An Integrated Reliability and Physics-based Risk Modeling Approach for Assessing Human Spaceflight Systems

Susie Go\textsuperscript{*a}, Donovan Mathias\textsuperscript{a}, Chris Mattenberger\textsuperscript{b}, Scott Lawrence\textsuperscript{a}, and Ken Gee\textsuperscript{a}
\textsuperscript{a}NASA Ames Research Center, Moffett Field, CA, USA
\textsuperscript{b}Science and Technology Corp., Moffett Field, CA, USA

Abstract: This paper presents an integrated reliability and physics-based risk modeling approach for assessing human spaceflight systems. The approach is demonstrated using an example, end-to-end risk assessment of a generic crewed space transportation system during a reference mission to the International Space Station. The behavior of the system is modeled using analysis techniques from multiple disciplines in order to properly capture the dynamic time- and state- dependent consequences of failures encountered in different mission phases. We discuss how to combine traditional reliability analyses with Monte Carlo simulation methods and physics-based engineering models to produce loss-of-mission and loss-of-crew risk estimates supporting risk-based decision-making and requirement verification. This approach facilitates risk-informed design by providing more realistic representation of system failures and interactions; identifying key risk-driving sensitivities, dependencies, and assumptions; and tracking multiple figures of merit within a single, responsive assessment framework that can readily incorporate evolving design information throughout system development.

Keywords: PRA, simulation, physical modeling, reliability modeling, risk-informed design.

1. INTRODUCTION

Over the years, NASA has developed a working knowledge and body of standards to guide both the design and the evaluation of safe human space flight systems. These include specific requirements on procedures and best practices in addition to quantitative mission risk requirements. As an example, NASA’s Commercial Crew Program (CCP) has identified quantitative requirements for loss-of-mission (LOM) and loss-of-crew (LOC) probabilities that establish safety and mission success metrics. To analytically verify these requirements before a system is operational, NASA programs often depend on probabilistic safety analysis (PSA).

In addition to requirement verification, PSA techniques can also be incorporated into the design process to optimize safety and mitigate critical risk drivers throughout system development. The quantitative safety and reliability information from PSA enables a risk-informed design process that includes risk as a design metric along with mass, performance, and cost. By injecting such information into the design process, the post-facto assessment of a system’s LOC/LOM probabilities is avoided, and actionable insights are provided at a point in the program where changes can still be impactful.

Traditionally, PSA, or probabilistic risk assessment (PRA), is performed using logical probability models [1]. Most commonly used are the event tree/fault tree (ET/FT) models that have been employed for systems such as the Space Shuttle and International Space Station (ISS). These methods are effective at combining failure probabilities into a series of top events or specified end-states. More recently, simulation methods have been employed to extend the traditional methods such that physical factors and system dynamics are inserted directly into the risk models [2]. The goal of these approaches is to more accurately represent the interactions among on-board systems and between the system and its operating environment. Whereas classical ET/FT approaches strive to generate a very precise solution to an approximate, static problem, simulation methods may be able to better represent a system’s fully interactive dynamics by producing an approximate solution to a more realistic problem. The ultimate choice of the “best” analysis approach is determined by the problem specifics, questions being addressed by the analysis, resources available, and desired level of precision.

\* Susie.Go@nasa.gov

Probabilistic Safety Assessment and Management PSAM 12, June 2014, Honolulu, Hawaii
This paper presents an example of a simulation-based risk model for a human spaceflight mission. The model blends engineering analyses commonly used during the design process with traditional reliability approaches to produce LOM and LOC estimates that can be used for risk-based decision-making and requirement verification. The example illustrates how the simulation model can produce top-level risk estimates, such as LOC/LOM, while also representing the physical aspects of the system that often drive the risk and are meaningful to system designers. In addition, the model can track multiple figures of merit (FOMs), time-based risk intensities, and system interactions within the same assessment. This paper utilizes a generic space transportation system to illustrate these capabilities.

2. EXAMPLE PROBLEM

An example mission for the end-to-end risk model was developed with similar mission requirements as those defined for the CCP [3]. The baseline mission includes delivering four astronaut crewmembers and their equipment to the ISS and returning them to Earth at least twice a year. A fictitious, generic human spaceflight launch vehicle (LV) and spacecraft (SC), shown in Figure 1, were used for the reference mission in order to illustrate the risk modeling process with representative design and performance data [4]. The launch vehicle is a two-stage, liquid-propellant rocket with four liquid oxygen LOX/RP-1 engines on the first stage and two LOX/hydrogen engines on the upper stage. The spacecraft is a capsule with a five-meter diameter and a tractor-type ascent launch abort system, sized to support the crew for seven days in free-flight or to dock with ISS for a six-month stay. All risk data and physical models in this paper are based on this generic space architecture, which is presented not as a closed architecture, but rather as an example architecture used to demonstrate the risk assessment process. The baseline mission timeline for this architecture was assumed to include a 9.25-minute ascent phase before separation of the spacecraft from the upper stage, a one-day flight to ISS, a six-month stay docked to the ISS with weekly one-hour spacecraft checkouts, and a four-hour entry, descent, and landing (EDL) phase, as shown in Figure 2.

The top-level LOC/LOM metrics were computed for this example mission and the leading risk drivers called out. In addition, the time-based risk intensity was produced for the system. Key interactive examples where dynamics and physics are first-order risk drivers are shown. All of this information was produced by a single model, removing the need for multiple risk models to track specific FOMs.

Figure 1: Launch vehicle and spacecraft specifications [4].
3. DYNAMIC MISSION RISK MODEL

3.1. Risk models and required data

A space launch system travels through multiple environments with vastly different timescales, and the integrated risk model was developed with these variations in mind. Table 1 summarizes some of the key data required for a typical mission risk assessment model. The data covers multiple science and engineering disciplines, including vehicle design, blast physics, trajectory analysis, propulsion system performance, aerodynamic and aerothermal analyses, environment characterization, material responses, and reliability engineering. Traditionally, these modeling methods are used in the design process with an emphasis on optimizing system performance or defining a feasible system. However, the outputs from each of these analysis disciplines can also be focused on understanding system performance during off-nominal conditions or degraded states to provide valuable risk information for both designers and decision makers. In this way, success-based design models can be used to populate the risk model.

Table 1: Summary list of data requirements for a typical mission risk model.

<table>
<thead>
<tr>
<th>Data Requirement</th>
<th>Ascent Phase</th>
<th>On-orbit Phase</th>
<th>EDL Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Launcher design and reliability</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spacecraft design and reliability</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Integrated stack design and reliability</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trajectory data</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aerodynamic environments</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aerothermal heating environments</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space environments</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orbital and deorbit maneuvers</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abort conditions</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Material response</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Concept of operations and contingencies</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>
In order to best represent the space exploration system, the dynamic mission risk model uses a hybrid approach that integrates the outputs from each of these disciplines to calculate estimates of key system metrics. Figure 3 contains an overview of the model’s data structure. At the top of the model are the elements that represent the dynamic behavior of the system and track system-level parameters of interests. Examples of such dynamics include physical failure propagation, system failure response, and the ultimate evolution of hazardous failure environments. The dynamic simulation elements are only incorporated to the level required to resolve dynamics that drive the overall assessment. Below these dynamic levels are static reliability elements that seed the dynamic simulation with likelihood information. These probabilities are often generated by representing components using traditional FT-like approaches. As a result, the simulation model handles the system dynamics and the full FT cut-sets while the static aspects live below the dynamics and can be incorporated back into the mission results through post-processing.

Figure 3: Example of multi-disciplinary, hierarchical failure data.

Starting at the bottom of Figure 3, representative fault trees can be seen to represent the component reliabilities. FT math is utilized when there are no important system dynamics or when a collective set of failures has a similar effect on the simulation, i.e. the simulation can’t tell the difference between failure types in the tree. This can occur when the system itself is accurately represented using static methods, when the data used to characterize the subsystem does not support dynamics, or when design fidelity is limited and a static representation serves as a reasonable proxy until additional information becomes available. Though the data in the trees may be static, the simulation can easily incorporate such data in a time-dependent way. For example, the simulation can apply a time-based duty-cycle that queries the tree data only during certain phases in the mission. Also, though FTs are often used to describe component reliability, the basic simulation elements can represent physical elements as well. For example, the current model incorporates offline EDL assessments, such as the one described in [5], at the basic simulation element level. This does not mean the EDL results are imported as a single probability, but that the results can be represented parametrically by a single element. Figure 4 shows a sample sub-element tree for the spacecraft’s pressure control system (PCS), based on component data from the master equipment list.
Moving upward through Figure 3, the level above the fault trees and basic simulation elements represents failure propagation through the system. The lowest levels, described above, determine the probability of a failure initiating. The middle levels of the model represent the translation of the initiators into cases that threaten the mission or crew. One such translation is performed by mapping the initiators to subsequent failure environments. This approach is relatively easy to implement and can be very effective in trade studies where a range of failure paths are to be considered. Consider the possible failure modes of the launch vehicle as shown in Figure 5 (from Ref. [6]). Each failure can progress down a number of paths, leading to a range of failure environments. The propagation table allows these “branches” to be assessed using offline models whose results can be directly incorporated into the mission risk model. Should the propagation analysis be updated, or have different assumptions applied, the values in the table can be updated without any change to the mission risk model. For more details on this approach, see Ref. [6].

Figure 5: Sample propagation table for launch vehicle failure modes [6].

†Though the example considers the elements at the bottom of the model to represent failures that trigger LOM, they can just as easily represent system failures that represent a loss of redundancy. Loss of redundancy can be mitigated by the design of the system, can trigger an abort in cases where the degraded system is considered too risky for mission continuance, or can represent a true loss of system function. In all cases, the dynamic simulation model can manage the occurrence and associated response.
While the previous example represents different discrete propagation paths resulting from a launch vehicle failure, continuous failure propagation models can also be incorporated. Consider a seal leak due to one of the PCS hatches shown in Figure 4. The simple, worst-case assumption would be that any PCS failure would lead to LOC immediately, but this assumption would be excessively conservative and drive the risk of the subsystem. To refine this assumption, the actual cabin environment can be modeled to allow the spacecraft to abort from orbit and return the crew safely. Reference [7] describes such a model, which has been incorporated in this assessment. With this additional modeling incorporated, the different leakage failure modes can now map to specific physical responses. For example, a valve that has failed opening in a leak of cabin pressure will not lead to immediate LOC. Instead, the time-to-LOC is predicted based on various mission parameters such as the size of the leak, available pressurant (itself a function of mission time and state), leak detection capability, and time required to safely depart ISS and return to Earth. The curves from [7] are shown in Figure 6 and represent parameterized responses to cabin leakage. The mission model selects from the curves, based on the appropriate pre-failure mission state, and determines a resulting time-to-LOC given a cabin leak. Again, because the physics of the failure propagation are imbedded in the dynamic mission risk model, changes to the design, abort policy, crew protection using pressure suits, or any number of other factors can be directly reflected in the analysis without re-architecting the model.

Figure 6: Example LOC hazard data for various cabin leakage scenarios [7].

Sometimes the physics itself is responsible for the initiator. For example, consider micrometeoroid orbital debris (MMOD) threats to the spacecraft in orbit. An analysis of the MMOD threat was performed using the BUMPER II code developed at NASA Johnson Space Center [8]. BUMPER II predicts the probability of MMOD penetrating the spacecraft in orbit, which is used to populate a basic simulation element in the model. If penetration occurs, then one would assess the damage done to underlying systems to determine the criticality of the impact. If the pressure vessel was compromised, the cabin physics model could be employed to determine the time-to-LOC, or the time the spacecraft could remain pressurized such that alternate action could be taken. In the current model, any full spacecraft penetration when docked to the ISS is assumed to result in LOM if unoccupied, or LOC if occupied. Of more interest is the case where an MMOD strike does not fully penetrate but causes damage to the thermal protection system (TPS). In this case, the impact damage is passed to the EDL element, described above, and the dynamic simulation model computes the LOC probability based on the heating environment, location of the damage, TPS material response, and temperature limits of the spacecraft structure. All of this is handled within the mission model automatically.

All of these examples begin with an off-nominal situation that initiates a potentially fatal chain reaction. The estimation of LOC depends on the ability of the space transportation system to respond in a manner that a) inhibits the failure propagation so that no threatening failure environment develops, b) protects the crew in spite of the failure environment, or c) can move the crew safely away from the hazard. In the second and third cases, the evolution of the failure environment must be characterized.
Much work has been done to estimate the blast overpressure, fireball, and debris environments resulting from an explosive system failure [9,10,11]. The current risk model incorporates these hazardous environments using parameterized tables of results computed outside of the risk model. In this way, the best estimates of the hazardous environments’ impact on the system are incorporated into the model and the resulting failure probabilities are computed using the state of the launch vehicle and spacecraft at the time of the failure.

A short list of the physic-based models needed to represent failure and propagation conditions encountered in a space mission risk model includes:

- LOC probability due to blast overpressure as a function of mission elapsed time (MET) and amount of abort warning time available.
- LOC probability due to debris striking the crewed cabin as a function of MET and amount of abort warning time available.
- Time to structural breakup of the launch vehicle in a loss-of-control situation, as a function of MET and failure mode (rapid versus slow control loss).
- Probability of penetration and partial penetration from MMOD as a function of TPS thickness and time in orbit.
- LOC probabilities due to TPS bondline temperature exceedance as a function of TPS thickness and TPS stack-up.
- Time-to-LOC cabin environment conditions as a function of failure severity, MET, and abort from orbit duration.

At the top of the model hierarchy shown in Figure 3 are the elements that track the failures, mitigations, and responses of the system. These top elements determine whether LOC, LOM, or any other simulation parameter of interest occurs. An attribute of simulation techniques is that multiple FOMs can be tracked within a single simulation. For example, one model can estimate LOC, LOM, failure time/risk intensity, failures that did not lead to abort, etc. These are the multiple top events shown in Figure 3. The parameter estimation is accomplished by simulating a large number of mission realizations, storing the end-state of each, and reporting the statistical results at the end of the simulation. Figure 7 shows a diagram of how the data items combine to produce these statistics.

Figure 7: Flow chart of mission risk data.
4. RESULTS

All results in this paper were run using a commercial Monte Carlo simulation code called GoldSim (www.goldsim.com/). A Monte Carlo simulation framework was chosen for its flexibility in incorporating probabilistic data from multiple sources and linking them together as event-triggered responses in a time-stepped model. The physics models were run offline in advance to generate response surface failure probability tables that are used as inputs to the simulation. A lookup table is created for each model and the tables are linked to different failure events along the mission timeline, as prescribed by the propagation maps. If a failure event is triggered during a simulation realization, the time of occurrence is captured and the appropriate value(s) are pulled from the relevant hazard lookup table(s). Each mission in a realization is “flown” to completion until a LOC event occurs. Unless otherwise noted, all simulations were run with 40,000 Monte Carlo realizations using Latin Hypercube sampling.

4.1. LOM and LOC for Baseline Design

First, consider the integrated mission LOC/LOM assessment typically associated with PSA. This type of analysis would normally be used for requirement verification. Figure 8 shows the mean LOM and LOC probability breakdowns by mission phase for a 183-day mission to ISS. The mean probability of LOM for this space system is 1 in 26 and the mean probability of LOC is 1 in 244 for a single Monte Carlo run. The stacked bar chart shows most of the LOM contribution occurring during the long orbital phase, followed by LOM contributions of ~1:80 during the ascent phase. The LOC contributions show most of the failures occurring during the ascent phase and the long orbital stay.

Figure 8: LOM and LOC estimates for generic spaceflight system mission to ISS.

![Mission LOM-LOC by phase](image)

Figure 9 shows a chart of the top 20 LOM and LOC risk drivers for this example spaceflight system, sorted by descending contribution to LOC and then LOM. The driver labels are the loss of critical function “top events” in the risk model data hierarchy and are coded “[LV-xx]” for launch vehicle and “[SC-xx]” for spacecraft. Here, LOM is further differentiated as either a mission abort or a loss of vehicle (LOV). The top-20 list shows that the largest contributor to LOC comes from pressure vessel leaks and catastrophic first stage engine failures. The top LOM driver during the on-orbit phase stems from critical MMOD strikes to the spacecraft: ~1% occurrence for a 183-day stay while docked to the ISS. Since the spacecraft is largely unoccupied during this time (checkouts with crew occur weekly and for only 1 hour while docked), these critical penetration failures specifically lead to an LOV and would require a rescue mission to return the crew safely back to Earth, unless some repair capability was defined.
4.2. Trade Studies and Uncertainty Studies

While a LOC/LOM assessment has value, the requirement verification aspect of such assessments often occurs fairly late in a program, after the design has frozen to the point where a complete analysis can be performed. This is a bit of a catch-22, since the risk information is most impactful while the design is still progressing. A model such as the one described above, however, can provide additional insight throughout the design process. While the results will not necessarily characterize the final system to a large degree of precision, they offer tremendous value by providing context for design choices. The model’s modular design allows sensitivity and trade studies to be easily performed by sampling the corresponding rows or columns in the lookup tables.

For example, consider a system trade study looking at ascent abort triggers that detect launch vehicle failures and recommend when an abort should be attempted. Suppose a large number of abort trigger options are being considered, but the system impact (e.g., cost, schedule, or complexity) is significant if the sensor suite and software must be re-engineered. Assume three options have been proposed for the abort trigger thresholds: the baseline option, a moderate option, and a sensitive option, as shown in Table 2. The baseline system has little or no abort capabilities (a large vehicle deviation rate trigger, no vehicle attitude trigger, and no early warning systems) and is compared against the two alternate abort systems.
Table 2: Ascent abort modeling assumptions used for abort trigger trade study.

<table>
<thead>
<tr>
<th>Abort system</th>
<th>Trigger</th>
<th>Baseline</th>
<th>Moderate</th>
<th>Sensitive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Launch vehicle rate trigger (deg/s)</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Launch vehicle attitude trigger (deg)</td>
<td>90</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Abort system latency (ms)</td>
<td>300</td>
<td>300</td>
<td>300</td>
</tr>
<tr>
<td></td>
<td>First stage warning time (s)</td>
<td>0</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Second stage warning time (s)</td>
<td>0</td>
<td>1</td>
<td>5</td>
</tr>
</tbody>
</table>

The risk model was run with a single failure environment assumption set and the three different abort trigger options. Figure 10, shows the 90% confidence intervals for LOC generated from the Monte Carlo simulations of each abort system. An abort system with the moderate design thresholds offers a large benefit to LOC, reducing the mean LOC probability from 1 in 660 missions to 1 in 1,220 missions compared to the baseline abort system. The sensitive abort system with even more responsive thresholds further reduces the occurrence of LOC to a mean of 1 in 2,830 missions. This does not imply that such systems are desirable or even feasible, but it allows designers and decision-makers to perform what-if studies and make risk-informed decisions about the specific design threshold values that are needed to achieve a certain level of protection.

Figure 10: Abort trigger trade study ascent LOC results ($5^{th}$, mean, and $95^{th}$ percentile confidence bounds).

While the impact of the abort trigger options could be very significant in meeting a LOC requirement, it is also important to consider whether the results reflect the best knowledge of the system. For example, assume previous studies have shown that the assumptions about the blast and debris environments are very impactful to ascent LOC results. The project team would benefit from insight into how effective different the trigger results would be under different explosion scenario assumptions. Uncertainty distributions could be assigned to the likelihood of the range of potential blast environments, but this would yield the same mean results shown above, assuming the mean of the distribution was used for the sensitivity study. The confidence intervals would likely be different, but it is unlikely that they would characterize the differences in a way that could meaningfully impact trigger selections. Alternately, the trigger trade study could be re-assessed under specific blast assumption cases. This would provide a direct connection between the trigger results and the blast assumptions and would enable a decision based on the belief of those assumptions. This takes traditional uncertainty quantification to another level and highlights whether a result is sensitive to the uncertainty. In this way model uncertainty is included in the assessment, not in a combined way that...
In addition, this model now allows a user to understand the range of possible conditions following a failure rather than choosing one specific set of assumptions.

Incorporating the risk assessment into the design process can also produce significant value by iteratively identifying where overly conservative or insufficiently defined assumptions may need to be refined. For example, Figure 9 shows the top LOC driver to be on-orbit seal leakage. This seems intuitively wrong, so this risk driver should be examined further. It turns out that the con-ops stated that an abort would be initiated in the event of any PCS failure, so the model had the crew enter a “leaky” spacecraft and return to earth. While this may sound ridiculous, is it the result of a bad model? Maybe not, as it is not unusual to have blanket policies such as these defined as placeholders during the early design phases. At this point, the hypothetical abort policy could be raised as an issue. While the LOC impact was certainly over-stated, the risk model highlighted an incomplete con-ops definition. By revisiting and refining this assumption, the LOC became LOV, and further definition of mission operations and repair capabilities could remove these LOV even further.

Similarly, the top LOM driver is MMOD penetration on orbit. As mentioned, this is LOV, not LOC, because a penetration can be isolated and the crew protected while the spacecraft is docked. However, the model allows simple trades to reduce the potential for these LOM (and LOV) scenarios. Changing the TPS design definition by increasing the TPS insulation thickness from 1 inch to 1.5 inches reduces the probability of a LOV from 1.04e-2 (+ 8.33e-4 at 90% confidence) to a probability of LOV 4.08e-3 (+ 5.24e-4 at 90% confidence).

5. CONCLUSION

A multi-disciplinary risk modeling approach was developed to merge the best information from each engineering field into a cohesive view of a space transportation system. A generic space architecture and mission were used to demonstrate how such a modeling approach can provide valuable risk-based insights during the design cycle rather than just at the end.

For the risk model to impact the design, including con-ops, it must tie to meaningful design parameters and be responsive to the design process. It is suggested that the real value of PSA for space system design is not as much from the LOC/LOM verification but from the insight it can provide throughout the development process. The example in this paper illustrated aspect of dynamics modeling, inclusion of physics, trade support, model uncertainty, and using the PSA to drive discussions that would be easy to discount and delete from a risk model. The example also showed how a model can highlight a gap, not necessarily in the design or data, but in the operation assumptions themselves. Assumptions causing high variations in the outputs of interest merit further exploration and refinement of the underlying inputs, while assumptions that do not drive the key results may be suitable to use as-is. As design data matures, an iterative feedback loop between designers and risk analysts can be established early, when trade studies are still being performed and hold the most promise of being implemented. In order for such value to be realized, the model needs to be responsive so that improved data, assumptions, policy, or design features can be readily incorporated. If the analysis cannot be updated frequently, then such things as the seal leaks causing LOC will either encourage the design team to discount the PSA altogether, or even worse, put design resources into fixing a non-problem. Responsiveness is a key aspect of the modeling approach presented in this paper.

In these ways, the risk model’s responsiveness and ability to represent meaningful parameters enable it to provide actionable, risk-informed design data. In addition to simply verifying requirements, the risk model provides an intelligent roadmap for a spaceflight program to plan future analysis studies and optimize design improvements throughout the system development lifecycle.

---

1 The architecture used for this study, as mentioned, was conceptual only and the abort policies were defined such that an abort was triggered on a loss of redundancy or function.
Acknowledgements

The authors wish to thank Samira Motiwala for developing the design and reliability data for the generic space launch system and spacecraft used in this paper.

References