunction examples, did in fact account for density field variations, and their findings showed similar variations in conjunction severity to those of the present analysis (which will be presented in section 7). A fully definitive adjudication of this general question would in principle begin with a comprehensive statistical investigation of the forecast errors present in space weather indices (the Bussy-Virat et al. study makes a good beginning at this but investigates only two years' worth of data and limits itself to F10 and Ap) and, using the statistical models generated from such an investigation, generate statistically consistent space weather index errors that could then be applied to the input indices. Density errors would then arise naturally from plausible space weather index forecast errors and vary appropriately by latitude and any other parameterizations of the atmospheric density model employed. When such a statistical examination has been conducted for all of the space weather indices presently utilized (such as the more recent S10, M10, and Y10 EUV indices and the Dst geomagnetic index) over an appropriate time period, it would be very much desirable to reaccomplish a study such as the present one, employing this additional information. However, until that information is available, it is seen as adequate to proceed with the present simplified methodology, as it is believed that it is both broadly representative and consistent with other efforts.

After this adjustment, both objects were propagated forward in time, a new TCA was located, and a new Pc was computed. Propagation was accomplished using the USAF Astrodynamics Standards precision propagator. When this propagator is used in so-called "Vector Covariance Message" mode, an interesting approach is taken to atmospheric modeling: the original ballistic coefficient for each satellite is solved for using the Jacchia-Bowman-HASDM 2009 (JBH09) model, but the reported ballistic coefficient contained in the Vector

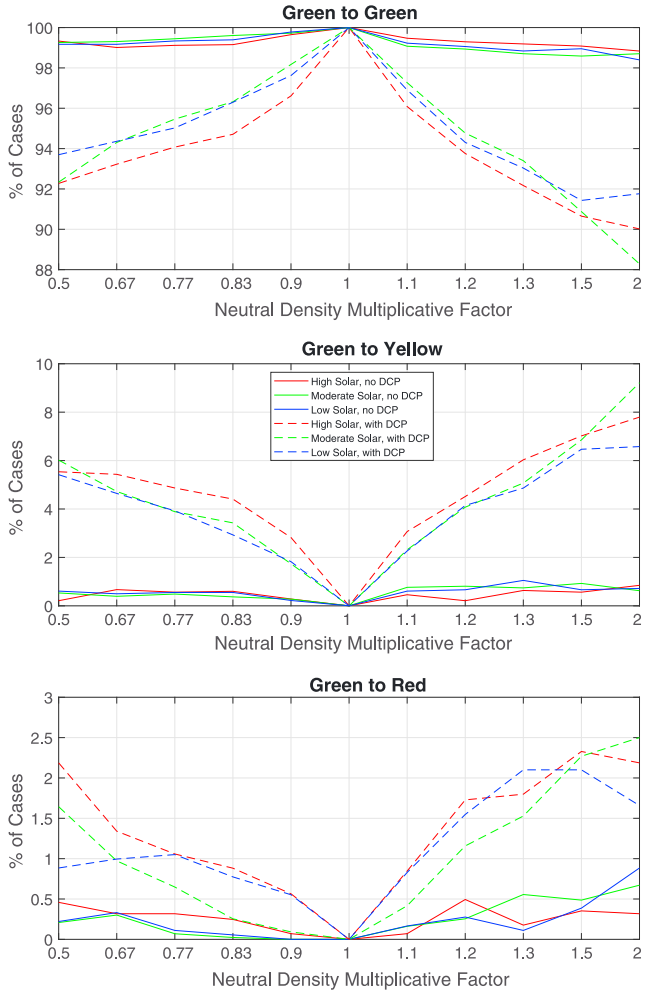

Figure 3. Color changes for green events as a result of density error addition.

Covariance Message has been regressed in order to reproduce the JBH09 output when propagated with the Jacchia 70 atmospheric model. This allows general users—and the particular experiment at hand—the ability to reproduce the output of the higher-fidelity JBH09 model with the more portable and computationally efficient Jacchia 70 model. As stated previously, the outcome of the present experiment is largely agnostic to the particular atmospheric model employed because state propagations with that particular model are defined as the truth state, to which density error is added and the results compared to that state, but nonetheless, it is desirable to use a model mode that is designed to emulate the performance of JBH09, the USAF's high-fidelity atmospheric model.

Finally, as part of the application of each scale factor to each conjunction event, two covariance propagation possibilities were executed. The first was the simple propagation of the two covariances as provided, without alteration. The second involved the addition of the model error percentage variance to the ballistic coefficient variance, in accordance with the DCP methodology described in section 5.2. This methodology allows for the consideration of atmospheric density modeling error in the Pc calculation.

The specific questions this experiment wishes to answer are the following:

1. What percentage of events are color-recategorized (e.g., green to yellow and red to green) at what levels of density error? Are the percentages significant even at low levels of density error?
2. What is the significance of the types of recategorizations observed? Are the majority false alarms (sometimes called Type I errors), in which a nonserious (say green) event is mistakenly categorized as a significant event (say red), or are they missed detections (sometimes called Type II errors), in which a serious event is mistakenly recategorized as a nonserious event? The Type I/II terminology arises from a hypothesis testing context and relates to whether or not the null hypothesis is appropriately dispositioned. In applying this terminology to the present problem, the null hypothesis is chosen to be of the form "The estimated distance between two space objects at their point of closest approach is greater than the combined hard-body radius of the objects." Thus, a rejection of the null hypothesis is the conviction that a dangerous satellite proximity event will occur. A Type I error is an errant rejection of this hypothesis, that is, the belief that a collision is likely when the evidence actually indicates otherwise; it is thus a false alarm. A Type II error is an errant retention of this null hypothesis, that is, the belief that a collision is unlikely when the evidence actually indicates that it is likely; it is thus a missed detection. This latter type is of course more worrisome, as it represents a situation of serious risk that passes undetected and thus unanalyzed/unremediated. Nonetheless, the significance of false alarms should not be understated, as they can command a significant workload, wasting time and diverting analytical resources from other events that are truly serious. Are these recategorizations sufficient in quantity to serve as a motivator to develop more precise modeling (which would include improvements in both forecast error of solar indices and intrinsic model error)?
3. Does the incorporation of the model error in the covariance (through the use of a DCP) improve the stability of the results? Of the recategorizations observed, does the including of model error shift the error production to more "favorable" types (i.e., Type II to Type I errors), and does it reduce miscategorizations overall? Do the experiment results promote this approach as more desirable generally?

7. Experimental Results

7.1. General Remarks

The present experiment takes actual CA events and then reprocesses these events with graduated additions of atmospheric density error in order to determine the degree to which the event appears more or less risky

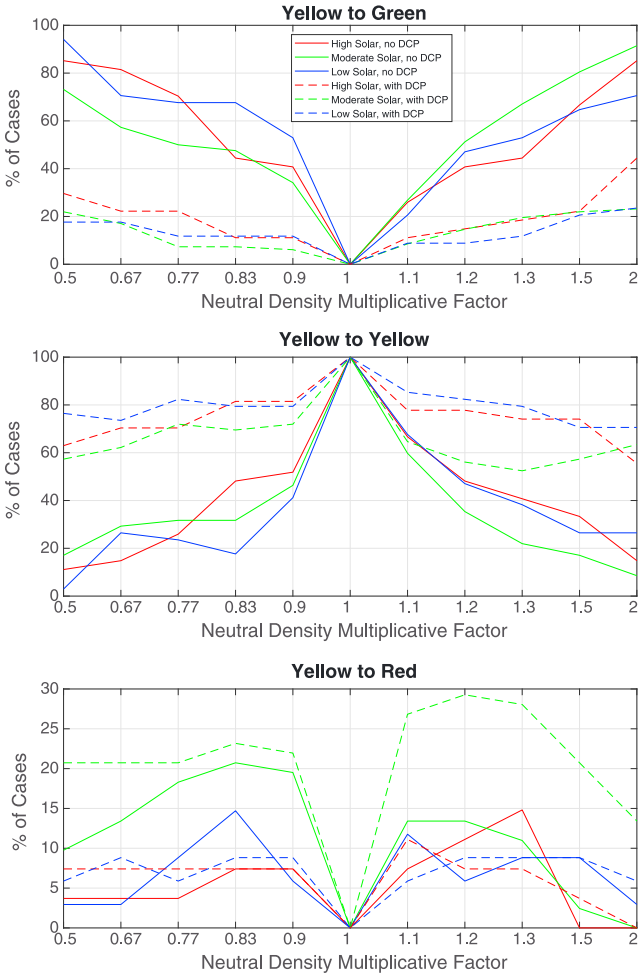

Figure 4. Color changes for yellow events as a result of density error addition.

in the presence of this error. As such, the original categorization of each event, without any density error introduced, is considered the "truth" result, and the differences between reprocessed versions of each event and this truth result determine the degree of increased or decreased risk wrought by the introduction of this error. The three sets of plots appearing as Figures 3–5 are a graphical attempt to communicate these variations in apparent risk due solely to increased atmospheric density model error, and each set is devoted to the group of events of a certain color (i.e., severity) before altering the situation: there is one set for events that in their unaltered state are green, one for events that in their unaltered state are yellow, and one for events that in their unaltered state are red. For each color group, there are three plots shown, each of which describes a certain type of recategorization. Taking as an example the "green" color group (Figure 3), that is, those events that without any added atmospheric density error are green events, there is one plot that, after adding atmospheric density error, remain green ("green to green", top); one plot for those events that, after adding the error, become yellow ("green to yellow", middle); and one plot for those events that, after adding the error, become red ("green to red", bottom). In each plot, a total of six lines appear. The three different line colors indicate the three different levels of solar activity represented in the sample data sets (red indicates high solar activity, green indicates moderate solar activity, and blue indicates low solar activity). The two different line styles indicate whether the density error was included in the covariance when the state covariances were propagated and the Pc calculated: a solid line indicates that there was no such modification of the covariance by the addition of a DCP in order to incorporate the expected density error, and a dashed line indicates that the covariance was in fact modified in this way. For each plot, the *x* axis gives the scale factor by which the density estimate has been multiplied in order to introduce synthetic error (ranging from 0.5 to 2) and the *y* axis gives the percent of cases that were recategorized by the application of each of the scale factors. As an overall example in reading these graphs, the bottom plot of Figure 3 (green to red) has a red solid line that shows a *y* value of 0.5 at an *x* value of 1.2. This means that for the high-solar-activity cases examined that were originally green events, when the covariance is not inflated to attempt to incorporate atmospheric density error and the density value is multiplied by a factor of 1.2, 0.5% of these cases become recategorized as red events.

In each plot, the ideal situation is for no events to be recategorized by adding atmospheric density error—this would mean that CA is insensitive to density errors, and therefore it is not necessary to work to improve atmospheric density models for the sake of this mission. Ideal results would thus be 100% in any plot that addresses color retention (e.g., green to green and red to red) and 0% in any plot that reports color change (e.g., green to yellow and yellow to red). The actual results are not always smooth or monotonic; this is believed to be due largely to the limited sample size but in a few cases may represent more systematic effects. While some plot anomalies can be explained straightforwardly, not all have such explanations. It is to be remembered, however, that the purpose of the study is not to establish individual recategorization numbers but general trends and the assessment of whether these trends are significant, and for this purpose, the experiment's results are believed to be adequate, if occasionally somewhat noisy.

Trying to explain each plot in prose in order to draw out conclusions of interest becomes quite tedious. As an aid to explanation of general trends, Table 3 below gives, for each plot, the recategorization percentage values at scale factors 0.5, 0.9, 1.1, and 2.0. Examining the situation this way shows the change in the "near field" of small density errors (0.9–1.1), which are easily understood to be possible with current forecasting and modeling, and also the "far field" of large (*r*) density errors (0.5–2.0), which previously discussed analysis shows to be within the realm of possibility. "High," "medium," and "low" refer to the levels of solar activity; "covariance comp." reveres to whether the covariance has been inflated with the DCP, and row and

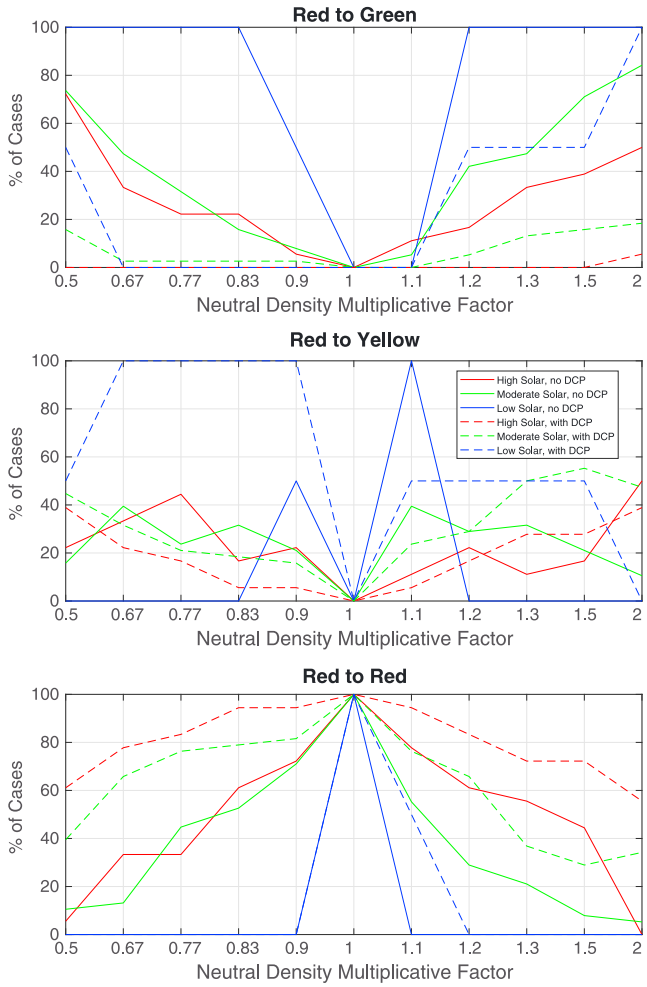

Figure 5. Color changes for red events as a result of density error addition.

column designations provide indices for easier reference to particular table values. The following streamlined discussion of individual results makes use of this table's summary of the results.

7.2. "Green" Results Set

Because of the great span of Pc values that constitute green events (from 0 to 1E−07), one would expect a relatively small number of recategorizations with modified density values; for example, a change in Pc value from 1E−50 to 1E−10, while large in a ratio sense, at both extremes still qualifies as a green event. The results bear this out: fewer than 2% of the cases without DCP (solid lines) are recategorized (1 AD; row 1, columns A and D in Table 3), and over the whole range of scale factors 8–10% of the cases with DCP (dashed lines) are recategorized (2 AD). While the amount of recategorization may be small (in a percentage sense at least), it is interesting and somewhat disappointing that the use of the DCP here produces notably less favorable results than for those cases for which the covariance was uncompensated, since compensation for atmospheric density error has been advocated as an expected improvement for CA calculations. Furthermore, because there is a large number of green events, relatively modest percentage levels of miscategorizations can produce nontrivial numbers of actual miscategorized conjunctions—additional red and yellow events that require manual review and analysis. However, it should also be emphasized that these miscategorizations here are all Type I errors, that is, false alarms. They do drive additional workload, and this outcome must not be neglected or trivialized; but it is not as serious of a situation as introducing Type II errors, in which high-risk events are passing undetected because they have falsely been given a less severe status. Further treatment of the trade-off between Type I and Type II errors will be given in the discussion section (section 8; Figure 3 and rows 1–6 in Table 3).

7.3. "Yellow" Results Set

It is always best to begin with the "nonrecategorized" results set for each color set; here it is the middle graph of Figure 4 and rows 9–10 in Table 3, representing yellow events that remain yellow. One notices immediately that the relative performance of the no-DCP and with-DCP cases is reversed with respect to the previously discussed green set: the with-DCP (dashed lines) now perform much better (~60–80% of cases preserved as yellow at the ends of the full span of scale factors [9A and 9D]) as opposed to 20% to almost 0% of the non-DCP cases over the same span (10A and 10D). Regarding these non-DCP recategorizations, most of these are reduced to green status: the percentages of recategorization in the yellow-to-green pane are substantial (7A–7D), whereas the yellow-to-red recategorizations are much smaller and more muted (11A–11D). The DCP-enabled cases show steady and reasonable performance in both of these cases (8A–8D and 12A–12D), with the possible exception of the moderate solar activity result for the yellow-to-red situation (12C and 12D, "Med" column; Figure 4 and rows 7–12 in Table 3).

Here the virtues of the DCP compensation are evident. Yellow-to-green recategorizations represent Type II errors, in that events that should receive increased monitoring and manual attention as well as mission notification are pushed into green status and thus not processed actively at all. Performance between the non-DCP and with-DCP cases is similar for the yellow-to-red situation, and while the DCP case performs somewhat worse for the moderate solar activity level (12C and 12D, Med columns), the effect here is to produce additional Type I errors—not a desirable outcome but certainly less serious than generating additional Type II errors.

7.4. "Red" Results Set

This set of results is more graphically angular, due mostly to the sample size being smaller (actual numbers of events considered were presented earlier in Table 1). Again, the nonrecategorization results set, here red-to-

Table 3
Summary of Experiment Results (Consolidation of Figures 3–5)

			Column A			Column B			Column C			Column D		
			SF = 0.5			SF = 0.90			SF = 1.1			SF = 2.0		
Row	Color Change	Covariance Comp	High	Med	Low	High	Med	Low	High	Med	Low	High	Med	Low
1	Green to Green	No DCP	99.3	99.3	99.2	99.6	99.7	99.8	99.5	99.1	99.2	98.8	98.7	98.4
2		With DCP	92.3	92.3	93.7	96.6	98.2	97.6	96.1	97.3	96.9	90.0	88.3	91.8
3	Green to Yellow	No DCP	0.2	0.5	0.6	0.3	0.3	0.2	0.5	0.8	0.6	0.8	0.6	0.7
4		With DCP	5.5	6.0	5.4	2.8	1.7	1.8	3.1	2.3	2.3	7.8	9.2	6.6
5	Green to Red	No DCP	0.5	0.2	0.2	0.1	0.0	0.0	0.1	0.2	0.2	0.3	0.7	0.9
6		With DCP	2.2	1.6	0.9	0.6	0.1	0.6	0.8	0.4	0.8	2.2	2.5	1.7
7	Yellow to Green	No DCP	85.2	73.2	94.1	40.7	34.1	52.9	25.9	26.8	20.6	85.2	91.5	70.6
8		With DCP	29.6	22.0	17.6	11.1	6.1	11.8	11.1	8.5	8.8	44.4	23.2	23.5
9	Yellow to Yellow	No DCP	11.1	17.1	2.9	51.9	46.3	41.2	66.7	59.8	67.6	14.8	8.5	26.5
10		With DCP	63.0	57.3	76.5	81.5	72.0	79.4	77.8	64.6	85.3	55.6	63.4	70.6
11	Yellow to Red	No DCP	3.7	9.8	2.9	7.4	19.5	5.9	7.4	13.4	11.8	0.0	0.0	2.9
12		With DCP	7.4	20.7	5.9	7.4	22.0	8.8	11.1	26.8	5.9	0.0	13.4	5.9
13	Red to Green	No DCP	72.2	73.7	100.0	5.6	7.9	50.0	11.1	5.3	0.0	50.0	84.2	100.0
14		With DCP	0.0	15.8	50.0	0.0	2.6	0.0	0.0	0.0	0.0	5.6	18.4	100.0
15	Red to Yellow	No DCP	22.2	15.8	0.0	22.2	21.1	50.0	11.1	39.5	100.0	50.0	10.5	0.0
16		With DCP	38.9	44.7	50.0	5.6	15.8	100.0	5.6	23.7	50.0	38.9	47.4	0.0
17	Red to Red	No DCP	5.6	10.5	0.0	72.2	71.1	0.0	77.8	55.3	0.0	0.0	5.3	0.0
18		With DCP	61.1	39.5	0.0	94.4	81.6	0.0	94.4	76.3	50.0	55.6	34.2	0.0

red, is the most helpful initial graph to view. One observes that levels of recategorization for the non-DCP cases are again very high (almost complete recategorization by the ends of the scale factor interval [17A and 17D]), and the with-DCP cases always perform everywhere at least as well, and in most situations better, than the non-DCP cases. Note that for this plot set, all of the recategorizations are Type II errors and therefore represent situations in which high-risk situations are either degraded (to yellow) or effaced entirely (to green). In examining the specific recategorizations, one must separate the low solar activity results (blue lines) from the other two solar activity levels (red and green lines). For the medium and high solar activity cases, the more "mild" red-to-yellow recategorizations show recategorizations of 10 to 50% by the extremes of the interval, and both the non-DCP and with-DCP cases perform roughly similarly in that in some cases one fares better than the other, and vice versa (15–16A and 15–16D). For the more severe cases of red-to green recategorization, at each solar activity level, the with-DCP substantially outperforms the non-DCP situation (17–18A and 17–18D). The low solar activity situation presents an unusual graphical appearance, especially in the red-to-yellow plot. To understand the behavior, it is helpful to look at all three plots together. In the red-to-yellow graph, one can then see that the spike from 0 to 100% for the non-DCP case at a scale factor of 1.1, and then back down to 0% at scale factor 1.2, takes place because at the scale factor of 1.2, all of the red events that were recategorized as yellow at scale factor 1.1 have now become further recategorized to red; the dropping back to 0% therefore is thus actually an indication of a worse rather than a better outcome. Similar behavior is seen for this level of solar activity for scale factors less than unity: a quick spike that then retreats to zero when 100% of the cases are recategorized as green. In both situations, the with-DCP results remain within the yellow realm, thus manifesting better performance (Figure 5 and rows 13–18 in Table 3).

8. Discussion

The number of different cases evaluated above (e.g., different color changes, with and without DCP, large and small spans of scale factors, and different levels of solar activity) can make it difficult to draw direct and durable conclusions. In trying to simplify the situation, one observes that in the main, the miscategorization behavior among the three different solar activity levels in any given situation is not all that different; on

Table 4
Experiment Results Applied to a Typical Screening

		Scale factor 0.9–1.1				Scale factor 0.5–2.0			
		Unweighted		Weighted		Unweighted		Weighted	
Color change	Severity	No DCP	With DCP	No DCP	With DCP	No DCP	With DCP	No DCP	With DCP
Green to Yellow	1	1.9	10.1	1.9	10.1	2.6	29.3	2.6	29.3
Green to Red	3	0.3	2.4	1.0	7.1	2.0	8.0	6.0	23.9
Yellow to Green	2	9.1	2.6	18.1	5.2	22.5	7.2	45.0	14.4
Yellow to Red	3	2.9	3.7	8.8	11.1	0.9	2.4	2.6	7.2
Red to Yellow	8	2.8	2.3	22.8	18.7	1.1	2.6	9.2	20.5
Red to Green	10	0.9	0.0	9.3	0.3	5.6	2.2	56.0	22.1
Totals		18.0	21.2	61.9	52.5	34.7	51.7	121.3	117.5

most plots, the three lines representing the three different levels of solar activity stay in a reasonably bounded corridor. In order to provide summary data therefore it seems an acceptable stratagem to take the mean of those three values to represent miscategorization performance for a given scale factor value. Additionally, the use of percentages of total CA cases, while helpful when combining data from multiple screening events, obfuscates the degree to which certain outcomes actually are problematic for CA operations. In order to move to the more accessible framework of numbers of events, these percentages can be applied to the event distribution that arises from a typical screening. To determine this, typical CARA screening results from 2017 (for an ensemble set of ~42 protected payloads with perigee heights below 1,000 km) were profiled, and a typical screening produced, at the median value, 433 green events, 27 yellow events, and 7 red events. Applying the previously determined percentages to these event loading levels allows the actual expected number of miscategorized events of each color type for a typical screening to be calculated. Finally, some effort must be made to determine the relative severity, in terms of operational burden and worry, of the different types of recategorizations. The basic distinction, as described earlier, is between Type I errors (false alarms) and Type II errors (missed detections). The fact that the latter is more problematic than the former is clear, but the degree of increased severity is not, and it is further expected to be a function of the particular type of miscategorization (e.g., green-to-red versus yellow-to-green). Any such ranking of severity will of necessity be subjective, but long-term experience with operations suggests the following rules:

1. The Type I error associated with a miscategorization will always be less severe than the comparable Type II error (e.g., red-to-yellow worse than yellow-to-red), and often much less severe.
2. Both miscategorizations of an event to a red level (green-to-red and yellow-to-red) are equally severe; they are problematic in that they induce additional work, but such miscategorizations will often resolve themselves by approximately two days to TCA when propagation intervals and density forecast errors are both smaller.
3. A green event miscategorized as yellow is the least worrisome miscategorization; a yellow miscategorized as green is of course more worrisome but not particularly so, as yellow events are merely monitored and refined but not acted upon directly.
4. Red events miscategorized to a lower level are the most problematic outcome; for in such cases, the proper level for the event will probably not be established until approximately two days to TCA (if then), at which point it is very difficult to perform maneuver planning and arrange for maneuver ephemerides to receive a CA screening. Red events miscategorized as green is the very most problematic situation; red miscategorized as yellow is somewhat less so but still much more serious than the other recategorization types.

From the above, a set of relative weights of event severity is suggested and appears in Table 4, along with the number of miscategorized events per screening, as a function of scale factor size (the smaller range of 0.9 to 1.1, which may be even smaller than a routine level of density error; and the larger range of 0.5 to 2, which represents error possibilities in a perturbed situation) and the presence or absence of covariance compensation in the form of a DCP The "unweighted" column set gives raw numbers of recategorized events; the

"weighted" column set gives the product of the severity weight and the raw numbers of recategorized events. In both cases, smaller numbers denote superior performance. Rows that report Type I errors are shown in blue and Type II errors in orange.

Before discussing the significance of the numbers of miscategorized events that will be presented to an operational unit, it is helpful to examine the relative differences between the cases that omit or employ the DCP. For both the small-error (scale factor from 0.9 to 1.1) and large-error (scale factor from 0.5 to 2.0) situations, the use of the DCP increases the overall number of miscategorizations, only slightly in the small-error cases but substantially in the large-error case. However, a quick examination reveals that this increase is almost always of Type I errors. When the severity factors are applied to calculate the "weighted" sums, employing that the DCP improves the situation for both the small- and large-error cases: while employing the DCP does increase the number of Type I errors, which do result in unnecessary work and potential ops confusion, the DCP's reduction in Type II errors—errors that tend to create "late-notice" events for which there may not be enough time to plan and execute remediation—makes its use advantageous overall. For both the small- and large-error cases, the weighted sums of number of events are lower for the DCP-enabled cases (52.5 versus 61.9 and 117.5 versus 121.3, respectively), bearing out this conclusion. Admittedly, this result is dependent on the severity levels that are assigned to each of the different miscategorization types; if the severity for the Type II errors, especially for those that involve red-event miscategorization, is lessened relative to the Type I errors, then the DCP advantage becomes much smaller and can even disappear. However, operational experience with late-notice events reaffirms the seriousness, both technical and psychological, of the Type II situations and counsels the relative levels of severity chosen here.

Regardless of the covariance compensation approach used (i.e., with or without DCP), the numbers of miscategorized events are still significant. For the small-error case without DCP, 5.2 Type I and 12.8 Type II errors are generated, and the situation is essentially inverted when the DCP is added (16.2 and 5.0, respectively). While not all of these miscategorization types produce substantial operational perturbation (e.g., 9.0 of the 12.8 Type II errors for the non-DCP case are yellow-to-green miscategorizations), nonetheless this is not a trivial number of miscategorizations (around 20) to impose on operations—three additional red-coded events that are not truly red is a notable increase in operational labor, and 2–3 miscategorizations of red events to lower levels introduces significant potential for late-notice events that catch missions off-guard and often do not allow for desired remediation. The situation is only worsened in the large-error case, in which the totals of miscategorized events are larger, and many of these are significant (almost seven serious Type II errors for the no-DCP case, and only slightly better for the with-DCP situation).

Given these results, it is clear that CA operations would benefit from atmospheric density modeling improvements for both the quiescent and perturbed cases. In all situations, however, it is very important not just to improve the modeling but to generate a durable estimate of the model's actual error in any given circumstance; for CA performance in both small- and large-error cases is enhanced when this information is available and incorporated into the CA calculations.

9. Conclusions

The principal takeaway from the present experiment is that the ability to characterize the inherent error in neutral atmospheric density models, so that this error can be incorporated into space objects' state estimate covariances, can have a substantial effect on mitigating the effects of such errors on CA risk assessment. While this practice does increase the Type I error rate, it can notably improve the Type II error rate and prevent serious events from being dismissed outright. Most atmospheric models presently in operational use do not include any embedded error estimation so that error information, tailored to the particulars of the situation, can be provided to users. This estimation should be a standard feature for future models and would be an extremely useful independent study and enhancement effort for presently employed operational models.

The experimental results also show the overall value in reducing neutral atmospheric density estimation errors and, in particular, bringing them down to a bounded set of values about the true value. The larger error values explored in the experiment, which can represent commonly encountered situations during space weather events such as coronal mass ejections, can cause quite serious miscategorizations of events even when compensation for these errors is included in the objects' state estimate covariances. Furthermore, while some of these miscategorizations, such as the "red-to-yellow" Type II error (Figure 5, middle plot), may be

addressable with current operational procedures, this approach will not be tenable with large space catalogues, as there will be simply too many conjunctions of this type to examine individually. A robust future for CA risk assessment, in which both larger space catalogues and the full range of space weather events can be addressed, requires methods that bring estimation errors down to relatively small and bounded values over the entire range of space weather conditions.

Acknowledgments
Both authors have performed this research as part of their general duties in support of NASA's Robotic Conjunction Assessment Risk Analysis program. Neither author presently holds any other professional affiliations, contractual, or otherwise, and neither supports, or in the future expects to support, efforts to perform atmospheric modeling research and development. The data set used to perform the present analysis is a set of satellite conjunction screening results data from the NASA CARA program archive. Such screening results provide satellite state and covariance information for two conjuncting satellites at TCA. These data originate from the USSTRATCOM Joint Space Operations Center and are a Controlled Unclassified Military Information (CUMI) datatype; as such, they are not publicly releasable and are exempt from Freedom of Information Act (FOIA) requests. Because of this article's general interest to the space weather community, it has been granted a special exception from the AGU's study data availability policy. Special appreciation is extended to Eric Sutton of Air Force Research Laboratory and Megan Johnson of a.i. solutions Inc. for helpful review of and suggestions for portions of this paper and to Daniel Pachura and Alyssa Hollander of Omitron Inc. for extracting and assembling the needed data sets. Finally, the excellent comments and suggestions from the two anonymous reviewers substantially improved both the technical content and expressive clarity of the present article.

References

Alfano, S. (2005). Relating position uncertainty to maximum conjunction probability. *Journal of the Astronautical Sciences*, *53*(2), 193–205.
Alfano, S. (2013). "Determining a probability-based distance threshold for conjunction screening." AAS Space Flight Mechanics Conference (Paper #13–352), Kauai, HI, February 2013.
Alfano, S., & Oltrogge, D. (2016). "Probability of collision: Valuation, variability, visualization, and validity." AAS Astrodynamics Specialist Conference (Paper # 2016–5654), Long Beach, CA, August 2016.
Alfriend, K. T., Akella, M. R., Frisbee, J. L., Foster, D.-J., & Wilkins, M. (1999). Probability of collision error analysis. *Space Debris*, *1*(1), 21–35.
Bruinsma, S. L., & Forbes, J. M. (2007). Storm-time equatorial density enhancements observed by CHAMP and GRACE. *Journal of Spacecraft and Rockets*, *44*(6), 1154–1159.
Bussy-Virat, C. D., Ridley, A. J., & Getchius, J. W. (2018). Effects of uncertainties in the atmospheric density on the probability of collision between space objects. *Space Weather*, *16*, 519–537. https://doi.org/10.1029/2017SW001705
Casali, S. J., & Barker, W. N. (2002) ."Dynamic calibration atmosphere (DCA) for the High Accuracy Satellite Drag Model (HASDM)." 2002 AISS/AAS Astrodynamics Specialist Conference, Monterey, CA, August 2002.
Cerven, W. T. (2011). "Covariance error assessment, correction, and impact on probability of collision." 2011 AAS Space Flight Mechanics Conference (paper # 11–154), New Orleans, LA.
Cerven, W. T. (2013). "Improved empirical covariance estimation." 2013 AAS/AIAA Astrodynamics Specialist Conference (paper # 13–768), Hilton Head, SC.
Chan, F. C. (2008). *Spacecraft collision probability*. El Segundo, CA: The Aerospace Press.
Cook, G. E. (1965). Satellite drag coefficients. *Planetary and Space Science*, *13*(10), 929–946.
Duncan, M., & Long, A. (2006). "Realistic covariance prediction for the Earth science constellation." 2006 AAS/AIAA Astrodynamics Specialist Conference (Paper # 06–6293), Keystone, CO.
Emmert, J., Byers, J., Warren, H., & Segerman, A. (2014). "Propagation of forecast errors from the Sun to LEO trajectories: How does drag uncertainty affect conjunction frequency?" 2014 AMOS Technical Conference, Wailea, HI.
Emmert, J. T., Warren, H. P., Segerman, A. M., Byers, J. M., & Picone, J. M. (2017). Propagation of atmospheric density errors to satellite orbits. *Advances in Space Research*, *59*(1), 147–165.
Foster, J. L., & Estes, H. S. (1992). "A parametric analysis of orbital debris collision probability and maneuver rate for space vehicles." NASA/JSC-25898.
Frisbee, J. H. (2009). "Re-examining probability dilution." 2009 AAS Astrodynamics Specialist Conference (Paper #09–413), Pittsburg, PA.
George, E. R., & Chan, F. K. (2012). "Covariance-based pre-filters and screening criteria for conjunction analysis." 2012 AMOS Technical Conference, Wailea, HI.
Hejduk, M. D., & Johnson, L. C. (2016). "Approaches to evaluating probability of collision uncertainty." 2016 AAS Space Flight Mechanics Meeting (Paper # 16–241), Napa, CA.
Hejduk, M. D., Laporte, F., Moury, M., Kelso, T. S., Newman, L. K., & Shepperd, R. (2017). "Consideration of collision 'consequence' in satellite conjunction assessment and risk analysis." 26th International Symposium on Space Flight Dynamics (Pembroke, Ont.), Matsuyama, Japan.
Hejduk, M. D., & Pachura, D. A. (2017). "Conjunction assessment screening volume sizing and event filtering in light of natural conjunction event development behaviors." 2017 AAS/AIAA Astrodynamics Specialists Conference, Stevenson, WA.
Hejduk, M. D., Plakalovic, D., Newman, L. K., Ollivierre, J. C., Hametz, M. E., Beaver, B. A., & Thompson, R. C. (2013). "Trajectory error and covariance realism for launch COLA operations." 2013 AAS/AIAA Space Flight Mechanics Meeting (paper # 13–355), Kauai, HI.
Hoots, F. R., Crawford, L. L., & Roehrich, R. L. (1984). An analytic method to determine future close approaches between satellites. *Celestial Mechanics*, *33*(2), 143–158.
Mehta, P. M., McLaughlin, C. A., & Sutton, E. K. (2013). Drag coefficient modeling for GRACE using direct simulation Monte Carlo. *Advances in Space Research*, *52*(12), 2035–2051.
Mehta, P. M., Walker, A., Lawrence, E., Linares, R., Higdon, D., & Koller, J. (2014). Modeling satellite drag coefficients with response surfaces. *Advances in Space Research*, *54*(8), 1590–1607.
Moe, K., & Bowman, B. R. (2005). "The effects of surface composition and treatment on drag coefficients of spherical satellites." 2005 AAS/AIAAA Astrodynamics Specialist Conference (paper # 05–258), Lake Tahoe, CA.
Moe, K., & Wallace, S. D. (1998). Improved satellite drag coefficient calculations from orbital measurements of energy accommodation. *Journal of Spacecraft and Rockets*, *35*(3), 266–272.
Moe, M. M., & Wallace, S. D. (1993). Refinements in determining satellite drag coefficients: Method for resolving density discrepancies. *Journal of Guidance, Control, and Dynamics*, *16*(3), 441–445.
Moe, M. M., & Wallace, S. D. (1995). Recommended drag coefficients for aeronomic satellites. In *The Upper Mesosphere and Lower Thermosphere: A Review of Experiment and Theory*, *Geophysical Monograph Series* (Vol. 87, pp. 349–356). Washington, DC: American Geophysical Union.
Montenbruck, O., & Gill, E. (2005). *Satellite Orbits: Models, Methods, and Applications*. Berlin: Springer Verlag.
Narvet, S., Frigm, R. C., & Hejduk, M. D. (2011). "Assessment of uncertainty-based screening volumes for NASA robotic LEO and GEO conjunction risk assessment. 2011 AAS Astrodynamics Specialist Conference (Paper #11–432), Girdwood, AK.
Newman, L. K., Frigm, R. C., Duncan, M. G., & Hejduk, M. D. (2014). "Evolution and implementation of the NASA Robotic Conjunction Assessment Risk Analysis concept of operations." 2014 AMOS Technical Conference, Kihei, HI.
Newman, L. K., Hejduk, M. D., & Johnson, L. C. (2016). "Operational implementation of a Pc uncertainty construct for conjunction assessment risk analysis." 2016 AMOS Technical Conference, Kihei, HI.
Pardini, C., Anselmo, L., Moe, K., & Moe, M. M. (2009). Drag and energy accommodation coefficients during sunspot maximum. *Advances in Space Research*, *45*(5), 638–650.

Pilinski, M., & Palo, S. (2011). Drag coefficients of satellites with concave geometries: Comparing models and observations. *Journal of Spacecraft and Rockets*, *48*(2), 312–325.

Poore, A. B. (Ed.) (2016). "Covariance and uncertainty realism in space surveillance and tracking." Report of the Air Force Space Command Astrodynamics Innovation Committee, 27 JUN 2016.

Schamberg, R. (1959). Analytic representation of surface interaction for free molecule flow with application to drag of various bodies. In D. J. Masson (Ed.), *R-339 Aerodynamics of the Upper Atmosphere* (pp. 12–1–12–41). Santa Monica, CA: Rand Corporation.

Storz, M. F., Bowman, B. R., & Branson, J. I. (2002). "High accuracy satellite drag model." 2002 AISS/AAS Astrodynamics Specialist Conference, Monterey, CA.

Tapley, B. D., Schutz, B. E., & Born, G. H. (2004). *Statistical orbit determination*. New York: Elsevier Inc.

Vallado, D. A., & Finkleman, D. (2014). A critical assessment of satellite drag and atmospheric density modeling. *Acta Astronautica*, *95*, 141–165.

Vallado, D. A., Hujsak, R. S., Johnson, T. M., Seago, J. J., & Woodburn, J. W. (2010). "Orbit determination using ODTK version 6." Fourth International Conference for Astrodynamic Tools and Techniques, European Space Astronomy Centre (ESA/ESAC), Madrid, Spain.

Wang, W., Burns, A. G., Wiltberger, M., Solomon, S. C., & Killeen, T. L. (2008). Altitude variations of the horizontal thermospheric winds during geomagnetic storms. *Journal of Geophysical Research*, *113*, A02301. https://doi.org/10.1029/2007JA012374

Wiedemann, C., Stabroth, S., Vörsmann, P., Oswald, M., & Klinkrad, H. (2005). Size distribution of NaK droplets released during RORSAT reactor core ejection. *Advances in Space Research*, *35*(7), 1290–1295.

Zaidi, W. H., & Hejduk, M. D. (2016). "Earth observing system covariance realism." 2016 AAS/AIAA Astrodynamics Specialists Conference, Long Beach, CA.

Zhang, S. P., & Shepherd, G. G. (2000). Neutral winds in the lower thermosphere observed by WINDII during the April 4–5, 1993 storm. *Geophysical Research Letters*, *27*(13), 1855–1858.