Validation of a Scenario-Based Approach to Assess Gaps in Earth Observations

Katharine Burn
Bailey Ethridge
Marie Ivanco
Systems Analysis and Concepts Directorate
NASA Langley Research Center
5 Langley Blvd.
Hampton, VA 23681
katharine.z.burn@nasa.gov
bailey.e.ethridge@nasa.gov
marie.l.ivanco@nasa.gov

Abstract—**With the increased reliance on spaceborne Earth observation data among the Earth science community and other end users, it is important that efforts are made to promote data continuity for a range of parameters of interest. Continuity gaps may occur between missions measuring like parameters due to mission development delays or early termination and introduce the potential of increased uncertainty for retrieved parameters. To inform portfolio-level decisions for Earth observing missions, Ivanco et al. developed a method that enables the assessment of the probability of continuity gaps for multi-mission architectures and provides a framework to assess this probability in the context of multiple scenarios that represent possible future states of the architecture [Ivanco et al., "A Scenario-Based Approach to Assess Continuity Gaps in Earth Observations," IEEE Aerospace Conf., 2024].**

While this method was previously applied to assess continuity gaps for a specific multi-mission architecture, it had not yet been validated with historical data. This paper builds on the work of Ivanco et al. by utilizing data obtained from past NASA Earth science mission formulation documents to retroactively estimate mission timelines and gap probabilities for a selected multi-mission architecture. The model results are compared with the true outcomes of the missions in an attempt to validate the probabilistic gap assessment method. Discussion in this paper is limited to the probabilistic analysis portion of the method, which estimates the likelihood of gaps for a baseline mission architecture measuring a specific parameter of interest. The latter steps of the method assess the impact of decisions on measurement continuity, given the estimated gap probabilities, by use of scenario analysis and are therefore out of scope for this validation effort.

The efforts to validate the stochastic portion of the scenario-based method are not an exhaustive validation for all use cases of the method, but rather a case study of a singular architecture comparable to the architecture defined in Ivanco et al.'s analysis. The validation case study concludes that the method adequately identifies gap potential for a case where a gap was not planned in the mission architecture. A sensitivity analysis is performed to evaluate the effect of the epistemic uncertainty of input distributions on the model and demonstrates the impact of variations of input distributions related to missions in development. The validation and sensitivity analysis efforts inform future modeling improvements which will impact the effectiveness of subsequent applications of the method.

This paper first summarizes the method used to assess the probability of gaps in multi-mission architectures. It then describes the analysis performed by the authors to assess the validity of the method and quantify its predictive accuracy and presents the results of the sensitivity analysis. Lastly, insights gained from the validation efforts as well as proposed improvements pertaining to model input parameters are discussed.

TABLE OF CONTENTS

1. INTRODUCTION

Many Earth science researchers, applied scientists, and other end users, such as private companies, globally utilize spaceborne Earth observation data for monitoring and researching the Earth system. If gaps are introduced between missions in Earth observing multi-mission architectures, the accuracy of predictions and analysis could be diminished as the continuous record of measurements may be compromised. With operational mission lifetimes varying from one year to extended operations of 10 years or more, the probability of data continuity gaps needs to be assessed well in advance of their occurrence to support strategic decision-making. Decision-makers for NASA's Earth Science Division (ESD) face significant uncertainty in this process and must consider many factors including funding, technology maturation, and schedule shifts, while considering many potential outcomes that are often difficult to predict [1].

The scenario-based method described by Ivanco et al. was developed to estimate the probability of potential continuity gaps—defined in this context as any length of time during which there are no operational missions collecting a particular type of data—and to assess this probability in the context of multiple possible future states of the architecture [2]. Spaceborne multi-mission architectures often constitute systems of systems (SoS), defined as “a set of systems or system elements that interact to provide a unique capability that none of the constituent systems can accomplish on its own" [3]. Because of the complexities associated with the SoS nature of multi-mission architectures, the scenario-based method departs from NASA’s traditional systems engineering decision analysis process [4] and instead incorporates selected ideas from the Military Decision-Making Process (MDMP) [5].

This paper builds on the work of Ivanco et al. by describing validation efforts performed to assess the predictive accuracy of the stochastic portion of the scenario-based method. The intent is not to perform an exhaustive validation of all use cases of the method. Rather, the study is an initial effort investigating the validity and adequacy of the method, by providing insight into the effectiveness of the distributions and models used to assess the probability of gaps in a multi-mission architecture.

To validate the method, data are obtained from past NASA Earth science mission formulation documents to estimate mission timelines and gap probabilities according to the state of knowledge during mission formulation. The model results are compared with the true outcomes of the missions to assess the predictive capabilities of the model. A sensitivity analysis is also performed to further examine how variations in input parameters affect model outputs. The paper first summarizes the method used to assess the probability of gaps in multi-mission architectures. It then describes validation efforts and their results and presents the results of the sensitivity analysis. Lastly, insights gained from the work and proposed improvements pertaining to model input parameters are discussed.

2. CONTINUITY ASSESSMENT METHOD

Discussion in this paper will be limited to the first part of the method developed by Ivanco et al. to focus on the validation of the probabilistic analysis, which estimates the likelihood of a gap for a baseline scenario consisting of missions measuring a parameter of interest and their respective assumed timelines. The latter steps of the method assess the impact of decisions on measurement continuity, given the estimated gap probabilities by use of scenario analysis, and are therefore out of scope for the validation effort.

The first step of the stochastic method involves defining a set of adverse events that could cause a gap in continuity. For a mission in formulation, there are three types of adverse events that may impact continuity:

1) Delayed mission launch
2) Unsuccessful mission start
3) Premature mission termination

A mission launch is considered delayed when the launch occurs on a later date than nominally scheduled. In this framework, initiating events resulting in launch delay constitute events that result in delay of launch readiness date, such as technology development and supply chain delays, as well as delays attributed to launch vehicles and launch scheduling, such as weather delays. A mission start is considered unsuccessful if there is a loss of vehicle at launch, unsuccessful orbit insertion, or failure of the satellite or sub-systems to retrieve measurements that satisfy science requirements. A mission is considered to end prematurely if the mission’s true lifetime is less than the expected design lifetime. This could be caused by technical failures, accelerated orbit decay, or damage caused by space debris. For operational missions, the first two outcomes do not apply as the mission success status and launch date are known and do not need to be modeled probabilistically; only the premature mission end should be considered.

To estimate the probability of adverse events occurring and the resulting impact on the mission timeline, we must be able to identify the comprehensive set of initiating events that may trigger the adverse event. In the Probabilistic Risk Assessment Procedures Guide for NASA Managers and Practitioners, initiating events are defined as events that “perturb the system (i.e., cause it to change its operating state or configuration), representing a deviation in the desired system operation” [6]. Adverse events are decomposed into one or more failure types, then failure causes, and finally initiating events. Initiating events are documented as nodes in the lowest tier of the decomposition. A notional diagram of the decomposition is shown in Figure 1.

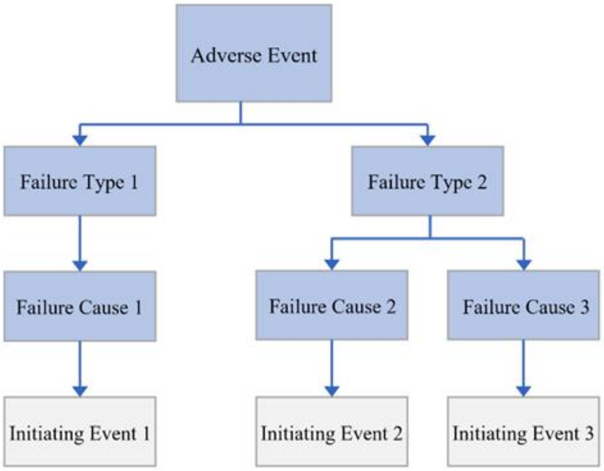

Figure 1. Notional diagram of the hierarchical decomposition of adverse events

Once adverse events have been identified and modeled, the probabilities of the occurrence of each initiating event are quantified when feasible. For initiating events causing a shift in mission timelines, the potential impact of the event in terms of launch delay duration and mission lifetime deficit are modeled as distributions. Technology Readiness Level (TRL) of sub-systems and technology heritage inform the selection of distributions when relevant. In cases where data are unavailable to accurately parameterize a node, the nodes are omitted from the assessment and documented in the assumptions.

The probability of a gap over time is then assessed by use of Monte Carlo simulation. The simulation generates probability distributions for mission start and end dates, resulting in modified mission timelines for the cases of delayed launch or premature mission termination. For cases of an unsuccessful mission start, the simulation categorizes the result as a mission cancelation. A manifest is then constructed as the aggregate of each mission's simulated outcome.

Performing N Monte Carlo runs results in N potential manifests, from which the probability of a gap in a given year, j, can be calculated as

$$P\left(Gap_j\right) = \frac{\sum_{i=1}^{N} I\left(m_{ij}\right)}{N}, \qquad I\left(m_{ij}\right) = \begin{cases} 1, m_{ij} = 0 \\ 0, m_{ij} \neq 0 \end{cases} \qquad (1)$$

where m_{ij} is the number of active missions in year j predicted by simulation run i and $I(m_{ij})$ is an indicator variable which is equal to one when m_{ij} is zero and zero otherwise. The current study assesses gaps for a mission architecture with a single parameter of interest. For more complex studies with multiple parameters of interest, Eq. 1 would require an additional index to represent assessing the likelihood of gaps in each individual parameter.

3. VALIDATION

Verification and validation processes "quantify the confidence and predictive accuracy of model calculations" [7] and are essential parts of the modeling process. Results obtained from verification and validation efforts inform decision-makers of the predictive ability of the model used to guide decisions.

According to the AIAA "Guide for the Verification and Validation of Computational Fluid Dynamics Simulations," verification is "the process of determining that a model implementation accurately represents the developer's conceptual description of the model and the solution to the model" [8]. There are multiple approaches to perform model verification depending on the nature of the model. A common approach is to use the model to compute an analytical problem with a known solution; the output of the model is then compared to the known solution to verify the accuracy of the model and ensure that it was implemented correctly. Validation is defined as "the process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model" [8].

Naylor and Finger developed a multi-stage approach to validation that is routinely cited in relevant literature [9], [10], [11], [12]. The first stage of the approach involves building a model with high face validity. Analysts select variables and relationships to represent the system of interest and subsequently assess the ability of the designed model to represent the real-world system. Observations and expert judgement are used to perform the assessment. The second stage of the approach entails validating model assumptions, which are identified and assessed. Assumptions are empirically tested where possible. The third and final stage of the approach ensures that the model's output data have "sufficient accuracy for the model's intended purpose" [12]. In this stage, analysts assess the predictive accuracy of the model and whether the achieved level of accuracy is adequate for the intended use of the model.

This paper focuses on the third stage of Naylor and Finger's validation approach as it was of particular interest to the authors to gain insights into the ability of the model to accurately estimate mission timelines and predict the likelihood of gaps to occur in multi-mission architectures. The authors therefore focused their efforts on comparing the output of the simulation with realized historical outcomes to characterize the predictive accuracy of the model and inform future modeling improvements.

This work was motivated by the relevance of examining epistemic uncertainty of the system of interest. Multiple data sources inform node parametrization and the design of the mathematical model. While historical data and quantitative analysis were favored to inform the selection of distributions, select nodes were informed by expert opinion found in mission review documents. The model must consider factors such as sub-system heritage, anticipated technology developments, and assessed supply chain risk. These factors can be contributors to initiating events that might cause adverse events and impact the likelihood of gaps between missions. However, available data to assess these factors for specific missions and inform the parametrization of nodes tend to be qualitative in nature. Examples include the following:

- Sub-systems can be described as having "significant heritage."
- Experts can estimate the number of years needed to reach the next level in the TRL scale.
- Confidence in a selected approach may be stated.
- Risk associated with the supply chain can be assessed as high.

In addition to the inherent subjectivity of expert opinion, the use of text-based sources introduces challenges to adequately represent uncertainty [13]. Analysts interpret text-based, qualitative statements to inform mathematical models. This introduces additional uncertainty in the model as interpretation and design decisions may vary among analysts. Additional challenges are also associated with potentially missing information. Nodes were parametrized when data were available; however, select nodes identified as potential initiating events could not be parametrized due to a lack of data.

These challenges motivated this study and the authors' interest in comparing model output with the true outcomes of past missions to assess the validity of the probabilistic gap assessment approach.

Validation Approach

To implement stage 3 of the validation approach, data obtained from past NASA Earth science mission formulation documents are used to generate distributions of predicted mission start and end dates. The outputs of the simulation are compared to the realized launch and end of life (EOL) dates of the past missions.

When setting up the validation, the model must be constructed under the same assumptions that would have been used for the nominal assessment of missions in operation and formulation. To simulate this environment, a historical set of two missions measuring the same parameter of interest is selected to mimic the case of one mission being in operation while the second is in formulation. A notional timeline of these missions is shown in Figure 2.

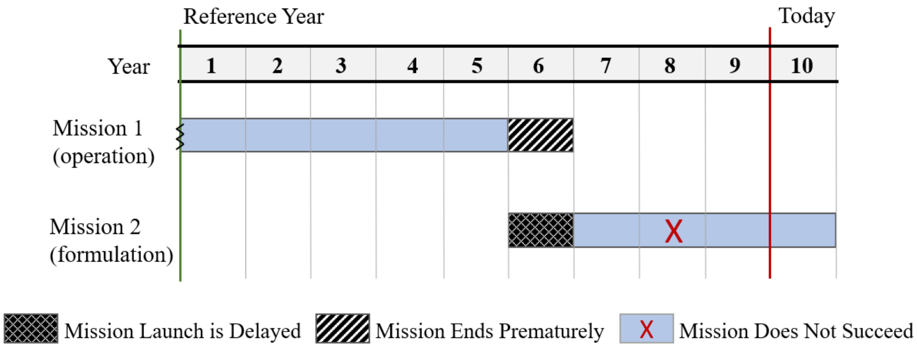

Figure 2. Notional mission timeline used for validation

Documents created in a particular year during Mission 2 formulation, also termed the "reference year," are obtained for both missions. From these records, information is gathered to inform the hierarchical decompositions built for the adverse events based on the state of knowledge at the time of mission formulation. It is important to ensure the knowledge gathered during data collection efforts and used in the validation analysis is traced to the state of knowledge of the reference year, and no insight about the true outcomes of the mission is used in the event decomposition.

Three adverse events are identified for the two missions included in the analysis: Mission 1 ending prematurely, Mission 2 being delayed, and Mission 2 failing to succeed in obtaining measurements. The result of the stochastic analysis will be an estimation of the EOL for Mission 1, the launch date for Mission 2, and the probability of Mission 2 failing to succeed. From these results, the probability of a continuity gap between the two missions can be derived.

Probabilistic distributions are derived for each initiating event node relevant to the mission of interest. The majority of the nodes that were selected to be omitted from the validation analysis are consistent with the original analysis performed by Ivanco et al. Any differences in the nodes used between the two studies is due to the different instrumentation on the validation missions compared to the original analysis. Certain nodes related to tech development instrumentation are turned on in the validation analysis to account for this difference and accurately represent the systems with increased risk in the mission architecture.

Distributions used in the original analysis are reassessed for the validation analysis to account for changes in the availability of data and individual system risks. Since this effort is targeting validation of the method and not the specific datasets used in the original analysis, historical data is filtered to only include records available prior to the reference year to more closely mimic the results of the method applied during mission formulation.

Four types of distributions are used to model initiating events. For events such as orbital debris collision, internal analysis is leveraged to estimate a rate of occurrence which is used to parameterize an exponential distribution giving the time to event. For other low probability events such as instrument failures, historical data are used to construct empirical distributions for the lifetime deficit and delay durations resulting from those events. Triangular distributions used by NASA to estimate cost and schedule growth are used to model cost- and schedule-related nodes respective to system risk levels. Finally, events resulting in mission cancellation are represented with single-point probabilities derived from historical data.

After defining the probabilities and distributions of the initiating events, the Monte Carlo simulation is run to estimate distributions for the occurrence of the three adverse events (mission cancellation, mission delay, or early termination). From these distributions, estimates for each mission's start and end date are generated. These predictions serve as a proxy for what could have been predicted for mission lifetimes during the formulation of Mission 2, and the results are compared to the dates of the realized mission.

Comparison of Monte Carlo Outputs with Realized Outcomes

Distributions of the Mission 1 EOL and Mission 2 launch dates estimated by the Monte Carlo simulation are plotted in Figure 3. The distribution of EOL dates for Mission 1 is wide and highly negatively skewed, with a standard deviation of 228 days and a skewness coefficient of -5.70. Low outliers—ranging to almost six years prior to the true EOL—comprise 8.45% of the points. In contrast, there is a sharp cutoff on the right side of the distribution. This is due to extended mission lifetimes not being considered in the model based on the assumptions made in the original assessment. Aside from the truncated righthand side, the shape and location of the distribution indicates realistic estimates produced by the simulation. The mean EOL estimate is found to be within two days of the true EOL date. Additionally, the probability of failure is non-zero at any point from the analysis start date, which is reflected by the long left tail, and becomes more significant closer to the nominal EOL.

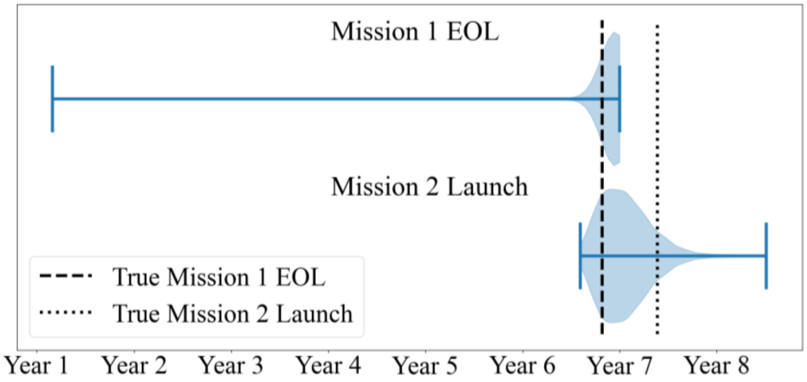

Figure 3. Probability density of Monte Carlo outputs for Mission 1 EOL and Mission 2 launch date values

The Mission 2 launch dates plotted in Figure 3 have a narrower, positively skewed distribution, with a standard deviation of 102 days and a skewness coefficient of 1.14. The mean simulated launch date underestimated the true launch delay by 117 days, suggesting that the input distributions for launch delay may have been too narrow. However, the true launch date is within the predicted distribution, falling within 1.15 standard deviations of the mean, with 10.35% of simulation runs falling after the true date. A prediction error of four months is relatively small compared to the remaining six years of formulation, and the magnitude of the predicted delay (177 days) is indicative of the true delay (294 days), which informs decision-makers of gap potential.

These distributions should not be considered independently, as each simulation run generated an EOL and launch date pair where overlap between these two dates meant that continuity was preserved. Figure 4 shows these points with respect to the true Mission 1 EOL, represented by the vertical dashed line, and Mission 2 launch date, represented by the horizontal dashed line. The color saturation of each point corresponds to its observed frequency, and the red X's spanning the bottom of the plot signify runs in which Mission 2 experienced a cancelation.

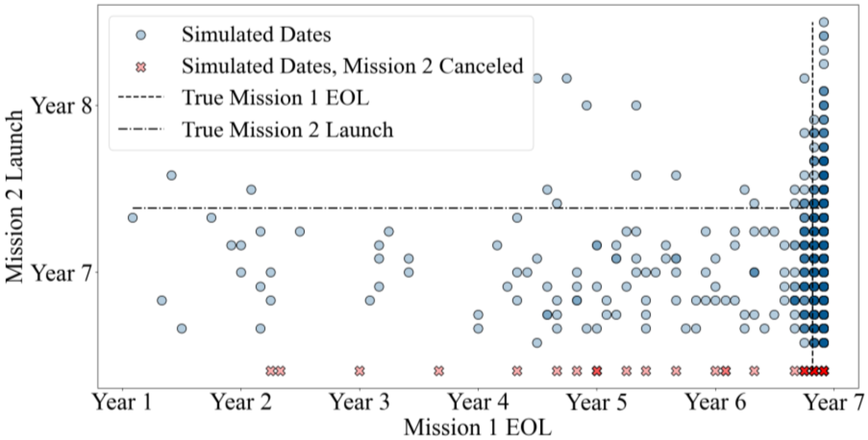

Figure 4. Monte Carlo results for Mission 1 end of life and Mission 2 launch plotted against the true EOL and launch dates

Simulation outputs are concentrated in the bottom-right quadrant of Figure 4, indicating that the model tended to overestimate the lifetime for Mission 1 and underestimate the delay for Mission 2. Two thirds (66.6%) of all points fall into this quadrant; these points overestimate Mission 1 EOL by an average of 48.4 days and underestimate Mission 2 launch delay by an average of 145 days. A higher concentration of outputs in the bottom-right quadrant could lead to a less conservative characterization of potential gaps and a greater risk of false negatives where a true gap is not predicted. However, the magnitude of these discrepancies is acceptable, especially when interpreted within the context of multi-year mission durations. Of all simulation runs, 54.4% are within 6 months of both the Mission 1 EOL and the Mission 2 launch date, and 80.0% of runs fall within one year of the true dates.

Comparison of Planned, Actual, and Simulated Mission Counts

Figure 5 compares the count of operational missions over time for the planned outcome, the outcome predicted by the stochastic model, and the true outcome of the missions. The simulated mission counts are calculated as the mean number of active missions at each point in time across all simulation runs. From year 1 to the end of year 6, the simulated mission count drops from 1 to 0.94 as an increasing number of simulation runs predict EOL of Mission 1. Once the nominal Mission 2 launch date is reached, the expected mission count rises to a peak of 1.079 with the addition of the second mission, then drops significantly at the presumed EOL of Mission 1 in year 7. Following Mission 1's end, the expected count increases as Mission 2 is launched in an increasing number of simulation runs until it finally converges at 0.84, the percentage of simulations in which Mission 2 was not canceled. The simulated mission counts show a potential area of continuity risk between Mission 1 and Mission 2.

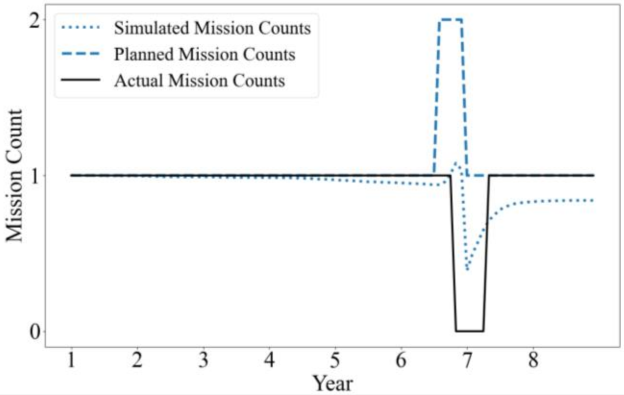

Figure 5. Comparison of mission counts for planned dates, stochastic analysis predictions, and true dates

As shown in Figure 5, the stochastic model more closely predicts the true outcomes of the missions than what was initially planned during formulation. The purpose of the stochastic analysis is to identify areas in a mission architecture that have a risk of a gap. This plot shows that the stochastic model accurately predicts where the gap would occur between the missions and demonstrates that the intended purpose of the model is met.

It is important to note that this validation effort only applies the continuity assessment method to a single historical mission architecture. While the results documented here suggest adequacy of the model to predict mission timelines and likelihood of gaps for this test case, more testing is necessary to truly indicate the adequacy of the model for every situation.

4. SENSITIVITY ANALYSIS

A sensitivity analysis is performed to evaluate the effect of epistemic uncertainty on the model results. Epistemic uncertainty, according to Oberkampf et al., "derives from some level of ignorance, or incomplete information, of the system or the surrounding environment." [14] Epistemic uncertainty is highly relevant to this model due to the inherent lack of knowledge available to inform the selection of distributions used as inputs to the model. The model addresses aleatory uncertainty, or the "inherent variation associated with the physical system or the environment under consideration," [14] by using probabilistic distributions for the input parameters. The goal of the sensitivity analysis is to observe how adjustments to subjective input parameters affect model results.

First, a sensitivity analysis is performed for a continuous distribution of schedule delay based on the selection of a qualitative risk level. An internal NASA cost and schedule risk analysis guide defines triangular distributions for nine levels of risk varying from "Low" to "Extra High." A change in risk level selection changes the 10% and 90% points of the triangular distribution, which is used as an input to the continuity assessment model for select nodes. Due to the subjectivity of selecting a risk level based on qualitative information, there can be variations in the input distributions selected by any given analyst. To study the sensitivity of the selected input, the distribution is both increased and decreased by a single level of risk according to the NASA-provided risk chart, and the effects of the associated triangular distribution changes are assessed.

The results of the sensitivity analysis are visualized by plotting simulated mission counts for different risk cases against each other. Figure 6 shows a comparison of the mission counts for the case studies of increased and decreased schedule risk for Mission 2. The figure depicts the dataset for the base case compared to both cases of increased and decreased schedule risk level. Table 1 provides the minimum and maximum simulated mission count values for each case study. When comparing the minimum simulated mission count values for each case, there is a 20.1% decrease in simulated mission count from the base case to the case of increased risk, meaning an increased likelihood of a gap between the missions. There is a 46.2% increase in minimum simulated mission count values from the base case to the case of reduced risk. There is also an increased likelihood of mission overlap, as shown by the 16.7% increase in the maximum simulated mission count values from the base case to the case of reduced risk. The larger percent change for the case of reduced schedule delay risk suggests that underestimating the risk of the system could have a larger impact on the model results than adding additional risk margin to the inputs.

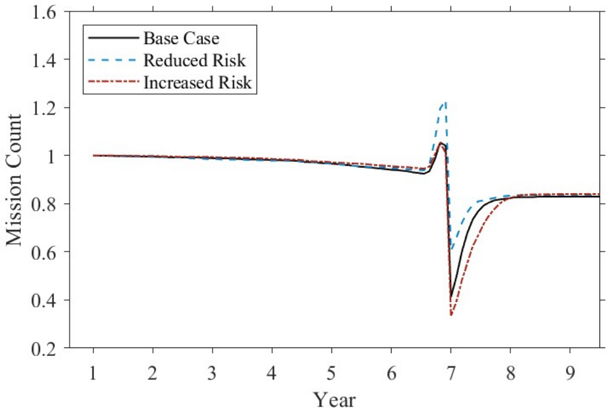

Figure 6. Comparison of simulated mission counts for three case studies of schedule uncertainty inputs

Table 1. Minimum and maximum simulated mission count values for cases of adjusted schedule uncertainty risk, plotted in Figure 6

	Min Simulated Mission Count	*Max Simulated Mission Count*
Base Case	0.41	1.05
Increase in Risk Level	0.33	1.05
Decrease in Risk Level	0.60	1.23

Some nodes model the likelihood of mission delay due to the development of technologies for the mission of interest. To observe the change in model results when omitting a node addressing instrument development, the second case study of the sensitivity analysis omits a node for a new instrument technology that was originally included in the validation analysis. Figure 7 shows a 62.8% increase in the minimum value of simulated mission counts when this new technology node is not included, and a 23.5% increase in the maximum value relative to the base case. The minimum and maximum simulated mission count values for this case study are listed in Table 2.

In the study by Ivanco et al. as well as this validation study, the parameterization of heritage technology nodes was omitted from the analysis. This was motivated by the quick turn-around time required to complete the analysis, which prompted the authors to focus on the parametrization of nodes representative of the development and production of new technologies. Adverse events are more likely to occur due to new technologies than technologies with spaceflight heritage. While the level of risk associated with a heritage technology and a new technology differ, the sensitivity analysis performed for the selected node suggests that the omission of heritage nodes in the original study should be revisited to determine the impact of the omission of low-probability events on the model results.

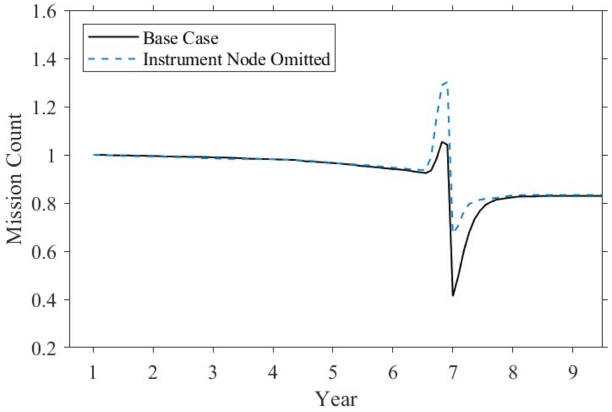

Figure 7. Comparison of simulated mission counts for the case of omitting a technical development node from the analysis

Table 2. Minimum and maximum simulated mission count values for the case of omitting a technical development node, plotted in Figure 7

	Min Simulated Mission Count	*Max Simulated Mission Count*
Base Case	0.41	1.05
Omitting Tech. Dev. Node	0.67	1.30

The plotting of simulated mission counts shows that adjustments to the model's input distributions related to the delay of Mission 2 have an impact on the extent to which a gap can be expected due to mission development delays. This is observed by the width of the distributions of minimum mission count plotted for each case study. The adjustment of input parameters pertaining to Mission 2 development does not largely influence the placement of the highest identified gap likelihood, but it does impact the length of the gaps that are predicted for the mission architecture of interest. This is indicative of the model's sensitivity to variations in input parameters associated with missions in development to accurately estimate the likelihood of a gap in a given year due to development delays.

5. PROPOSED MODEL IMPROVEMENTS

The validation and sensitivity analysis studies identified areas of improvement that should be revisited for future applications of the continuity assessment method. The first use of the method, documented by Ivanco et al. [2], did not consider extended mission lifetimes, which results in the truncation of the right side of the probability density plot of the Monte Carlo outputs. Introducing the ability to model extended mission lifetimes is planned for future implementations of the model. This change will improve the model's predictive accuracy by including the probability that missions may operate past their expected EOL date.

The current model predicts gaps based on the operational availability of a mission in a given year. Future model improvements will provide higher fidelity analysis by assessing the relative ability of missions to address a given science parameter.

The validation study, as well as the study by Ivanco et al., focused on nodes that were known drivers affecting mission timelines. Nodes representative of the development and production of new technologies were parameterized, while nodes representative of heritage technologies were not. The sensitivity analysis performed for a non-parametrized node demonstrated the impact of the omission of new technologies on the predicted likelihood of a gap in a mission architecture. For future studies, it is recommended to quantify the impact of not including heritage technologies. Including the development and production risk for heritage technologies would account for all causes of initiating events, even those with low probability.

System redundancy should also be considered in future model revisions. While system redundancy is considered to determine risk levels used to select probability distributions, the failure of a redundant system that is not mission ending is currently not considered. Future model revisions should include the capability to simulate the effects of non-mission ending failures to a multi-mission architecture.

6. CONCLUSION

A method was introduced by Ivanco et al. [2] to characterize the probability of continuity gaps for multi-mission Earth observing architectures and to assess this probability in the context of multiple scenarios that represent possible future states of the architecture. In this paper, the authors use data obtained from past NASA Earth science mission formulation documents to retroactively estimate mission timelines and gap probabilities for a selected multi-mission architecture. The model results are then compared with the true outcomes of the missions.

The efforts to validate the stochastic portion of the scenario-based method are not an exhaustive validation for all use cases of the method, but rather a case study of a singular architecture comparable to the architecture defined in Ivanco et al.'s analysis. The intent is to provide insight into the effectiveness of the distributions and models used to generate mission timeline estimates and inform potential future model improvements.

The purpose of the stochastic analysis is to identify areas of increased likelihood of gaps between missions in a multi-mission architecture. The shape and location of the probability density of the Monte Carlo outputs indicate the estimates produced by the simulation are adequate in predicting mission timelines and likelihood of gaps for the selected mission architecture. While planned mission counts during mission formulation anticipated an overlap between Missions 1 and 2, the simulated counts indicate there is likelihood of a gap between the two missions, mimicking the actual events of the mission architecture.

The mean EOL estimate for Mission 1 is found to be within two days of the true EOL date. The mean simulated launch date for Mission 2 underestimates the true launch delay by 117 days. This suggests that the input distributions for launch delay may need to be revisited; however, a prediction error of four months is relatively small compared to a six-year mission formulation period and is within the model uncertainty. The 177-day predicted mission delay is indicative of the 294-day true delay, which informs decision-makers of gap potential, meeting the overarching goals of the method.

The sensitivity analysis examines the impact of risk levels selected to inform the parametrization of cost and schedule delay nodes. It demonstrates that underestimating the risk associated with the system could have a larger impact than increasing the risk margin of the model inputs. The sensitivity analysis also examines the impact of not parameterizing nodes that are driving initiating events. The results showcase that the omission of driving factors impacts the model results, and it is recommended that future studies further examine the impact of not parameterizing nodes that are low probability and are not considered to be driving initiating events, as their omission may also cause an underestimation of the likelihood of gaps to occur.

The sensitivity analysis also demonstrates that adjusting the input distributions associated with Mission 2 launch delay has a greater effect on the magnitude of the likelihood of a gap in the mission architecture than on the years in which a gap is expected to occur. This means that variations in distributions for missions in development affect the predicted length of a gap but do not affect the model's overall ability to predict gap occurrence in a mission architecture.

Planned model improvements include increased modeling capability to assess the probability of extended mission lifetimes and the inclusion of the relative ability of missions to address a given science parameter. Additional work could include the addition of modeling low-probability events in the Monte Carlo simulation and modeling of the failure of redundant systems in non-mission ending failure cases.

REFERENCES

[1] National Academies of Sciences, Engineering, and Medicine, "Continuity of NASA Earth Observations from Space: A Value Framework," Washington, DC: The National Academies Press. 2015. https://doi.org/10.17226/21789.

[2] M. Ivanco et al., "A Scenario-Based Approach to Assess Continuity Gaps in Earth Observations," IEEE Aerospace Conference, Mar. 2024.

[3] *Systems and software engineering – System of systems (SoS) considerations in life cycle stages of a system*, International Standard 21839, 2019-07: ISO/IEC/IEEE 21839:2019.

[4] *NASA Systems Engineering Handbook*, National Aeronautics and Space Administration, Washington, DC, USA, 2020. Available: https://www.nasa.gov/sites/default/files/atoms/files/nasa_systems_engineering_handbook_0.pdf

[5] Military Decision-Making Process, Center for Army Lessons Learned, Handbook No. 15-06. Available: https://usacac.army.mil/sites/default/files/publications/15-06_0.pdf

[6] M. Stamatelatos, et al. "Probabilistic Risk Assessment Procedures Guide for NASA Managers and Practitioners," 2011, NASA/SP-2011-3421.

[7] B. H. Thacker, et al. *Concepts of Model Verification and Validation*, Los Alamos National Laboratory, Los Alamos, New Mexico, 2004.

[8] American Institute of Aeronautics and Astronautics, *Guide for the Verification and Validation of Computational Fluid Dynamics Simulations*, AIAA-G-077-1998, Reston, VA, 1998.

[9] T. H. Naylor, J. M. Finger, "Verification of Computer Simulation Models," *Management Science* 14 (2): B92-B101, 1967.

[10] R. G. Sargent, "Verification and Validation of Simulation Models," in *Proceedings of the 2010 winter simulation conference*. IEEE, 2010 pp. 166-183.

[11] X. Xiang, et al., "Verification and Validation of Agent-based Scientific Simulation Models," in *Agent-directed simulation conference*, vol. 47. San Diego, CA, USA: The Society for Modeling and Simulation International, 2005, pp. 47-55.

[12] N. Nayani, M. Mollaghasemi, "Validation and Verification of the Simulation Model of a Photolithography Process in Semiconductor Manufacturing," in *1998 Winter Simulation Conference*, vol. 2. 1998, pp. 1017-1022.

[13] A. D. Nesmith, "Quantitative Modeling of Text-Based Intelligence Source Uncertainty," M. S. thesis, Dept. of Physics, Air Force Institute of Technology, Wright-Patterson Air Force Base, OH, 2023.

[14] W. L. Oberkampf, et al., "Challenge problems: uncertainty in system response given uncertain parameters," in *Reliability Engineering & System Safety*, vol. 85, no. 1-3, pp. 11-19, 2004.

BIOGRAPHY

Katharine Burn *received a B.S. in Mechanical Engineering from Utah State University in 2019 and began her career as a propulsion engineer at Northrop Grumman Space Systems. She then received an M.S. in Aeronautics and Astronautics from Purdue University in 2022 with a focus in aerospace systems analysis. She now works as an aerospace engineer and systems analyst in the Systems Analysis and Concepts Directorate at NASA Langley Research Center, focusing her work on portfolio-level analysis of NASA's Earth science missions.*

Bailey Ethridge *serves as a mission analyst within NASA Langley Research Center's Systems Analysis and Concepts Directorate, where she develops analytical tools and methodologies to support strategic decision-making for both Earth science and human spaceflight mission applications. Bailey received B.A.s in cultural anthropology and applied mathematics from the University of North Carolina at Asheville and is currently pursuing her M.S. in computational modeling and simulation engineering at Old Dominion University.*

Marie Ivanco *is an Aerospace Engineer working in the Systems Analysis and Concepts Directorate of NASA Langley Research Center. Her work encompasses many aspects of space mission design and analysis, with a special focus on Decision Science applied to complex aeronautics projects. Mrs. Ivanco received her B.S. and M.S. degrees in Mechanical Engineering from Old Dominion University. Mrs. Ivanco also holds a B.A. in Political Science from Institut d'Etudes Politiques, Lyon, France and a M.A. from La Sorbonne Law School, Paris, France.*