Decision Support Tool for Risk Assessment & Maneuver Planning in Collision Avoidance

Alexander Ryan
Lead Aerospace Engineer
Angus Leung
Aerospace Engineer
David Shteinman
CEO/Managing Director
Mark Yeo
Lead Software Engineer
Dr. Matthew Hejduk
Chief Engineer
NASA Conjunction Assessment Risk Analysis (CARA)

ABSTRACT

As the quantity of orbital debris continues to grow, so too does the rate of conjunction messages that suggest possible collisions between high value payloads and debris. The abundance of these conjunction messages, and eventual misses, has led to a culture of ignored alerts, and an increase in satellite operation costs as a result of the frequent need to plan resources for maneuver planning and execution. The loss of "trust" in conjunction alerts is due to the poorly characterized evolution in probability of collision (Pc) as time approaches the time of closest approach (TCA) between two objects, as well as the interpretation of Pc in the context of maneuver planning. To address these problems, and in collaboration with the NASA Conjunction Assessment Risk Analysis (CARA) program, the Industrial Sciences Group has developed a novel Maneuver Decision Support System (MDSS) to assist satellite operations in conjunction assessment and Maneuver planning. It provides a meaningful and intuitive Urgency metric for actionable maneuver decisions, based on the physical dynamics of conjunctions. It is based on a forecast of the evolution of Pc over time and represents an advance over current methods that are in use for satellite conjunction monitoring and planning. The result is to give satellite operators a validated decision support systems to plan for maneuver execution or mitigation or monitoring up to 3 days before TCA

1. INTRODUCTION

As is well known there is a need for the development of a rigorous Space Situational Awareness (SSA) framework, which includes the goal of keeping track of objects in orbit by measuring and predicting their path with a regular cadence, and in turn predicting potential collisions. However, given the significant number of objects that are currently operating within the low earth orbit (LEO) regime, in addition to the numerous pieces of existing debris (>36,000 objects greater than 10cm in size), the task of deciding when to maneuver a satellite can be challenging for an operator [1].

To tackle the emerging problem of reliable conjunction analysis, Industrial Sciences Group (ISG) has developed a decision support tool to provide a recommended course of action up to 2.5 days ahead of a maneuver commitment point, derived from computed forecasts of future probability of collision (Pc), following the detection of a conjunction. The MDSS is comprised of three distinct components: 1) Prediction of the future statistical evolution of position covariance of both objects; 2) characterization of the potential change in miss distance between both objects; 3) interpretation of predicted trends into actionable maneuver decisions using a weighted urgency metric. The formulation of these metrics was tuned extensively in collaboration with members of the NASA CARA operations team to ensure that the decision support tool produced reliable and predictive recommendations to operators.

1.1 Tool capabilities

The MDSS automates a complex analytical workflow and increases the confidence in predictions of collision likelihood, allowing satellite operators to make confident, timely decisions for satellite maneuvers irrespective of the scale of satellite operations. The MDSS software includes the systematization of a previously subjective analysis and incorporates several innovative statistical methods into conjunction analysis. The MDSS automatically incorporates metrics that would currently require a skilled analyst's expertise to produce recommendations on a case-by-case basis.

The overall structure of the decision support tool developed is outlined in Fig. 1. The research article for the modular software package is written in MATLAB, as per the recommendations from NASA CARA; but there is now also a Python instantiation. The parameters for tuning or making any adjustments to the underlying functionality of the package (urgency metric functions and weightings, number of bootstraps etc.) can be found in the "set framework parameters" function. The core function is "Decision tool," as this imports CDM data and exports the final urgency metric (with associated metadata such as individual metrics). The core of the tool "PredictPc" uses a network of 13 Gaussian Process Regression (GPR) models, trained on a large dataset of CDMs, to individually fit primary and secondary object covariance elements, as well as miss distance volatility.

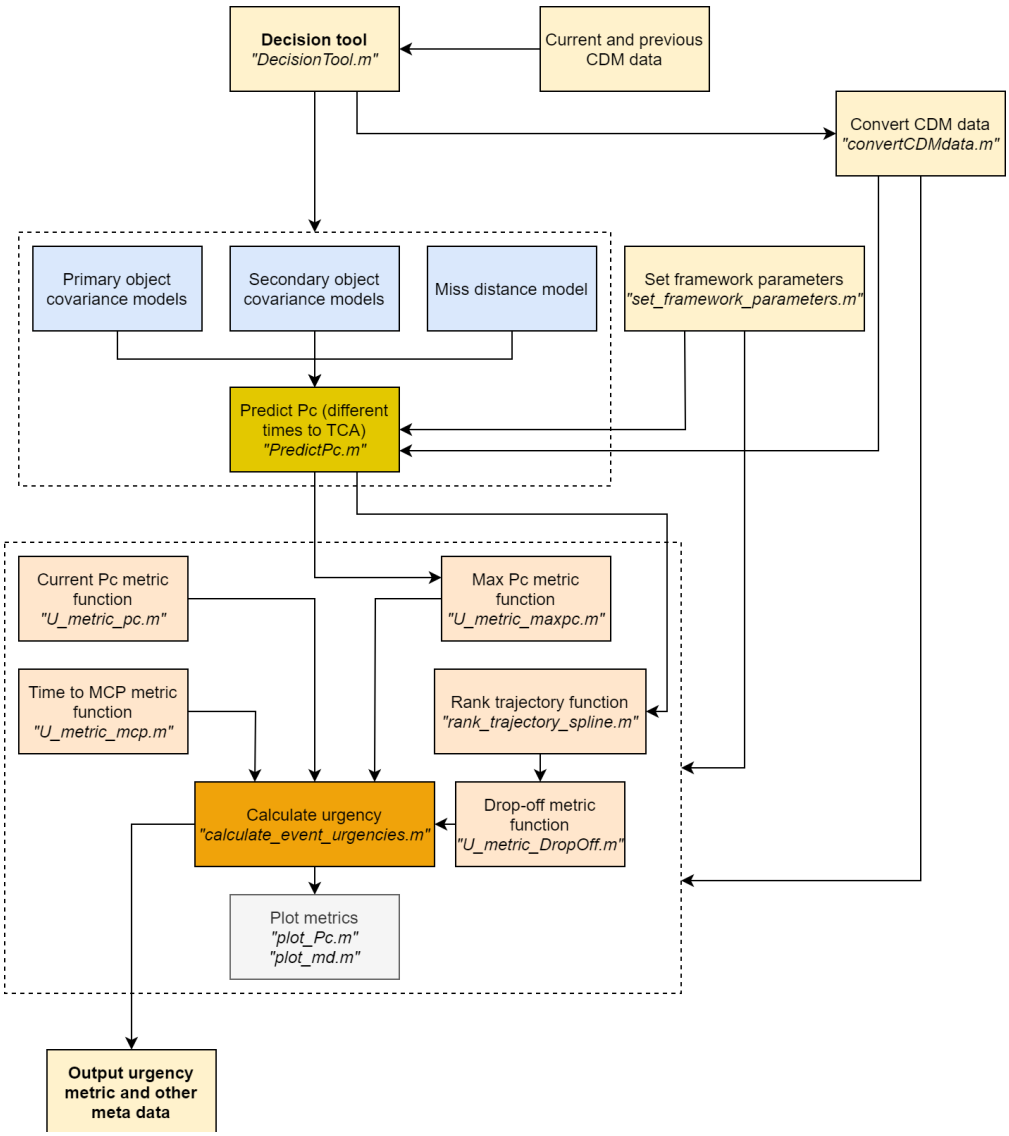

Fig. 1 Decision Support System (MDSS) structure.

The main driver behind the determination of a conjunction event's urgency is the Pc between the primary and secondary object. The primary object is the active spacecraft, and the secondary object can be either another intact satellite or a piece of space debris. Pc calculation has emerged as the standard way to assess the threat of a collision, and the means of filtering events for further assessment. In a single value, it captures the effect of *covariance size, miss distance, the orientation of the covariance ellipsoid and the size of the objects.* However, there are many sources of uncertainty that make Pc estimation change over the course of time including:

- Uncertainty in the model dynamics
- Sensor errors such as measurement noises and navigation biases
- Numerical errors and approximations
- Propagated uncertainties
- Spacecraft size uncertainty
- Atmospheric density forecast and space weather error
- Satellite frontal area stability uncertainty

There are ongoing efforts to negate the effects of these errors; however, they can never be eliminated. A key approach used by ISG to address them is the use of a supervised machine learning model called Gaussian Process Regression (GPR) to forecast Pc. By training on a database of CDMs from historical conjunction events, the tool predicts the covariance behavior of the two objects of interest over time, as well as the volatility of the miss distance (and hence the posterior distribution of change in miss distance at TCA).

1.2 Gaussian Process Regression

To determine the behavior of a conjunction event over time, the future evolution of covariance and miss distance must be forecast. As part of the MDSS we fit a total of thirteen Gaussian Process Regression models, six models for each object's covariance and one for miss distance volatility. This method was selected as it allows for the quantification of uncertainties in each Pc computation, and no restrictive assumption about the data is imposed. The crucial component that allowed us to develop this tool is the satellite event database supplied by CARA. This set of more than 400,000 conjunction data messages contains information on satellite ephemeris, drag, energy dissipation rate, etc. needed to fit time series regression models. These were collected over a four-year period, containing conjunctions involving 25 high-value NASA science satellites.

Gaussian Process is the generalization of a multivariate Gaussian distribution to infinitely many variables [2,3]. It is an extension of the idea of taking a probability distribution over numbers, to a *probability distribution over functions.* This idea is to regress a training set of input $\boldsymbol{X} = \{X_1, X_2, X_3, \dots, X_n\}$, where every X is a vector with dimensions equal to the number of regressors, to a corresponding set of outputs $\boldsymbol{Y}$. This is achieved by having a prior distribution over functions, which in Bayesian statistics represents the prior belief over the type of functions one expects before seeing any of the outputs.

This zero-mean prior distribution is visualized in Fig. 2 with a set of randomly drawn sample functions. Zero mean indicates that the average of the function so far does not depend on $\boldsymbol{X}$ and averages to 0. The grey band signifies the 2-σ standard deviation with a given input $\boldsymbol{X}$. When a corresponding output $\boldsymbol{Y}$ is observed with a corresponding $\boldsymbol{X}$, only functions that go through or pass near the 2 observations are now considered, as shown in Fig. 2b. The dotted lines now show a subset of the posterior distribution over functions, and the solid line is the mean of the posterior distribution. This mean function is denoted as $\mu(\boldsymbol{X})$, and the specification of the prior regarding its functional form is denoted as $K(\boldsymbol{X}_1, \boldsymbol{X}_2)$. The function $K(\boldsymbol{X}_1, \boldsymbol{X}_2)$, referred to as the kernel function, constitutes the core of the training process and plays a major role in its prediction accuracy. The estimated mean function is then used to make predictions $\boldsymbol{Y}_{new}$, on new incoming data $\boldsymbol{X}_{\boldsymbol{new}}$. For this tool, the 1-dimensional $\boldsymbol{X}$ in this example is extended to a 13-column vector.

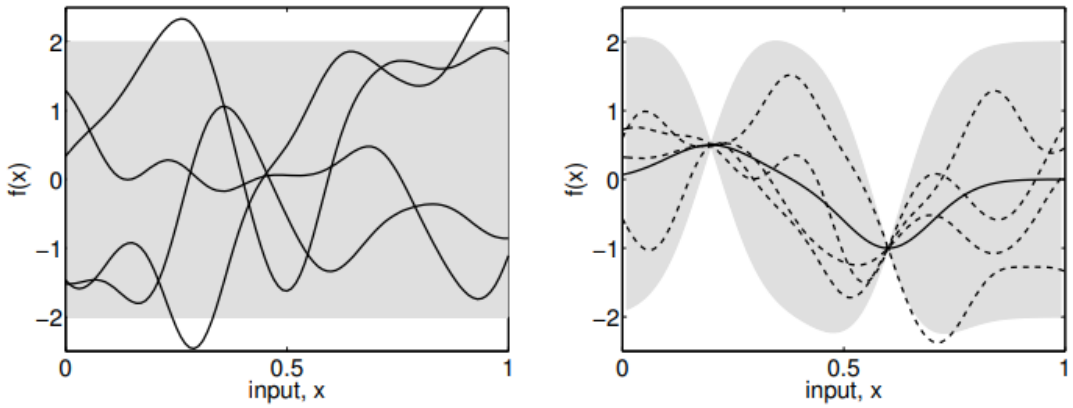

Fig. 2 Four samples drawn from a prior distribution (left) and a mean prediction (solid line) with four samples from the posterior distribution (right). The convergence of the posterior functions with 2 observations is also indicated (right) [2].

To get a higher resolution on the estimated mean function and lower the uncertainty bounds, a sufficiently-sized sample is required. Using the large data set provided (NASA CARA CDMs), we developed a mean function that can cover the (up to) 13-dimensional solution space for each model, which significantly contributes to the predictive power of the tool.

A significant part in the process of developing a GPR model is choosing the covariance function and setting the best hyperparameters. The kernel function below is the ARD (variable length scale) exponential kernel we used for covariance modelling:

$$\boldsymbol{K}(\boldsymbol{X},\boldsymbol{X}) = k\left(x_i, x_j \middle| \theta\right) = \sigma_f^2 \exp\left[-\sqrt{\sum_{m=1}^{d} \frac{\left(x_{im} - x_{jm}\right)^2}{\sigma_m^2}}\right]$$

The covariance function $\boldsymbol{K}(\boldsymbol{X},\boldsymbol{X})$, is a square matrix with dimensions in accordance with the training set size i.e., a model trained with 12 000 rows will have $\boldsymbol{K}(\boldsymbol{X},\boldsymbol{X})$ size 12 000 x 12 000. Each entry of $\boldsymbol{K}(\boldsymbol{X},\boldsymbol{X})$ contains a kernel function $k\left(x_i, x_j \middle| \theta\right)$, which transforms the difference 2 observations x_i and x_j into an exponential relationship. σ_m is the variance or length scale of each individual regressor, and σ_f is the function noise. The choice of this covariance function was based on research across different kernels, with ARD squared exponential demonstrating the minimum RMSE in predicting a consecutive covariance entry.

1.3 Covariance prediction

There are many formulations that exist currently that compute high-value collision probability in various ways. In our tool, the set of assumptions and simplifications are in accordance with Alfano in 2005 [4]. Both primary and secondary objects are assumed to be spherical, with the diameter defined by the maximum dimensions of either object. This eliminates complexities in relation to the attitude of the object when calculating Pc. The relative motion is considered linear, as even though both objects are orbiting the Earth, but the velocity is much greater in magnitude in comparison to its acceleration. The uncertainty of their positions is assumed to be described by a 3-D zero mean, static gaussian distribution.

The calculation of collision probability using the encounter plane is shown in the equation shown below:

$$P = \frac{1}{2\pi\sigma_{\boldsymbol{X}}\sigma_{\boldsymbol{y}}} \int_{-HBR}^{HBR} \int_{-\sqrt{HBR-x^2}}^{\sqrt{HBR-x^2}} \exp\left[\left(\frac{-1}{2}\right)\left[\left(\frac{x+x_m}{\sigma_{\boldsymbol{x}}}\right)^2 + \left(\frac{y+y_m}{\sigma_{\boldsymbol{y}}}\right)^2\right]\right] dy\, dx$$

Alfano (2005)

where:

- P – Probability of collision
- $\sigma_{\boldsymbol{x}}$ – Combined covariance's semi-minor axis in the encounter plane
- $\sigma_{\boldsymbol{y}}$ – Combined covariance's semi-major axis in the encounter plane
- HBR – Hard body radius
- x_m – miss distance along the semiminor axis
- y_m – miss distance along the semimajor axis

The HBR is derived from a circular simplification of object profile, and the miss distance values are results of propagation based on current observations. $\sigma_{\boldsymbol{x}}$ and σ_y are the semi-minor and semi-major axis of the projected 2-D covariance that lies in the encounter plane. The derivation of these quantities from the 3-D covariance matrix will not be listed, as the details can be found in Alfano's paper.

As previously mentioned, the problem with using Pc as a primary factor for determining a course of action leading up to TCA, is not knowing how Pc will evolve over time. A high value Pc 3 days before TCA will often see a significant decrease before 1 day to TCA. Currently CARA operators will receive approximately 1000 CDMs at approximately 8-hour intervals. Most are not significant events with very low Pc. However, as time progresses and increasing numbers of RSOs are launched, the chance of a significant event occurring will become exponentially higher. Then the issue of deciding which events to prioritise becomes much more difficult (with compounding limitations on decision making in a complex multi-variable environment under time pressure).

To perform forecasts of covariance for an object using CDM data, each covariance model comprised a series of GPR sub-models for each entry of the symmetrical correlation matrix. The correlation matrix is simply a transformation of the covariance matrix, which was introduced to eliminate asymptotic calculations and to address model stability when using logarithmically-scaled covariances.

$$Corr(\boldsymbol{X},\boldsymbol{Y}) = \frac{Cov(\boldsymbol{X},\boldsymbol{Y})}{\sigma_{\boldsymbol{X}}\sigma_{\boldsymbol{Y}}} = \begin{bmatrix} \frac{\sigma_1{}^2}{\sigma_1{}^2} & \frac{\sigma_{12}}{\sigma_1\sigma_2} & \frac{\sigma_{13}}{\sigma_1\sigma_3} \\ \frac{\sigma_{21}}{\sigma_1\sigma_2} & \frac{\sigma_2{}^2}{\sigma_2{}^2} & \frac{\sigma_{23}}{\sigma_2\sigma_3} \\ \frac{\sigma_{31}}{\sigma_1\sigma_3} & \frac{\sigma_{32}}{\sigma_2\sigma_3} & \frac{\sigma_3{}^2}{\sigma_3{}^2} \end{bmatrix} = \begin{bmatrix} 1 & \frac{\sigma_{12}}{\sigma_1\sigma_2} & \frac{\sigma_{13}}{\sigma_1\sigma_3} \\ \frac{\sigma_{21}}{\sigma_1\sigma_2} & 1 & \frac{\sigma_{23}}{\sigma_2\sigma_3} \\ \frac{\sigma_{31}}{\sigma_1\sigma_3} & \frac{\sigma_{32}}{\sigma_2\sigma_3} & 1 \end{bmatrix}$$

As seen in the equation above, there are 6 distinct terms that make up the matrix. These terms contain information on positional uncertainty in the RIC frame, as well as their respective correlations. From examining the database of CDMs, general trends for each correlation matric parameter were determined. By performing a detailed sensitivity analysis for covariance outputs across the CDM database, a list of 17 potential variables was chosen as inputs to the covariance model. These variables contain information about object energy dissipation, orbital data, and time to closest approach.

To produce the collision probabilities of the future CDMs, the tool performs a time series prediction of each of the elements of the correlation matrix, then reconstructs the covariance matrix. After a transformation of the covariance predictions for each object and then into the encounter plane, each prediction forms a 2-D covariance ellipse. An example of the forward covariance propagation is shown in Fig. 3. This figure is produced at ~3 days before TCA for

a particular event. The uncertainty ellipsoid for the conjunction event at 3 days before TCA is represented by the most outer ellipse. All the encompassing ellipses are the predicted uncertainty behavior, all the way up to 0.25 days to TCA. The red circle in the figure represents the combined hard body object. Since the red circle remains fully encompassed by the 1- σ covariance ellipse, an operator is given greater confidence that the conjunction should be taken seriously.

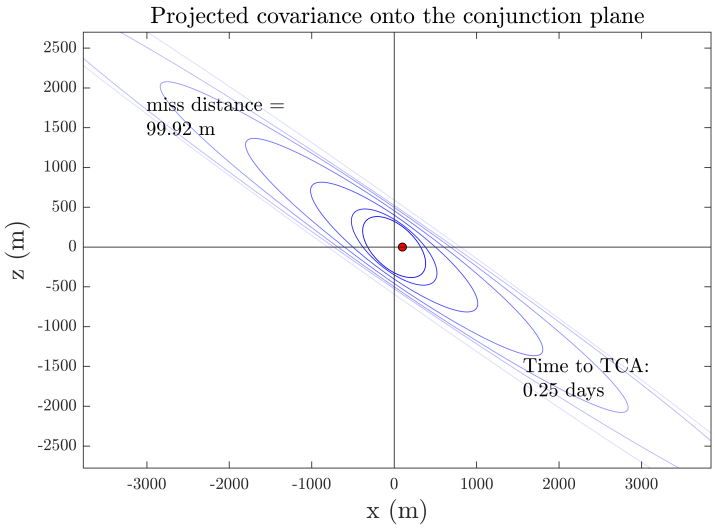

Fig. 3 Combined covariance prediction from 3 days in increments of 0.5 days up to 0.25 days before TCA. The most outer ellipse represents the combined covariance at 3 days, and the smallest ellipse is the predicted covariance at 0.25 days to TCA.

1.4 Miss distance volatility model

The calculation of Pc also involves another critical input, miss distance. Miss distance is the calculated distance between the 2 objects of interest at TCA. A significant factor impacting the effectiveness of the traditional use of Pc to decide on a maneuver several days ahead of TCA is the volatility over time in the predicted miss distance vector.

There are many factors that make predicting miss distance difficult, with the most critical being the effect of space weather on atmospheric density, hence drag and propagated position. Accurate atmospheric density models are an area of active research, however there is still significant uncertainty when it comes to predicted future solar activity. There are limitations when predicting the effect of solar activity over 2 days to TCA. Due to the rotation of the Sun and a lack of visibility of the opposing side of the Sun to earth, there are no detailed and continuous measurements of the Sun's behavior. This problem may evolve as our understanding of the Sun improves and satellites are launched to monitor activity in more detail.

Rather than predicting miss distance itself, we modelled the miss distance volatility and solar indices to determine a '*cone of expected changes in miss distance*". Rather than outputting a deterministic result using a constant miss distance, leading to a single Pc prediction at each time increment (as covariance forecasts are assumed as point predictions), the MDSS tool calculates the miss distance variance and truncates over time. Bootstrapping is then used to obtain a distribution of values over time. This is illustrated below in Fig. 4, where an initial CDM at ~3 days with a miss distance of ~3km is propagated to TCA.

As the atmosphere models have a limited predictive power as well as limitations in secondary object observations, the predicted miss distances tend to drift from a particular observation, as represented in the growing variance in the miss distance distribution over time. By using this method, the tool is able determine a level of likelihood of an event to drop off in collision probability, as discussed further in the following section.

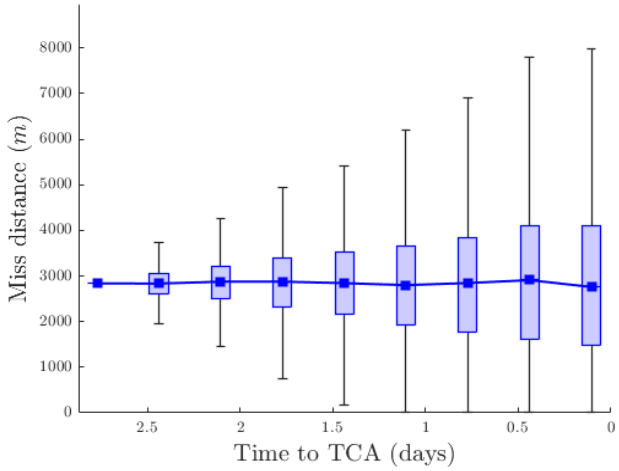

Fig. 4 Miss distance prediction, demonstrating potential miss distance ranges based on conditions at 3 days before TCA.

1.5 Pc prediction

Combining the covariance and miss distance predictions, the tool generates a distribution of future Pc as demonstrated in Fig 5. The example showcases the tool's prediction of the Pc behavior from 2.6 days to TCA, using only the CDM pair, received at 2.6 and 2.9, days before TCA. The whiskers of the blue solution represent the Pc distribution resulting from the miss distance distribution mentioned in the previous section. A shorter box and whiskers plot provides a higher level of confidence in the Pc prediction over time, given that the miss distance does not deviate significantly over the 2.6 days period.

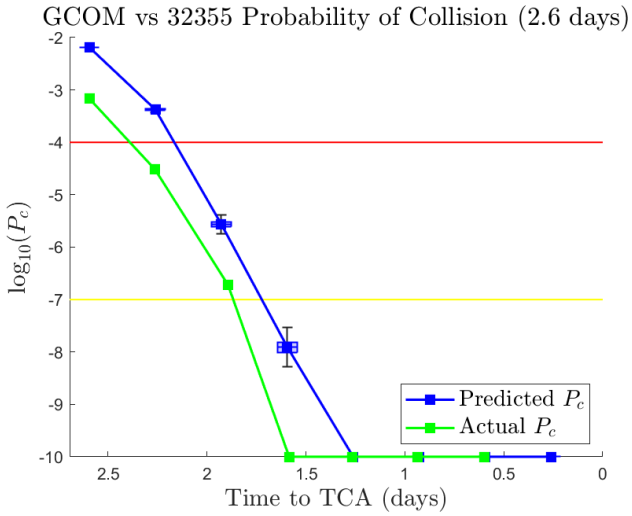

Fig. 5 Pc prediction, demonstrating potential Pc ranges based on conditions at 2.6 days before TCA in blue, and compared with the measured Pc evolution in green.

2. URGENCY METRIC

The predictions produced by our models, well as other meaningful conjunction information such as current Pc and time to maneuver commitment point, are quantified as time variable metrics in the range [0, 1]. Each of these metrics was combined using a weighted average to produce a single quantified **Urgency Metric**. However, an urgency value on its own can be ambiguous in its meaning. Therefore, the tool employs a useful and actionable recommendation. Each total urgency value (U_T) is categorized into one of the 5 color-coded bins, each one corresponding to a specific advised action. The advised actions are (from least to most urgent):

- Event dismissal (dark green): event can be dismissed as Pc is very unlikely to increase to a dangerous level at MCP
- Event monitoring (light green): Maintain monitoring of event but is currently not of great concern
- Preparatory mitigation planning (yellow): Begin initial phase/preparation for, potential maneuver planning
- Substantial mitigation planning (light red) Screen maneuver to ensure it does not lead to other collision events and prepare to act at any point
- Mitigation execution (red): Execute planned maneuver

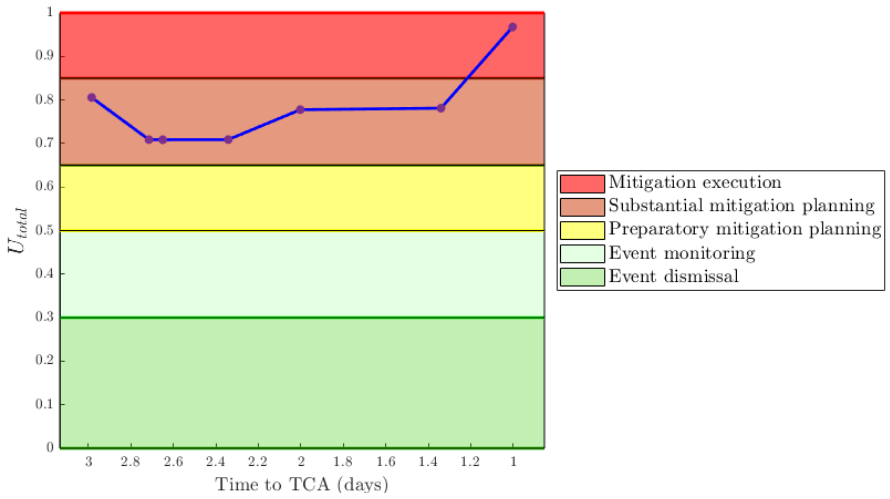

Fig. 6 Example urgency graph of time leading up to a conjunction event.

As mentioned, a core utility of the MDSS is to provide enhanced actionable decision support compared to traditional methods. For satellite operators that are new to the field or constrained by a budget, it is important to prioritise events that are significant and mitigate 'wasted' preparation time on insignificant events. With the categorisation of recommended actions, an owner can execute a precise action ahead of time (compared with standard MCP procedure) to ensure maximum operational efficiency. To maximise the amount of information extracted out of a single CDM, each metric encompasses different aspects of event data, both from the CDMs and ISG's models:

- Drop-off metric encapsulates the percent chance of a Pc drop off prior to the maneuver commitment point
- Current Pc metric: urgency based on the Pc of most current CDM
- Maximum Pc metric: the highest predicted median Pc up until TCA
- Time to maneuver commitment point: the time difference between current CDM to the MCP

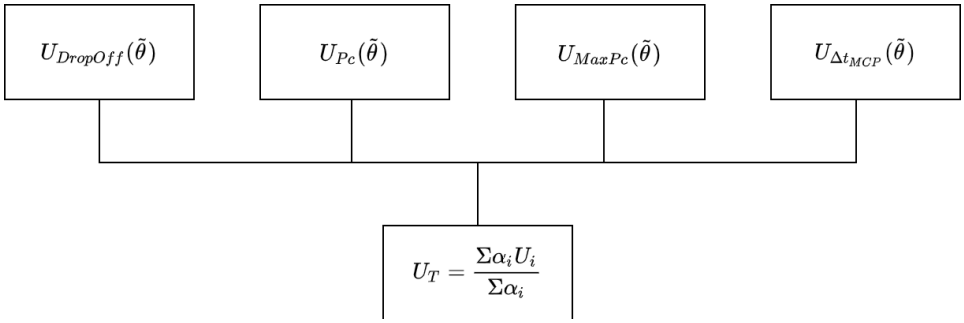

Fig. 7 Total urgency U_T flow chart.

2.1 MDSS parameter tuning

The MDSS tool is composed of a total of 17 tuneable parameters α_i, associated with the metric functions, metric weightings, and recommendation bins. To ensure that the action recommendation reflects the typical action that an experienced operator would normally take, ISG collaborated with the NASA CARA team to ensure that outputs reflect an intuitive, interpretable representation of conjunction information and provide advisory actions that correspond with those that would be taken by experienced operators. Operations in different companies may have different strategies when it comes to satellite management, so this process can be repeated with the same modelling framework for more 'bespoke' solutions.

As adjusting 17 independent settings is a high-dimensional problem, the space-filling technique of Latin hypercube sampling (LHS) is used. Compared to standard Monte Carlo sampling, LHS uses an orthogonal sampling technique that covers an input space more broadly and efficiently. This is illustrated in Fig. 8, where a sample is drawn from a bivariate distribution. Comparing the samples from 7a & 7b, we see that 7b tends to produce samples that are a greater distance from its closest neighbour.

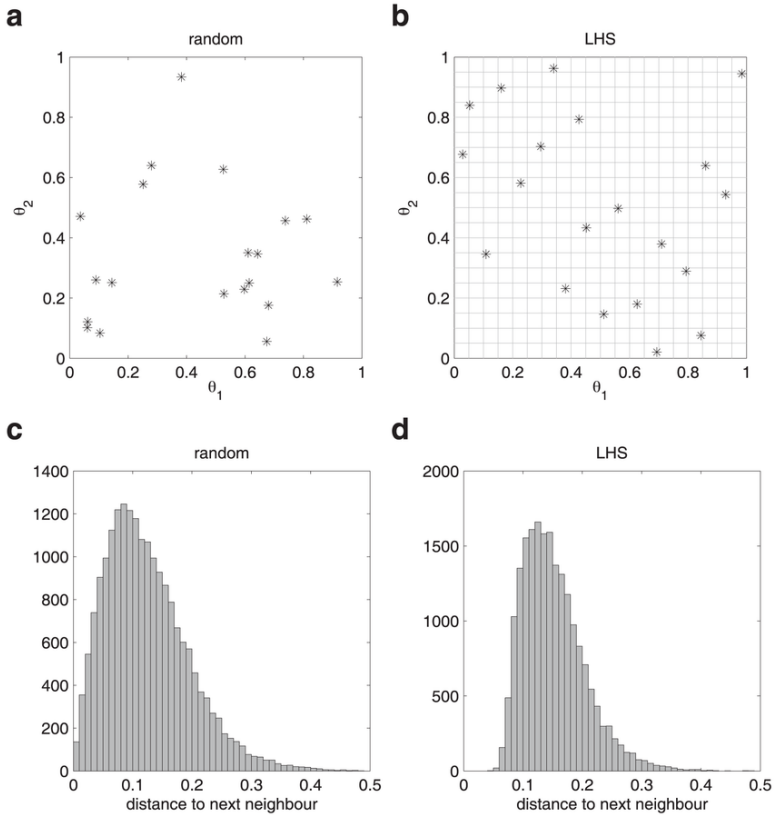

Fig. 8 Comparison of random sampling and Latin hypercube sampling (LHS) for the generation of initial parameter guesses [5].

Using this method, a comprehensive search and optimization of the best parameter settings could be performed for different time increments to TCA. The performance of the best combination of the 17 parameters is discussed in the results section.

2.2 Drop-off metric

The drop off metric can be derived from the uncertainty in the timeseries of Pc prediction data, as shown in Fig. 9. A cubic-spline interpolation is used to interpolate each rank of Pc predictions over time. As discussed previously, the Pc distribution is the result of the bootstrapping method that generates a range of miss distances and the covariance shrinkage. The collection of blue lines in the graph is a spread of individual Pc predictions based on a distribution of miss distance. For this metric, the distribution is ranked according to its magnitude. Each rank is then interpolated with the same rank of a different time prediction to form a function. This is illustrated by the individual blue trajectories in Fig. 9.

This sample was a prediction of the Pc at 3.5 days from TCA. By taking a vertical cross-sectional slice of the distribution at MCP, which is chosen to be at 0.5 days from TCA, a histogram such as in Fig. 10 is collected.

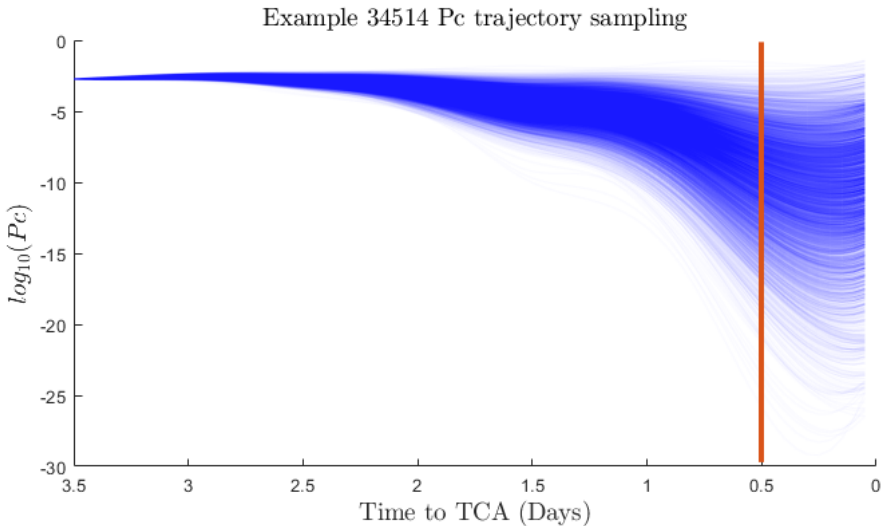

Fig. 9 *S*pline interpolation of the Pc prediction.

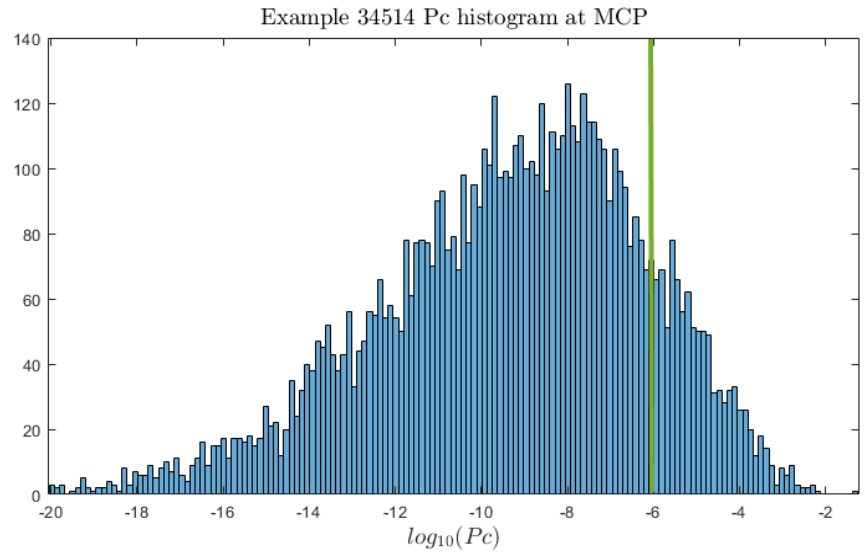

Fig. 10 Cross-section of the Pc distribution (red line indicating drop-off threshold).

The x-axis of the histogram represents the Pc value, and the y-axis counts the number of predicts that lie in a certain Pc range. With a selected drop-off threshold, summing all the counts of data points that are lower than the said threshold will give the percentage chance of seeing a drop off at MCP. For example, at this point, the drop off threshold is set to a value of 1×10^{-6}. To calculate the percent chance of drop off, the sum of the number of data points left of the green line is measured as a percentage of the total number of samples in the miss distance bootstrapping process. However, to reduce the number of false alarms and increase prediction accuracy, this drop-off threshold is set as a function of time to TCA.

2.3 Current Pc metric

The current Pc metric simply takes the current Pc value, contained in the current CDM dataset, and assigns it an associated urgency. Following discussion with CARA, it was determined that at different stages to TCA, the same value of Pc will raise a different level of concern to the operator. When the conjunction event is more than 3 days away, Pc values in the region of 3×10^{-5} may be of concern. Whereas approaching MCP, CARA operators tend to have a higher threshold of 1×10^{-4} as a maneuver planning threshold.

2.4 Maximum Pc metric

The Maximum predicted Pc metric works in a similar fashion to the current Pc metric. It examines the highest point prediction (mode/peak of the Pc distribution) from the predicted intervals. This metric is used because of the known shape of the Pc vs. time curve, where the Pc tends to increase then decrease over time when miss distance is static. If the Pc of an event goes up far into the red region ($>10^{-4}$), operators tend to pay more attention to the event, even if they think it is likely to drop off. This metric captures that potential rise in Pc, considering the geometry of the event and miss distance volatility. Note that even though the prediction of the tool is a distribution, this metric uses the median value of the Pc prediction interval at each prediction step.

2.5 Time to MCP metric

In general, as indicated by CARA operators, planning for a conjunction event takes place roughly 2-3 days before TCA. The Time to MCP metric is implemented to reduce the urgency of an event when it is a long period away from TCA. High-Pc events will still contribute to a relatively suitable course of action, but the Time to MCP metric allows for a smooth transition of the overall urgency value as TCA is approached.

3. RESULTS AND DISCUSSION

To assess the performance of the MDSS, an unseen set of the CDM database was used. Although the decision to maneuver is binary, the actions leading up to the decision point are not. Recall the 5 bin categories of recommended action: event dismissal (dark green), event monitoring (light green), preparatory mitigation action (yellow), substantial mitigation planning (light red) and mitigation execution (dark red).

To understand how well the tool performed at different times to TCA, the analysis segmented the CDMs into 5 time slots: 3, 2.5, 2, 1.5 and 1 days to TCA. Furthermore, each conjunction event (relating to all CDMs of the same conjunction) is labeled according to its Pc produced by the CDM immediately before the MCP. Although the conjunction behavior can still change after the MCP, it is irrelevant in a real-life scenario as you cannot operate via hindsight. These labels are green ($\text{Pc} < 1 \times 10^{-7}$), yellow ($1 \times 10^{-7} < \text{Pc} < 1 \times 10^{-4}$) and red ($\text{Pc} > 1 \times 10^{-4}$).

$$\text{Precision} = \frac{tp}{tp + fp}$$

$$\text{Recall} = \frac{tp}{tp + fn}$$

tp = true positive

tn = true negative

fp = false positive

fn = false negative

There are 3 prediction metrics that were used to quantify the performance: recall, precision, and accuracy. The formal mathematical definitions for precision and recall are shown above. To illustrate their definitions, example calculations are shown in Fig. 11 and 12, derived from a set of 2282 predictions at approximately 2 days before TCA. All the results will be shown in a similar format as the 2 tables, with the rows representing prediction category and the columns representing type of event (based on Pc at MCP). In Fig. 11, to calculate the prediction rate of light and dark red (event monitoring), recall would be calculated by performing a sum of the red column and dividing (84+2) by 115 = 75%. The purpose of recall is to answer the question “for a given type of event (green, yellow, red), what percentage of predictions is classified correctly?”. Alternatively, as demonstrated in Fig. 12, precision is determined by calculating the percentage of predictions that are correct. In the Fig. 12 Example the precision for the dark green prediction would be 12 divided by 12=100%

		Two-Day Prediction			
		Actual Pc at 0.5 Day MCP			
		Green	Yellow	Red	
Predictions	Predicted Dk Grn	12	0	0	12
	Predicted Lt Grn	1263	80	9	1352
	Predicted Yellow	272	152	20	444
	Predicted Lt Red	164	224	84	472
	Predicted Dk Red	0	0	2	2
	Total Events	1711	456	115	

Fig. 11 Recall calculation on red Pc events.

		Two-Day Prediction			
		Actual Pc at 0.5 Day MCP			
		Green	Yellow	Red	
Predictions	Predicted Dk Grn	12	0	0	12
	Predicted Lt Grn	1263	80	9	1352
	Predicted Yellow	272	152	20	444
	Predicted Lt Red	164	224	84	472
	Predicted Dk Red	0	0	2	2
	Total Events	1711	456	115	

Fig. 12 Precision calculation on red Pc events.

The MDSS tools capability at the 3 and 2 days epoch before TCA is demonstrated in Fig 13. As shown, when considering recall the tool is able to identify 74.7% of all red events to be of significance 2 days before TCA (the 2 green entries in the red Pc column) and 53.3% at the 3 day mark. The green-highlighted entries are combined, as they represent the type of predictions that operators prioritise. The top left 2 squares represent the green events that are correctly being categorised as low-importance events. The bottom right 2 squares represent events that ended with high Pc and are correctly identified. The advantage this provides is that not only do these CDMs indicate a high Pc, but by having higher urgency indicators (light red and dark red predictions) the tool provides greater confidence in decision making. Examining the green Pc event column, MDSS also demonstrates an ability to eliminate 65.5% and 74.5% of all green events from being prioritized, reducing unnecessary preparation.

		Three-Day Prediction			Two-Day Prediction		
		Actual Pc at 0.5 days MCP			Actual Pc at 0.5 days MCP		
		Green	Yellow	Red	Green	Yellow	Red
Recall	Predicted Dk Grn	0.0%	0.0%	0.0%	0.7%	0.0%	0.0%
	Predicted Lt Grn	65.5%	30.5%	6.7%	73.8%	17.5%	7.8%
	Predicted Yellow	25.6%	43.9%	40.0%	15.9%	33.3%	17.4%
	Predicted Lt Red	8.9%	25.6%	53.3%	9.6%	49.1%	73.0%
	Predicted Dk Red	0.0%	0.0%	0.0%	0.0%	0.0%	1.7%
Precision	Predicted Dk Grn	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%
	Predicted Lt Grn	90.1%	9.4%	0.6%	93.4%	5.9%	0.7%
	Predicted Yellow	67.5%	25.9%	6.7%	61.3%	34.2%	4.5%
	Predicted Lt Red	49.5%	31.8%	18.7%	34.7%	47.5%	17.8%
	Predicted Dk Red	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%

Fig. 13 Prediction results in recall and precision at 3 and 2 days before TCA.

When performing an analysis of precision, the tool is able to achieve a success rate of 90.1% and 93.4% for light green predictions. This improves operator confidence when using the tool, as whenever ‘event monitoring’ is advised they

can allocate less resources to that conjunction. At 3 days to TCA, we implemented a risk reduction mechanism to prevent false rejection early on during the event, effectively reducing the predicted dark red columns to 0%.

One of the goals during the development phase is to ensure that the possibility of the following 2 predictions is reduced:

- Prediction of mitigation execution (dark red) for an actual green event
- Prediction of event dismissal (dark green) for an actual red event

The first is defined as a type I error and the second as a type II error. Consulting the acceptable range of error rate with the CARA group, it was concluded that both rates would have to be close to ~0%, as they are the most disruptive to the whole operation. The Type I error would lead to significant attention placed on an event that was not dangerous (proven by the CDM at MCP). The worst case would be a type II error, where a dangerous conjunction was falsely dismissed too early. The consequence of the latter would lead to a catastrophic outcome if the two RSOs collided, resulting not only in the loss of satellites but also polluting the space environment. For this reason, the tool is tuned to safeguard against providing event dismissal advice unless the event can truly be dismissed. As shown in the Fig. 13 in the bottom left and top right of each table, the error entries are kept at the ideal 0.0% minimum.

		Two-Day Prediction			
		Actual Pc at 0.5 Day MCP			
		Green	Yellow	Red	
Predictions	Predicted Dk Grn	12	0	0	12
	Predicted Lt Grn	1263	80	9	1352
	Predicted Yellow	272	152	20	444
	Predicted Lt Red	164	224	84	472
	Predicted Dk Red	0	0	2	2
	Total Events	1711	456	115	2282

Fig. 14 Accuracy calculation

Furthermore, accuracy is defined as a percentage of the testing set that the tool predicted correctly in relation to the entire event test set. In this case, the definition for correct predictions before 2 days to TCA is defined as:

- For green Pc events: event dismissal and event monitoring (dark green and light green)
- For yellow Pc events: event monitoring, preparatory action and substantial mitigation planning (light green, yellow and dark red)
- For red Pc events: substantial mitigation planning and mitigation execution (light red and dark red)

An example calculation is illustrated in Fig. 14 by dividing the sum of entries within the red rectangles by the blue rectangle. Using the MDSS tool, a prediction accuracy of 70.9% and 85.5% at 3 and 2 days respectively is achieved. This level of accuracy indicates that the tool provides a high level of confidence for decision-making by separating high danger and insignificant events, as well as assisting with shifting the maneuver process earlier in the timeline. Outside of these >50000 events, there were only 3 cases where the Pc increased drastically at 1 day before TCA, resulting in a type II error. All three of these events were classified by experienced operators as events that could not have been predicted. These events were associated with secondary objects that were poorly tracked and did not receive sufficient tracking until very close to TCA.

One area of future work is to improve the precision performance of the light red category. Due to the imbalanced nature of the CDM database set, where most events are green and there are few red events in comparison, even a small proportion of green events predicted to be urgent will skew the precision metric. Another aspect that impacts the final output is the Pc calculation method, as method employed to determine Pc uses a combined hard body radius of 20m to perform the integration. This is a conservative method of determining Pc, as the integration encompasses a larger area than that of most RSOs and orbital debris. This conservative approach ensures an additional layer of safety against poor tracking; however, a higher-fidelity Pc calculation could lead to lower number of substantial mitigation planning predictions for green Pc event types.

4. CONCLUSION

The MDSS is a conjunction analysis tool that was successfully developed to assist operators in resource management and early maneuver planning. It introduces statistical rigor to satellite operations, regardless of experience, as part of a conjunction analysis routine. The output of the software is an urgency value that lies between the range of [0,1] and is binned into 1 of 5 predefined action recommendations. These predefined actions are calibrated to assist operators in making decisions, as part of an automated workflow, or as an added layer of analysis for high interest events that are more complex. The MDSS makes future Pc and miss distance predictions by utilizing a set of 13 Gaussian process models to predict time series covariance and miss distances distribution from a CDM epoch until TCA. Using the prediction results and combining it with additional CDM data, the tool calculates a weighted sum of 4 individual metrics: current Pc, maximum predicted Pc, percentage chance of Pc drop-off and time to MCP.

To convey the correct course of action for a given urgency, a series of workshops with NASA CARA were conducted to match the operators' intuition when considering past events. Using events from the CARA CDM database, the MDSS was able to achieve a prediction accuracy of 70.9% and 85.5% at 3 and 2 days before TCA respectively. Furthermore, the tool correctly predicted the suitable course of action 74.5% and 74.7% for green and red events, up to 2 days before TCA. Importantly, in a test set of over 2000 high interest events, the tool produced 0 false alarms for maneuver execution, as well as 0 events that were dismissed entirely but ended up with a $Pc > 1 \times 10^{-4}$ at MCP. This indicates that the utilization of this tool will reliably streamline operational resources and reduce the likelihood of a collision not being prevented. Thus, the MDSS tool can provide noticeable improvement to any SSA and satellite operations that seek to optimize their workflow and reduce resource expenditure. Furthermore, we anticipate that the MDSS will assist SSA operations in keeping pace with the growing number of conjunctions, and potentially introduce a standard practice for operators.

The MDSS is unique in the field of space traffic management as it combines physical and statistical models to provide operators with a rigorous recommended course of action well ahead of a maneuver commitment point. The tool is undergoing further refinement at CARA and by ISG to prepare for commercial deployment in later 2022.

5. REFERENCES

[1] ESA Space Debris Office. ESA's Annual Space Environment Report, *The European Space Agency*, 2022.

[2] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*, the MIT Press, 2006.

[3] R.B. Gramacy. *Surrogates*: *Gaussian Process Modeling, Design, and Optimization for The Applied Sciences.* Chapman and Hall/CRC, 2020.

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

[5] A. Raue, M. Schilling, J. Bachmann, A. Matteson, M. Schelke, D. Kaschek, et. al. Lessons Learned from *Quantitative Dynamical Modeling in Systems Biology*, PLoS ONE 8(9): e74335, 2013.