RESILIENT EXTRA-TERRESTRIAL HABITAT DESIGN USING A CONTROL EFFECTIVENESS METRIC

by

Meghan V. Cilento

A Thesis

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Master of Science in Aeronautics and Astronautics

School of Aeronautics and Astronautics

West Lafayette, Indiana

August 2022

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Karen Marais, Chair
School of Aeronautics and Astronautics

Dr. Shirley Dyke
School of Aeronautics and Astronautics

Dr. Leifur Leifsson
School of Aeronautics and Astronautics

Approved by:
Dr. Gregory Blaisdell

Dedicated to my parents, who may not always understand what I do, but who never fail to give me the love and support I need to see it through

ACKNOWLEDGMENTS

This material is based upon work supported in part by NASA under grant or cooperative agreement award number 80NSSC19K1076. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Aeronautics and Space Administration (NASA).

I would like to thank my parents, who have always supported me through my education and given me the means to reach for any goal I set my mind to. I would not be where I am today without them, and I cannot express the gratitude I feel to have them behind me. I would like to also thank my fiancé, Daniel Perrefort, who has been with me through the ups and downs of graduate school. His love and support pushed me to keep going, work hard, and never give up.

I would like to thank the people who made this work possible. The RETHi team, especially the MCVT model developers, who have worked closely with me throughout my involvement on the project, and were integral to the completion of my research. Yuguang Fu, Murali Krishnan, and Alana Lund were supportive team members who helped me understand our simulation development from a modeling perspective and without whom I would not have gained the collaboration and communication skills I now hold. In addition, the undergraduates who have worked with me; Hunter Mattingly, Kyle Alvarez, and Mark Zhang. Hunter and Kyle were a huge help in maintaining and expanding our database, and Mark was an essential part of developing the safety control process models. Finally, Masters student, Jacqueline Ulmer, who was invaluable in gathering MCVT data and completing the resilience metric assessment.

I'd also like to thank everyone in the VRSS lab, whose feedback and support helped make me the best researcher and speaker I could be. I've not only found colleagues in my lab mates, but also lifelong friends. Finally, I'd like to give a big thank you to my advisor, Dr. Karen Marais. I wanted to work with her before I was accepted at Purdue, and having had the opportunity to do so has been a privilege. Thank you for giving me so much support throughout my graduate school journey and for helping me become not only a better researcher, but a better student and person.

TABLE OF CONTENTS

LIST OF TABLES 8
LIST OF FIGURES 10
ABSTRACT 13
1. INTRODUCTION 14
1.1 Motivation 14
1.1.1 Resilience 14
1.1.2 Risk Assessment & Accident Modeling Techniques 16
1.2 Resilient Extra-Terrestrial Habitats institute (RETHi) 16
1.2.1 The Modular-Coupled Virtual Testbed (MCVT) 18
1.3 Thesis Objectives and Outline 18
2. CONTROL-THEORETIC APPROACH TO RESILIENCE 20
2.1 The State-Based Safety Model 20
Step 1: Identify Hazards 20
Step 2: Hazard Assessment 21
Step 3: Identify Safety Controls 21
Step 4: Safety Control Assessment 21
Step 5: Residual Risk Assessment 22
2.1.1 The System Safety Process 22
2.2 The State and Trigger Model 25
2.3 Previous Development of Disruptions, Hazardous States, and Safety Controls 28
3. DEVELOPMENT OF CONTROL EFFECTIVENESS 32
3.1 Safety Control Implementation Strategies 32
3.2 Safety Control Flaws and Generic Safety Control Flaws 33
3.3 Developing the Control Effectiveness Metric 36
Probability of Availability 38
Probability of Competent Design 39
Probability of Perfect Implementation 39
Response Margin 40
3.4 Developing Guiding Questions to Estimate Control Effectiveness 41

3.4.1 Color-Coding Control Effectiveness 46
3.5 Application: Assessing Safety Controls for Example Disruption and Hazardous State Scenario 47
4. DEVELOPING A CONTROL EFFECTIVENESS VALIDATION PLAN 53
4.1 The Four Step Validation Cycle 53
4.1.1 Assess Control Effectiveness for a Set of Safety Controls 54
4.1.2 Identify Relevant Performance Metrics 66
4.1.3 Obtain Performance Metrics as a Function of Time 75
The Modular Coupled Virtual Testbed (MCVT) 75
Modeling Variations in Control Effectiveness Values 77
4.1.4 Use Performance Curves to Assess Resilience 79
Resilience Metric Literature Review 80
Selecting Metrics for Quantitative Resilience Assessment 84
5. CONTROL EFFECTIVENESS VALIDATION RESULTS 94
5.1 MCVT v6 Current Capabilities and Limitations 94
5.2 Disruption Scenario 1: Meteorite Impact on Solar PV Arrays 95
5.2.1 Obtain Performance Metrics as a Function of Time 97
5.2.2 Use Performance Curves to Assess Resilience 102
5.3 Disruption Scenario 2: Meteorite Impact on Nuclear Radiator Panels 103
5.3.1 Obtain Performance Metrics as a Function of Time 104
5.3.2 Use Performance Curves to Assess Resilience 109
5.4 Disruption Scenario 3: Meteorite Impact on Structure at Location 2 110
5.4.1 Obtain Performance Metrics as a Function of Time 112
5.4.2 Use Performance Curves to Assess Resilience 120
5.5 Disruption Scenario 4: Fire Originating Near the Power Storage and Distribution Equipment 121
5.5.1 Obtain Performance Metrics as a Function of Time 123
5.5.2 Use Performance Curves to Assess Resilience 127
6. CONCLUSIONS 129
6.1 Summary 129
6.2 Key Findings 130

6.3 Limitations and Potential Improvement 132
REFERENCES 135

LIST OF TABLES

Table 1. Generic Safety Controls (Kitching, 2020) 30
Table 2. Generic Safety Control Flaws (Kitching, 2020) 35
Table 3. Definition of Control Effectiveness 37
Table 4. Guiding Questions to Evaluate Control Effectiveness 42
Table 5. Example Control Effectiveness Color-Coding Scheme 46
Table 6. Example "Risk Averse" Control Effectiveness Color-Coding Scheme 47
Table 7. Safety Control Implementation Strategies for Dust Accumulation Hazardous State 48
Table 8. Disruptions, Hazardous States, and Safety Controls Modeled in MCVT v6 57
Table 9. Implementation Strategies for Disruption Scenario Safety Controls 60
Table 10. Control Process Information for Safety Controls in MCVT v6 72
Table 11. Computation of $Re1$ for Each Resilience Curve Shape 87
Table 12. Computation of $Re2$ for Each Resilience Curve Shape 88
Table 13. Computation of $Re3$ for Each Resilience Curve Shape 89
Table 14. Computation of $Re4$ for Each Resilience Curve Shape 90
Table 15. Computation of $Re5$ for Each Resilience Curve Shape 91
Table 16. Computation of $Re6$ for Each Resilience Curve Shape 91
Table 17. Computation of $Re7$ for Each Resilience Curve Shape 92
Table 18. Summary of Analytic Computation of Resilience Metrics for Resilience Curve Shapes 93
Table 19. Categorization of High, Low, and Intermediate Safety Control Implementation Strategies for Disruption Scenario 1 97
Table 20. Safety Control Repair Rate Inputs for Agent Model in Disruption Scenario 1 98
Table 21. Resilience Metric Values for SC798 in Disruption Scenario 1 102
Table 22. Resilience Metric Values for SC800 in Disruption Scenario 1 102
Table 23. Categorization of High, Low, and Intermediate Safety Control Implementation Strategies for Disruption Scenario 2 104
Table 24. Safety Control Repair Rate Inputs for Agent Model in Disruption Scenario 2 105
Table 25. Resilience Metric Values for SC798 in Disruption Scenario 2 109

Table 26. Resilience Metric Values for SC800 in Disruption Scenario 2 110

Table 27. Categorization of High, Low, and Intermediate Safety Control Implementation Strategies for Disruption Scenario 3 112

Table 28. Safety Control Repair Rate Inputs for Agent Model in Disruption Scenario 3 113

Table 29. Resilience Metric Values for SC355 in Disruption Scenario 3 120

Table 30. Resilience Metric Values for SC798 in Disruption Scenario 3 121

Table 31. Resilience Metric Values for SC800 in Disruption Scenario 3 121

Table 32. Categorization of High, Low, and Intermediate Safety Control Implementation Strategies for Disruption Scenario 4 123

Table 33. Safety Control Repair Rate Inputs for Agent Model in Disruption Scenario 4 124

Table 34. Resilience Metric Values for SC797 in Disruption Scenario 4 (Temperature Metric) 128

Table 35. Resilience Metric Values for SC797 in Disruption Scenario 4 (Pressure Metric) 128

LIST OF FIGURES

Figure 1. The Resilience Curve 15

Figure 2. RETHi Representation of the Control-Theoretic Approach to Resilience (Dyke et al, 2018) 20

Figure 3. The Control-Theoretic Approach to Resilience (black text) mapped to the System Safety Process (red text) 23

Figure 4. The State and Trigger Model 26

Figure 5. Example State and Trigger Model 27

Figure 6. The Failure Network with Labeled Hierarchical Groupings (Kitching, 2020) 29

Figure 7. Control Effectiveness Results for Prevention Safety Controls 50

Figure 8. Control Effectiveness Results for Mitigation Safety Controls 50

Figure 9. Control Effectiveness Results for Intervention Safety Controls 50

Figure 10. Control Effectiveness Validation Procedure 54

Figure 11. General Structure of Events for a Disruption Scenario 55

Figure 12. General Disruption Propagation Matrix for MCVT v6 56

Figure 13. Control Effectiveness Values for SC798 62

Figure 14. Control Effectiveness Values for SC800 62

Figure 15. Control Effectiveness Values for SC785 62

Figure 16. Control Effectiveness Values for SC11 and SC822 63

Figure 17. Control Effectiveness Values for SC786 63

Figure 18. Control Effectiveness Values for SC795 63

Figure 19. Control Effectiveness Values for SC796 63

Figure 20. Control Effectiveness Values for SC823 and SC824 64

Figure 21. Control Effectiveness Values for SC355 64

Figure 22. Control Effectiveness Values for SC803 64

Figure 23. Control Effectiveness Values for SC802 64

Figure 24. Control Effectiveness Values for SC799 64

Figure 25. Control Effectiveness Values for SC801 65

Figure 26. Control Effectiveness Values for SC797 65

Figure 27. Control Effectiveness Values for SC787 65
Figure 28. Control Effectiveness Values for SC788 65
Figure 29. Control Effectiveness Values for SC789 65
Figure 30. Control Effectiveness Values for SC791 66
Figure 31. Control Effectiveness Values for SC792 66
Figure 32. Control Effectiveness Values for SC793 66
Figure 33. MCVT Architecture with Distribution of Disruptions, Safety Controls, and Installed Sensors 67
Figure 34. Control Process Model Adapted from Leveson (2004) 69
Figure 35. Control Process Model for Safety Control: *Ability to remove dust from solar PV arrays* 71
Figure 36. MCVT Layout of Habitat Subsystems (Dyke et al., 2022) 77
Figure 37. Nominal Resilience Curve 79
Figure 38. Visualization of Failure and Recovery Stages of the Baseline Resilience Curve 82
Figure 39. (Left) Performance Lost, (Right) Performance Gained 85
Figure 40. (Left) Bucket Shape, (Right) "V" Shape 85
Figure 41. (Left) Rigid Curve with No Recovery, (Right) Smooth Curve with No Recovery 85
Figure 42. (Left) Rigid "U" Shape, (Right) "Scoop" Shape 86
Figure 43. Multiple Minima 86
Figure 44. Schematic for Disruption Scenario 1: Meteorite Impact on Solar PV arrays 96
Figure 45. Solar Power Output for High, Low, and Intermediate Control Effectiveness Implementations of SC798 in Disruption Scenario 1 99
Figure 46. Nuclear Power Output for High, Low, and Intermediate Control Effectiveness Implementations of SC800 in Disruption Scenario 1 101
Figure 47. Schematic for Disruption Scenario 2: Meteorite Impact on Nuclear Radiator Panels 103
Figure 48. Solar Power Output for High, Low, and Intermediate Control Effectiveness Implementations of SC798 in Disruption Scenario 2 106
Figure 49. Nuclear Power Output for High, Low, and Intermediate Control Effectiveness Implementations of SC800 in Disruption Scenario 2 108
Figure 50. Meteorite Impact Locations on the Structural Protective Layer 110
Figure 51. Schematic for Disruption Scenario 3: Meteorite Impact on Structure Location 2 111

Figure 52. Hole Radius for High, Low, and Intermediate Control Effectiveness Implementations of SC795 in Disruption Scenario 3 114

Figure 53. Remaining Battery Cells for High and Low Control Effectiveness Implementation of SC355 in Disruption Scenario 3 115

Figure 54. Maximum Energy Storage for High and Low Control Effectiveness Implementations of SC355 in Disruption Scenario 3 116

Figure 55. Solar Power Output for High, Low, and Intermediate Control Effectiveness Implementations of SC798 in Disruption Scenario 3 117

Figure 56. Nuclear Power Output for High, Low, and Intermediate Control Effectiveness Implementations of SC800 in Disruption Scenario 3 119

Figure 57. Fire Starting Locations in the Interior Environment of the MCVT 122

Figure 58. Schematic for Disruption Scenario 4: Fire Originating Near the Power Storage and Distribution Equipment 122

Figure 59. Fire Radius for High, Intermediate, and Low Control Effectiveness Implementation of SC797 in Disruption Scenario 4 125

Figure 60. Interior Environment Temperature for High, Low, and Intermediate Control Effectiveness Implementations of SC797 in Disruption Scenario 4 126

Figure 61. Interior Environment Pressure for High, Low, and Intermediate Control Effectiveness Implementations of SC797 in Disruption Scenario 4 127

ABSTRACT

Extra-terrestrial habitats will be embedded in challenging environments and involve complex and tightly coupled combinations of hardware, software, and humans. Such systems will be exposed to many risks, both known and unknown, and anticipating all failures and environmental impacts will not be possible. In addition, complexity and tight coupling in these systems means space habitats are likely to experience system accidents, which arise not only from the failure of individual components but also from the interactions among components. Therefore, we propose a control-theoretic approach to resilient space habitat design, which is grounded in system safety engineering and goes beyond event and component-centric failure models underlying conventional risk-based design. We model the system from a state-based perspective where the habitat is in one of four distinct types of states at a given time: nominal, hazardous, safe, or accident. The habitat transitions from a nominal state to a hazardous state via disruptions, and further to safe and accident states via triggers. We use safety controls to prevent the system from entering or remaining in a hazardous or accident state, or to transition the system into a temporary safe state or back to a nominal state. We develop a safety control option space, from which designers choose the best control strategy to meet resilience, performance, cost, and other system goals. We show the development of a control effectiveness metric, which is defined to assess how well safety controls address the hazardous state or disruption for which they are designed. The control effectiveness metric is one dimension of the overall hazard mitigation evaluation, which should also include aspects like cost and launch mass. We validate this approach by assessing individual safety controls in the Modular-Coupled Virtual Testbed (MCVT). This physics-based habitat simulation models complex disruption scenarios which include unique combinations of hazardous states and safety controls. The MCVT allows for the activation of individual (and sets of) safety controls of varying control effectiveness values to evaluate habitat resilience under different control architectures. Using this simulation, we evaluate the control effectiveness metric to determine whether the definition is appropriate to select safety controls that lead to desired habitat resilience. Completing the validation of this metric is the first step towards the validation of the overall control-theoretic approach to resilient space habitat design.

1. INTRODUCTION

1.1 Motivation

Whether operating on the surface of the Moon or Mars, or in orbit around a planetary body, space habitats will be embedded in challenging environments and involve complex and tightly coupled combinations of hardware, software, and humans. These systems present *interactive complexity,* or the presence of unfamiliar or unplanned and unexpected sequences of events that are either not visible or not immediately comprehensible (Perrow, 1999). In addition, *tight coupling* in the habitat means each subsystem is closely connected to many other subsystems, and a change in one subsystem can rapidly affect the status of the others (Perrow, 1999). Space habitats will also be exposed to many risks during operation, both known and unknown, and anticipating all failures and environmental impacts will not be possible. Therefore, we need an approach that can (1) account for accidents that arise not only from component failures and (2) design systems that are resilient to both known and unknown hazards. To do this, we propose a control-theoretic approach that supports the development of habitat architectures that are *resilient* to its inevitable failures and (un)known risks.

1.1.1 Resilience

Resilience in engineering systems has been defined in many ways; however, at its core, each definition incorporates the ability of a system to react to, survive, and recover from disruptions. Resilience is dependent on the architecture of a system, which causes variations in how the system deals with threats or disruptions. The resilience of a system is often represented visually through resilience curves, which are a temporal sequence of the system's performance after a failure event or disruption (Figure 1). At the time of disruption, the system withstands the disruptive event ("Surviving the Disruption"), after which it recovers over time to its original nominal performance level ("Recovering from the Disruption"). The ability of a system to maintain functionality after a disruption can be termed "static resilience", while the recovery of the system after a disruption can be defined as "dynamic resilience" (Rose, 2005). Maximizing system resilience involves minimizing system degradation from a disruption (survivability) and decreasing the time between the system's performance level at the time of disruption to its regained performance level

(recoverability). This conceptualization is widely used in literature to depict the fundamental ideas behind resilience (Uday & Marais, 2015).

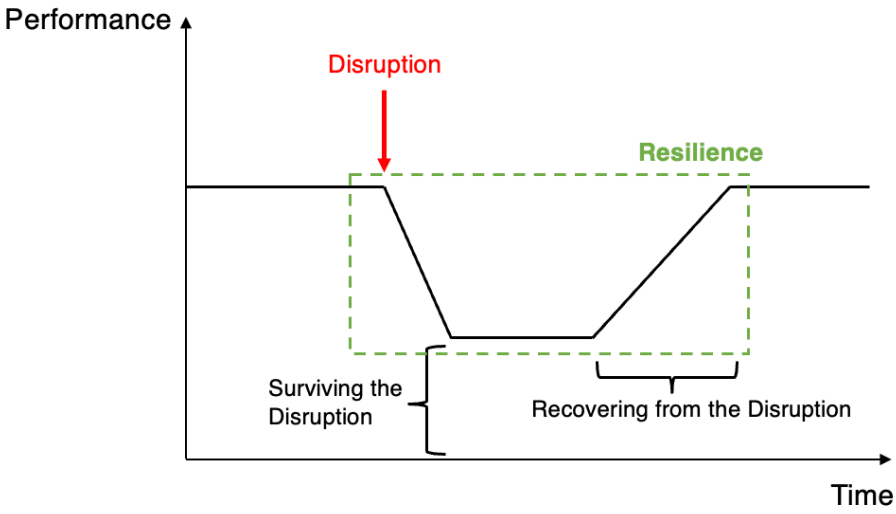

Figure 1. The Resilience Curve

It is common to study multiple resilience curves to gain a holistic view of resilience (Henry & Ramirez-Marquez, 2011). This may be necessary if it is not possible to understand the state of the system from a single performance metric in isolation. For example, a breach or leak in a space habitat's structure would likely cause a simultaneous decrease in the interior environment's temperature and pressure. Therefore, monitoring both temperature and pressure at different locations (or zones) in the system is necessary to detect, diagnose, and locate the breach or leak. Monitoring only temperature or pressure, or monitoring both at only one sensor location, would not be sufficient to understand the entire system state and pinpoint the location and severity of the breach or leak. Therefore, identifying several performance metrics for all known disruptions and cross-referencing their resilience curves is a necessary step in characterizing the entire system state and its resilience to simultaneous hazards. Such a holistic analysis allows for the identification of potential weaknesses in the safety structure that diminish resilience and the opportunity to enhance component, subsystem, or system design aspects that prove less resilient to disruptions than others.

1.1.2 Risk Assessment & Accident Modeling Techniques

Conventional risk assessment and accident modeling techniques rely largely on event-based and component-centric models that work well for loss caused by failures of physical components in relatively simple systems. However, rapid advancement in technology has resulted in the development of complex systems that are not fully understandable through existing techniques. The weaknesses and limitations in traditional approaches are well documented and focus on the inability of these methods to properly consider software, human interactions, and system accidents that arise not from component failure, but dysfunctional interactions among components (Kitching, 2020). Therefore, a new approach to risk assessment and accident modeling is required to understand and prevent accidents in complex systems such as deep space habitats.

Proposed by Nancy Leveson in 2004, the Systems-Theoretic Accident Model and Processes (STAMP) model offers a perspective on how to prevent and investigate system accidents. STAMP considers how accidents occur when disturbances, failures, or dysfunctional interactions among system components are inadequately controlled by safety-related constraints on the development, design, and operation of the system (Leveson, 2004). Rather than consider failures and faults and mitigate their effects, this type of approach views risk management from a control's perspective, where safety is considered a control problem. Such an approach offers a new technique to account for all types of accidents, including those that occur without any components failing.

1.2 Resilient Extra-Terrestrial Habitats institute (RETHi)

The work in this thesis was completed in conjunction with the Resilient Extra-Terrestrial Habitats institute (RETHi), a NASA-funded Space Technology Research Institute hosted at Purdue University and including representatives from the University of Connecticut, Harvard University, and the University of Texas at San Antonio. RETHi's vision is to "develop and demonstrate transformative smart autonomous habitats and related technologies that will adapt, absorb, and rapidly recover from expected and unexpected disruptions to deep space habitat systems without fundamental changes in function or sacrifices in safety" (Dyke et al., 2018). RETHi operates under the philosophy that resilience must be achieved by addressing the system as a whole, rather than as individual isolated components. As the system grows, complexity and connectivity increase the

risk of failure. Therefore, a comprehensive approach is needed that considers disruptions throughout the design process so that the system is capable of adapting to and rebounding from (un)foreseen events through a combination of preparedness, corrective actions, and autonomous interventions (Dyke et al., 2018).

For RETHi to accomplish this approach, we pursue three research thrusts. **Thrust 1** (Resilience Thrust) aims to develop and validate techniques that will establish a control-theoretic approach to resilience, as well as simulation environments to validate this approach. These simulation environments must be designed with the computational capabilities needed to capture complex behaviors and conduct trade studies to compare performance of different habitat architectures and on-board decision making. **Thrust 2** (Awareness Thrust) aims to develop and validate an intelligent and integrated health management (IIHM) system to enable detection and diagnosis of anticipated and unanticipated disruptions, and learn and predict future behaviors, needs, and responses of the habitat. **Thrust 3** (Robotics Thrust) aims to develop and demonstrate the technology needed for smart habitat interventions to realize autonomous response, repair, and recovery from disruptions through independent autonomous robots.

The objective of this thesis in the context of RETHi is to refine and validate the control-theoretic approach under development in **Thrust 1** that guides the design of resilient smart habitat architectures. Leveraging existing research on system accident modeling (STAMP), our control-theoretic approach is grounded in systems theory and considers the habitat as interrelated components kept in a dynamic state of equilibrium by feedback loops of information and control (Leveson, 2004). In the context of a deep space habitat, RETHi is interested in understanding how system resilience may be achieved while maintaining desired performance objectives and staying within specified safety constraints. Accidents in the system are therefore understood in terms of why the safety controls in place did not enforce the constraints that are designed to prevent, detect, and recover from inevitable failures and environmental disruptions. To design safety controls to properly enforce performance constraints, it is necessary to understand how we may determine whether a particular safety control (or combination of controls) is "good enough" at constraining system behavior. Therefore, a necessary part of our control-theoretic approach is the development of a *control effectiveness metric* which indicates how well a particular safety control addresses the

failure or disruption for which it is designed. Validating our approach therefore requires the validation of this control effectiveness metric, which we accomplish through the development of a testing procedure which uses a simulation environment with the computational capability to capture complex behaviors and conduct trade studies of different safety control combinations embedded in a hypothetical habitat architecture.

1.2.1 The Modular-Coupled Virtual Testbed (MCVT)

The Modular-Coupled Virtual Testbed (MCVT) is one of the simulation platforms under development to complete the testing plan for refining and evaluating control effectiveness and our control-theoretic approach to resilience. It is a virtual representation of the physical modules, or subsystems, of a hypothetical deep space habitat. Built from a conceptualization of a notional real habitat (NRH), the MCVT is designed from systems engineering approaches to allow for systematic model development and integration as requirements expand to include new functionalities necessary to explore RETHi's research objectives. Over the last two years, RETHi has steadily improved the MCVT's computational capabilities to capture the complex behaviors and interdependencies between the physical modules over time, while imposing changing environmental conditions, system architectures, and on-board decision making. The simulation is intended to validate and demonstrate the techniques and technologies proposed in all three research thrusts. Additional information on the development and design of the MCVT is provided in later chapters.

1.3 Thesis Objectives and Outline

In this thesis, we demonstrate how we develop and validate a control effectiveness metric to assess how well safety control strategies constrain system behavior and keep a habitat operating nominally. We emphasize the use of established system safety engineering processes and systems models to develop and assess control effectiveness. We use a physics-based habitat simulation for the modeling of individual safety controls with different control effectiveness values. Obtaining habitat resilience curves as a function of time allows for the computation of resilience metrics and the completion of a resilience assessment to evaluate how different habitat control structures respond to known disruptions.

In Chapter 2, we describe our control-theoretic approach for the design of resilient space habitats. We start with a discussion on how our approach maps to the traditional system safety process, followed by how the approach is based on a state-based accident model rather than models that emphasize components and their failures. At the end of Chapter 2, we summarize the work done by Purdue alumnus, Robert Kitching, who completed Steps 1–3 in our control-theoretic approach for his MSAA thesis (Kitching, 2020). In Chapter 3, we describe the development of the control effectiveness metric, including an example of evaluating the metric for a disruption and hazardous state scenario. In Chapter 4, we describe the development of a control effectiveness validation plan, which details the parallel development of the MCVT to support the evaluation of control effectiveness. In Chapter 5, we present the results of using the MCVT to evaluate control effectiveness for safety controls activated individually in response to known disruptions and hazardous states. Finally, Chapter 6 concludes this thesis.

2. CONTROL-THEORETIC APPROACH TO RESILIENCE

2.1 The State-Based Safety Model

To design resilient space habitats, we need an approach to resilience that (1) goes beyond the event and component-centric failure models underlying conventional risk-based design, and (2) helps identify designs that are prepared for both foreseen and unforeseen risks (Kitching, 2020). The control-theoretic approach to resilience proposed by RETHi includes five main steps necessary to mitigate risk and keep the system operating in a region of safe behavior. Figure 2 provides a visual representation of the approach, and details of the five main steps follow.

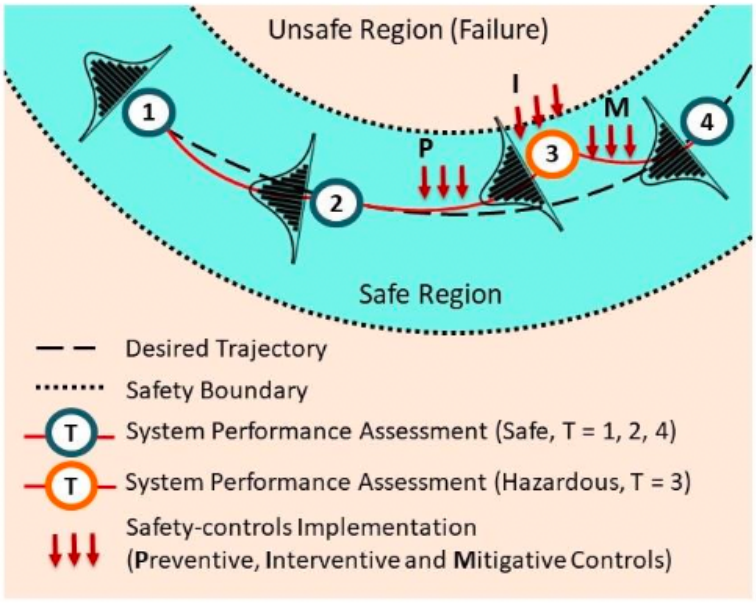

Figure 2. RETHi Representation of the Control-Theoretic Approach to Resilience (Dyke et al, 2018)

Step 1: Identify Hazards

Step 1 is to identify events (*disruptions)* that could cause the system to transition from a region of safe behavior (*nominal state*) to a region of unsafe behavior (*hazardous state)*. In addition to identifying these disruptions, this step includes the characterization of the hazardous states resulting from such disruptive events.

Step 2: Hazard Assessment

Step 2 is to conduct a hazard assessment to categorize disruptions and hazardous states based on several factors, including the severity of their consequences, the probability of their occurrence, and the time available to respond to a hazard before an accident occurs. Using a risk assessment matrix, we then group disruptions and hazardous states by their corresponding criticality and prioritize those hazards that provide the most risk to the habitat.

Step 3: Identify Safety Controls

Step 3 is to identify actions or design decisions (*safety controls*) that may return the habitat from a hazardous state to its original nominal state or prevent the transition of the system to an accident or loss (*accident state*). These safety controls are designed to address the disruptions and hazardous states identified in Steps 1 and 2. In parallel with identifying safety controls, we also identify corresponding generic safety controls principles that describe each control based on their method or principle of controlling the source of the disruption or hazardous state (e.g., REMOVE SOURCE FROM COMPONENT or REPAIR COMPONENT). These generic safety control principles can then be applied to other disruptions or hazardous states identified later in the design process, allowing us to put safety controls in place for both known and unknown disruptions or hazardous states.

Step 4: Safety Control Assessment

Step 4 is to assess the effectiveness of each safety control at constraining system behavior and keeping the system in a nominal state. This step is achieved through creating a *control effectiveness metric* that indicates how well a safety control addresses the disruption or hazardous state for which it is designed. We first identify several implementation strategies that describe how a particular safety control goal is achieved. We further identify flaws in these strategies and develop generic safety control flaws to describe how safety controls may be or become ineffective at constraining system behavior (e.g., SAFE CONTROL ACTION IS PROVIDED TOO LATE or SAFE CONTROL ACTION IS UNAVAILABLE).

Step 5: Residual Risk Assessment

Step 5 is to formally decide whether residual risk in the system design is acceptable. If risk is deemed acceptable, the habitat design is documented, and risk is periodically reviewed from that point onward. However, if risk is unacceptable, the system must be modified to further mitigate and control hazards, and the risks are then re-assessed.

2.1.1 The System Safety Process

The five-step control-theoretic approach to resilience is one part of the overall habitat development process. We are developing this approach in parallel with other system designers in RETHi to contribute to a larger design trade-off that incorporates multiple considerations such as launch mass, cost, system performance, resilience power, control effectiveness, and other requirements. Grounded in system safety engineering and fit within traditional risk management frameworks, our control-theoretic approach maps to the system safety process as for example described by Bahr (2016), the overall purpose of which is to identify hazards, eliminate or control them, and mitigate the residual risks. Descriptions of each step in the system safety process developed by Bahr (2016) are outlined next.

The first step in the traditional system safety process is to *define the boundary conditions or analysis objectives*, which provide the level of protection desired for the system. The primary question designers should ask is "How safe is safe enough?". In addition, understanding what classifies as a negligible, minor, critical, or catastrophic hazard in this stage is important to develop appropriate mitigation procedures later on.

The second step is *system description*, which involves understanding how the system works and how the hardware, software, people, and environment all interact. This step is essential to avoid flaws in the safety analysis and control structure later on.

The third step is *hazard identification*, which is a kind of safety brainstorming in which possible and credible hazards are identified. Without this step, attempts at safeguarding a system or controlling risks will be inadequate. The fourth step is a *hazard analysis*, which is a technique used

to study the cause/consequence relation of all identified hazards in a system. This step is important in assessing which hazards are important to control and requires the completion of a risk evaluation that considers the likelihood and consequence of each hazard.

The fifth step in the system safety process is *hazard control*, which involves controlling the effects of the identified hazards. Through engineering controls, we may change the hardware of the system to eliminate or mitigate hazards, whereas through management controls, changes in the organization itself are made to control hazards. This step includes the verification of controls, which requires a method of verifying that the controls in place actually control the hazards or mitigate risk to an acceptable level. Once verification is complete, the final stage is to make the formal decision that the residual risk in the system is acceptable. Figure 3 shows how our control-theoretic approach maps to this system safety process described by Bahr (2016).

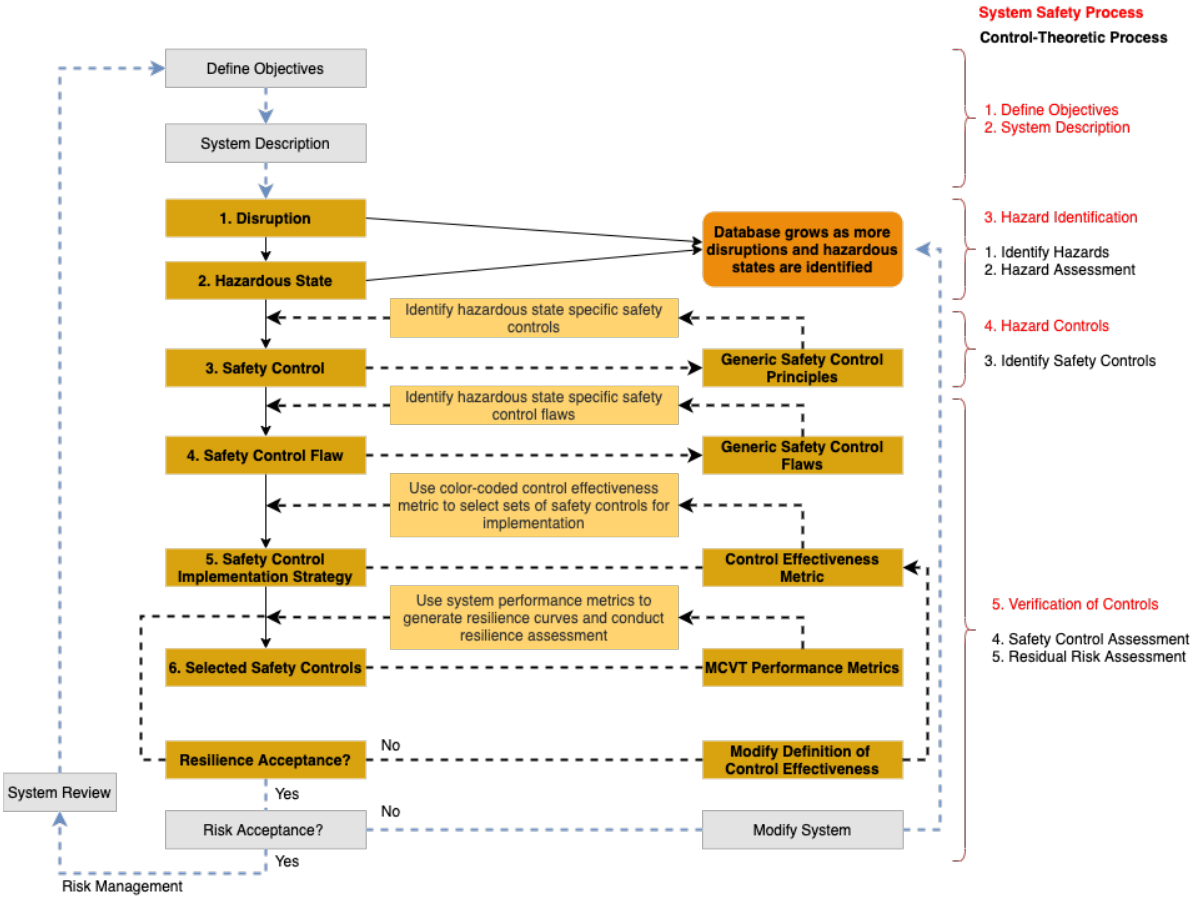

Figure 3. The Control-Theoretic Approach to Resilience (black text) mapped to the System Safety Process (red text)

Approaches for completing the system safety process described by Bahr (2016) are widely used in industry and across disciplines. Different industries approach safety in slightly different ways; so, there is much to be gained from understanding how other industries apply system safety engineering techniques (Bahr, 2016). The steps outlined by Bahr (2016) are an appropriate framework for understanding how safety is approached in industry. All approaches are rooted in identifying and understanding hazards, mitigating them, and then assessing how well the mitigative strategies worked. However, the techniques applied to complete this process differ across disciplines. For example, the manufacturing industry, although primarily compliance based, largely includes system safety engineering techniques such as safety checklists and process hazard analysis (specifically for industries that handle hazardous chemicals) (Bahr, 2016). In addition, the 1992 Occupational Safety and Health Administration (OSHA) regulations in the chemical, oil, and gas industries apply system safety engineering to the process industry (Bahr, 2016). This set of regulations formalized a standard safety analysis process and hazard management for these industries. In particular, the oil and gas industry adopted new safety tools such as HAZOP (Hazard and Operability Analysis) to create a structured technique for system examination and hazard identification. We can also consider the aviation industry, which uses system safety engineering and regulatory compliance. The four main analyses used to complete the system safety process are: functional hazard analysis, failure mode and effects analysis (FMEA), fault tree analysis, and zonal analysis (Bahr, 2016). In particular, the functional hazard analysis technique has informed the ARP (Aerospace Recommended Practice) standards, which provide guidelines on the development of aircraft with emphasis on safety. The aviation industry also incorporates human factor analysis tools such as the Human Factors Analysis and Classification System (HFACS), as much of the aviation industry is controlled by people. For more detailed information on system safety engineering and its evolution and application across industry, refer to Bahr (2016), as well as Gullo & Dixon (2018).

In addition to the examples discussed above, approaches have been developed which view safety as a control problem and emergent property of a system. Rather than approaching safety from a component-centric failure perspective, Rasmussen (1997) began efforts to use control-theory in accident modeling, which considers safety from the perspective of maintaining a system within

"boundaries of safe behavior". Others have followed this approach to safety. For example, as previously mentioned, the STAMP (Systems Theoretic Accident Model and Processes) model is a control-theoretic approach developed by Leveson (2004). This approach has been applied to several complex socio-technical accidents across several industries, including aerospace and water industries to demonstrate its use cases for studying accident causation (Leveson, 2012). Such control-theoretic approaches are relevant for the development of space habitats, because they account for the complexities and tightly coupled combinations of hardware, software, and humans in these systems.

2.2 The State and Trigger Model

Our control-theoretic approach considers the habitat system as being in one and only one of four states at a given time: nominal, hazardous, safe, or accident. A *nominal state* is when the system is within the boundaries of safe behavior. A *hazardous stat*e is when the system is in a state that, if left uncontrolled, will result in an accident or loss of life (*accident state*). A temporary *safe state* is a subcategory of the hazardous state, wherein an interventive action has occurred to stop the transition of the system to an accident state, but the primary source of the disruption must still be addressed to return the system to a nominal state. *Triggers* transition the system from one state to another. Each state must have at least one entering trigger. *Disruptions* are a type of trigger that instantiates transition to a sequence that includes hazardous or accident states. The safety-related constraints we consider in RETHi are termed *safety controls* and are defined as any aspect of the system design or operation that maintains the system in a nominal state, prevents the system from propagating to a hazardous state, or restores the system from a hazardous, safe, or accident state to a nominal state. To visualize the transition of the system between states, as well as how triggers and safety controls act to instantiate these transitions, we use a State and Trigger Model, as shown in Figure 4.

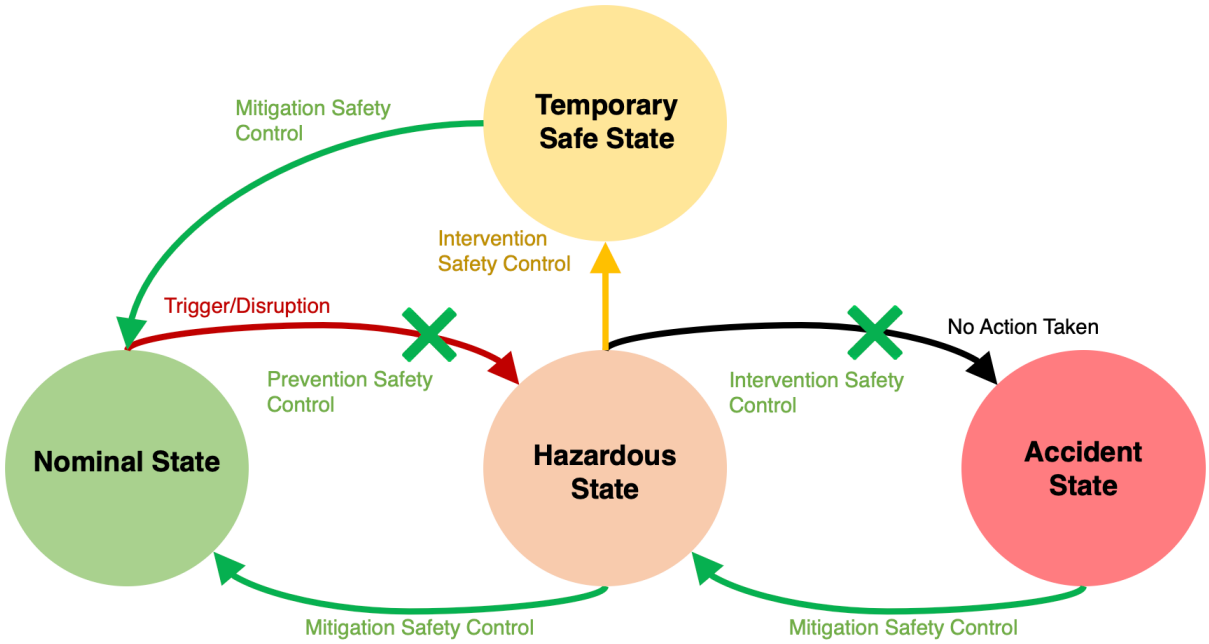

Figure 4. The State and Trigger Model

Our identified safety controls can be characterized as preventive, mitigative, or interventive, depending on where they exist in the State and Trigger Model. *Prevention* safety controls aim to maintain the system in a nominal state and prevent the system from propagating to a hazardous state. Conversely, upon transition to a hazardous state, *intervention* controls aim to prevent the system from propagating further to an accident state via the transition to a temporary safe state. We consider the safe state as a subcategory of a hazardous state, as the habitat is still operating outside the regions of safe behavior; however, performance degradation or cascading effects due to the disruption are temporarily delayed or alleviated to avoid the transition of the system into an accident state. A *mitigation* safety control is still required to enable the transition of the system back to the nominal state by directly addressing the source of the disruption. An example of a temporary safe state is to sequester section(s) of the interior environment to isolate a breach or fire. This would mitigate the cascading effects of such events and limit the portion of the habitat affected by the disruption. A second example of a safe state is when a redundant system or function is activated to provide support for the primary system operating in a hazardous state. A relevant case for this example is activating a secondary power generation system to supply critical loads to the habitat while the primary power generation system is repaired. In general, the use of a mitigation control is applicable if the system is in a hazardous, safe, or accident state, as such a

control acts to restore the system to its original nominal state. For all three of these safety control types, the control mechanism can be *passive*, and incorporated in the physical habitat design, or it can be *active* and require real-time response from agents or automated systems in the habitat. Classifying controls as prevention, intervention, and mitigation allows us to more easily make safety control selections in real time so we may appropriately respond to current system behavior. Moreover, identifying one or more safety controls in each of the three types is important for habitat resilience, as multiple layers of safety controls are needed if one or more fails or becomes significantly delayed in activation. Figure 5 shows an example State and Trigger Model corresponding to the hazardous state, *Degradation in solar power generation*.

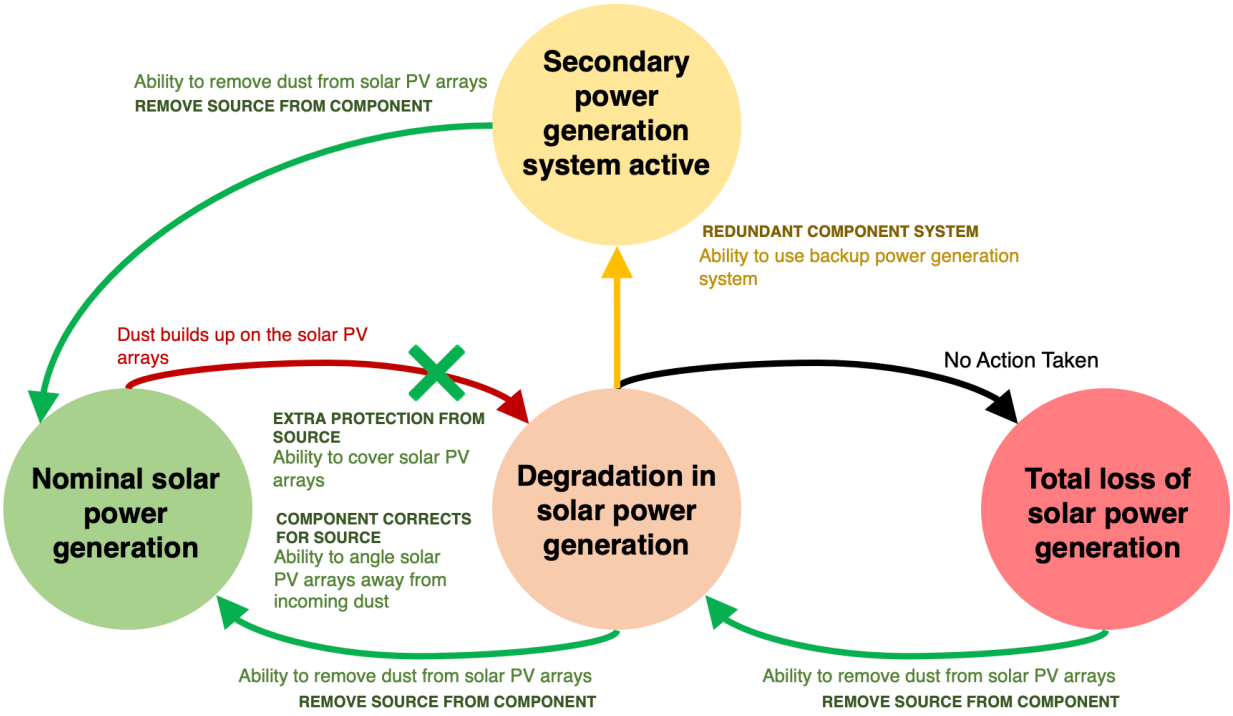

Figure 5. Example State and Trigger Model

The system begins in a nominal state, where solar PV arrays are clear of dust and power generation is normal. The buildup of dust on the solar PV arrays is modeled as an initiating disruption that causes the system to enter a hazardous state, in which solar power generation has degraded from the original nominal level. If no further action is taken, the habitat may further deteriorate into a state with total loss of solar power generation, which can have catastrophic effects on the essential

habitat systems which require constant power (e.g., Environmental Control and Life Support System).

When determining how to prevent the system from transitioning to a hazardous state, we must first consider the available prevention controls. The prevention safety controls shown in Figure 5 are represented by the green cross over the red trigger arrow. In this case, we can either provide protection from the dust by covering the solar PV arrays prior to an expected dust event, or orient the solar PV arrays away from the incoming dust. In the event that dust has already built up on the solar PV arrays, the mitigative action of removing dust would directly address the source of this disruption and act to transition the system back to its nominal state. However, the activation of the intervention control may be necessary if power generation degraded so severely that other habitat systems would be negatively affected in the time it takes to complete the mitigative action. Therefore, the intervention control would be activated to place the system in a temporary safe state in which a secondary power generation system is active while the mitigative safety control is completed. Upon removal of dust, the system would revert back to the primary power generation system and return to the nominal state. The layering of safety controls in this example demonstrates the need for multiple control actions to address one disruptive event and its potential consequences. In the event that a prevention control is inadequate at keeping the system in a nominal state, the mitigative control is in place to address the source of the disruption, and the interventive control is the final backup to avoid an accident in the event that the prevention and mitigative controls fail or experience a delay in activation.

2.3 Previous Development of Disruptions, Hazardous States, and Safety Controls

Using the State and Trigger Model as a basis, Steps 1–3 in the control-theoretic approach were completed by Purdue alumnus, Robert Kitching, as a part of his MSAA thesis (Kitching, 2020). Kitching developed a Microsoft Access database to organize and create relationships between the disruptions, hazardous states, and safety controls identified in Steps 1–3. Furthermore, working with undergraduate researchers, he generated a failure network to link the nominal state, disruptions, and associated hazardous states in a layered orientation to illustrate the propagation of the habitat from the nominal state to disruption, and through three levels of hazardous states

(subsystem, system, or habitat level). Figure 6 shows the failure network. Currently, the database houses 19 entering triggers (disruptions) and 221 hazardous states (spanning all three levels).

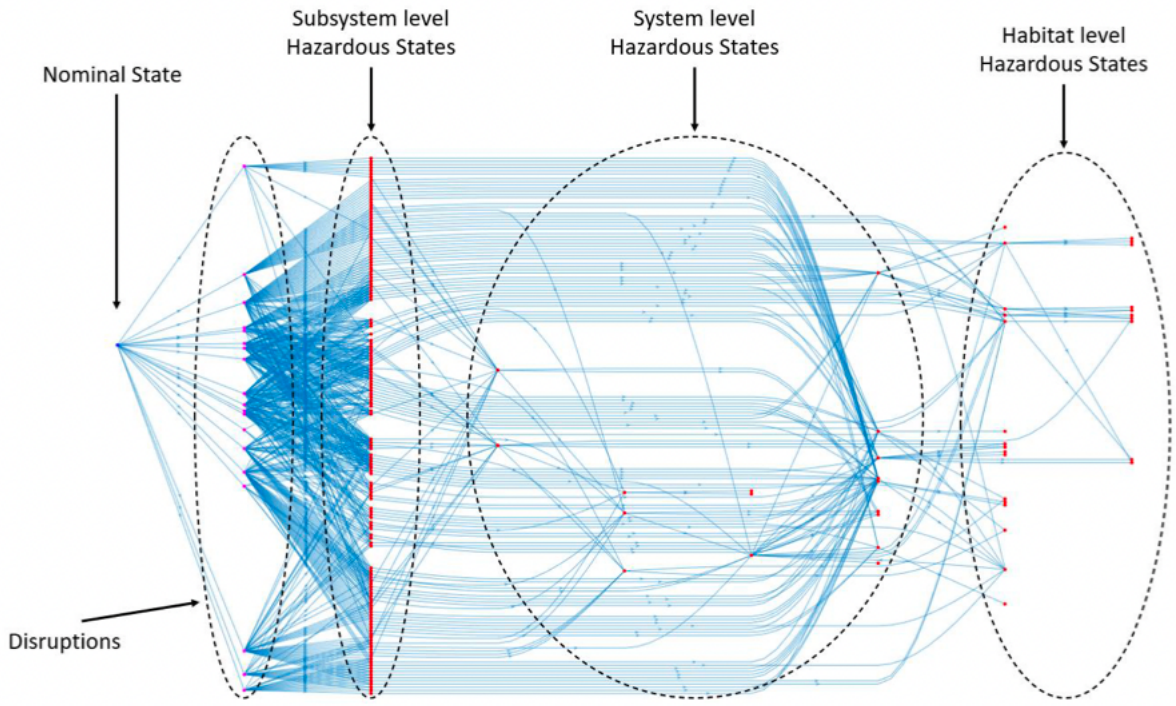

Figure 6. The Failure Network with Labeled Hierarchical Groupings (Kitching, 2020)

Kitching completed a hazard assessment using the failure network and associated network metrics to investigate the relationships between the disruptions and hazardous states in the database and prioritize controls. By leveraging knowledge in systems engineering, system safety, and of past accidents and incidents, he and the undergraduate researchers further identified safety controls to address those prioritized hazardous states and disruptions. The safety controls were developed throughout the design process, and the resulting set of potential safety controls formed an initial *safety control option space*. Currently, the database houses 784 safety controls to address the 19 disruptions and 221 hazardous states. For each of the identified safety controls, Kitching also developed *generic safety controls* to (1) identify more safety controls with various applicability to different disruptions and hazardous states, and (2) inform the evaluation of control effectiveness (covered in Chapter 3). For consistency, Kitching defined specific terms in the name and description of each generic safety control (as well as in generic safety control flaws) which is still used in later chapters of this thesis. These terms are as follows (Kitching, 2020): **ROBOT** refers

to the robot agent responsible for repairs, maintenance, inspections, and autonomous tasks that can be carried out without human intervention. **HUMAN** refers to the human agent, or crew member, responsible for maintaining, inspecting, and repairing the habitat when autonomous action is not sufficient or not possible. **COMPONENT** can refer to either the component, subsystem, or system in question that is being disrupted or is in a hazardous state. **SOURCE** refers to the source of the disruptive event or the hazardous state. For example, a dust storm is a source of a disruption, as is a micrometeorite impact. Table 1 describes all the generic safety controls identified by Kitching and currently used in the assessment of safety controls.

Table 1. Generic Safety Controls (Kitching, 2020)

Generic Safety Control	**Generic Safety Control Description**	**Transition to Safe State?**
REPAIR COMPONENT	The component or system affected by the source is repaired	No
REPLACE COMPONENT	The component or system affected by the source is replaced	No
ISOLATE COMPONENT	The component or system affected by the source is isolated to prevent further hazardous state	Yes
REMOVE COMPONENT FROM SOURCE	In the presence or anticipation of a source, the component or system is removed or shielded from the source	Yes
REMOVE SOURCE FROM COMPONENT	The source of the disruption is removed from the component or system	No
COMPONENT WITHSTANDS SOURCE	The component or system can function at a necessary level in the presence of a source	No
COMPONENT CORRECTS FOR SOURCE	The component or system adapts its function to protect against a source	No
REDUNDANT COMPONENT FUNCTION	The habitat can achieve the function of the component or system affected by the source using a different method	Yes
REDUNDANT COMPONENT SYSTEM	The habitat has another component or system to use when the operational component or system affected by the source can no longer by used	Yes
REDUCE COMPONENT LOAD	The component or system affected by the source is used less or at a lower capacity to ensure functionality	Yes
COMPONENT ROBUSTNESS	The component or system affected by the source is able to function in the presence or after being affected by a source	No
EXTRA PROTECTION FROM SOURCE	The habitat has the resources necessary to provide additional protection for vulnerable internal systems and/or crew members from the source of the disruption.	Yes
EXTRA PROTECTION FROM SOURCE	The habitat has the resources necessary to provide additional protection for vulnerable internal systems and/or crew members from the source of the disruption.	Yes
EVACUATE CREW	The human agents evacuate either to a part of the habitat that is not affected by the source, or leave the habitat entirely.	Yes

Table 1. Continued

RESUPPLY	The component, system, or resource produced or used by the component or system is resupplied from Earth	No
HUMAN VERIFIES SOFTWARE	The human agent verifies and confirms a process done autonomously by the habitat	No
COMPONENT DECENTRALIZED FUNCTION	A component or system can work independently to achieve a function that is done by a centralized system	No

Upon completion of Steps 1–3 in the control-theoretic approach, Kitching began the development of a procedure to assess the safety controls in the option space to accomplish *Step 4: Safety Control Assessment* in the control-theoretic approach. Kitching proposed a generalized procedure for completing this step of the approach, and his work sets the framework for the development of control effectiveness presented in **Chapter 3: Development of Control Effectiveness** of this thesis.

3. DEVELOPMENT OF CONTROL EFFECTIVENESS

In this chapter, we discuss the development of a control effectiveness metric to determine how well safety controls address their target disruption or hazardous state. Following the initial framework developed by Robert Kitching to assess safety controls (Kitching, 2020), we present the development and application of a modified definition of control effectiveness to fulfill *Step 4: Safety Control Assessment* in our control-theoretic approach to resilient space habitat design.

3.1 Safety Control Implementation Strategies

The safety controls discussed in previous chapters have been focused on *what* needs to be done to prevent transition to a hazardous or accident state, or return the habitat to a nominal state. To develop a better understanding of the effectiveness of each safety control, we must understand *how* each safety control achieves its goal. To do so, we consider the possible *implementation strategies,* which describe how a particular safety control goal is accomplished in the context of a particular disruption or hazardous state. For example, the mitigation safety control, *Ability to remove dust from solar PV arrays,* can be accomplished in (at least) the following three ways: (1) Human agent brushes dust off solar PV arrays, (2) Robot agent brushes dust off solar PV arrays, or (3) Built-in brush automatically removes dust from solar PV arrays. Although these three implementation strategies accomplish the same safety control goal, they differ in who or what removes the dust (human, robot, or automated mechanism), which subsequently affects the effectiveness of each strategy.

Upon identifying implementation strategies for all the safety controls in the option space, the next step is to determine how effective each implementation strategy is at accomplishing the safety control's goal (i.e., at addressing the target disruption or hazardous state). To do so, we developed a control effectiveness metric. Control effectiveness is one way of assessing our hazard mitigation techniques and can be used in combination with other control verification methods to confirm that our safety controls are adequate to maintain the system within boundaries of safe behavior. This metric is defined to allow us to make selections of appropriate safety controls from the safety control option space that will create habitat architectures with high resilience. Ultimately, control

effectiveness helps us answer the following questions: (1) How should we implement each safety control? (2) How might each safety control/implementation strategy be flawed? (3) Based on all available safety controls and considered implementation strategies, what control has the potential to be the most effective? Those safety controls (and implementation strategies) with high control effectiveness should therefore contribute positively to overall habitat resilience.

3.2 Safety Control Flaws and Generic Safety Control Flaws

As previously discussed, our control-theoretic approach to resilience relies on STAMP's principle that accidents occur when external disturbances, component failures, or dysfunctional system components are inadequately handled by the control system, that is, they result from inadequate safety constraints on the design, development, and operation of the habitat (Leveson, 2004). Therefore, understanding why accidents occur requires determining why the control system was ineffective, or why the controls in place did not detect or prevent performance changes that shifted the system towards hazardous or accident states.

Identifying several implementation strategies for each safety control in the option space helps us answer the first question: How should we implement each safety control? To then determine which of the identified strategies should be implemented in the habitat design and operation, we must answer the second question: How might each safety control/implementation strategy by flawed? To do so requires an understanding of the potential flaws in each implementation strategy that could cause the control to be ineffective at constraining system behavior. Therefore, we must identify safety control flaws and *generic safety control flaws* for each control and its corresponding implementation strategies. These flaws help us understand where we may need to reinforce or redesign safety controls to adequately address their target disruptions or hazardous states and keep the system within safe operating states. STAMP provides a classification of control flaws, which were developed for accident analysis or accident prevention activities (Leveson, 2004). In the context of our control-theoretic approach, the flaw classification is used to preemptively identify safety control factors that have the potential to make a safety control ineffective at keeping the system in safe operating states. The general classification of control flaws presented by Leveson (2004) is as follows:

(1) Inadequate Enforcement of Constraints (Control Actions)
    a. Unidentified hazards
    b. Inappropriate, ineffective, or missing control actions for identified hazards
        i. Design of control algorithm (process) does not enforce constraints
            1. Flaws in creation process
            2. Process changes without appropriate change in control algorithm (asynchronous evolution)
            3. Incorrect modification or adaptation
        ii. Process models inconsistent, incomplete, or incorrect (lack of linkup)
            1. Flaw(s) in creation process
            2. Flaw(s) in updating process (asynchronous evolution)
            3. Time lags and measurement inaccuracies not accounted for
        iii. Inadequate coordination among controllers and decision makers (boundary and overlap areas)
(2) Inadequate Execution of Control Action
    a. Communication flaw
    b. Inadequate actuator operation
    c. Time lag
(3) Inadequate or Missing Feedback
    a. Not provided in system design
    b. Communication flaw
    c. Time lag
    d. Inadequate sensor operation (incorrect or no information provided)

We focus on these control flaw classifications to develop our generic safety control flaws. When a safety control is not issued correctly, is inadequately executed, or does not provide feedback, we define that safety control as an *unsafe control action* (Kitching, 2020). Leveson identifies four ways that unsafe control actions can occur:

(1) A safe control action is not provided
(2) An unsafe control action is provided

(3) A safe control action is provided too late or too early

(4) A safe control action is stopped too soon or applied too long

Using Leveson's four generic classifications as a starting point, Robert Kitching originally mapped all identified safety control flaws to each of these generic categories. When an identified safety control flaw did not follow the principle of one of the four generic classifications from Leveson, the identified safety control flaw principle was mapped to a new generic safety control flaw to expand the final list. This process is similar to the identification of safety controls and generic safety controls previously described in the completion of Steps 1–3 of our control-theoretic approach. The final list of generic safety control flaws is shown in Table 2:

Table 2. Generic Safety Control Flaws (Kitching, 2020)

Generic Safety Control Flaw	**Generic Safety Control Flaw Description**
SAFE CONTROL ACTION IS NOT PROVIDED	Describes when the safety control is not implemented for any reason. This could be for example because it is not possible to be implemented, it is chosen not to be implemented, or that the safety control was attempted but not completed successfully.
SAFE CONTROL ACTION IS PROVIDED TOO LATE OR TOO EARLY	Describes when the safety control was stopped too soon or was activated too quickly. The safety control may not be completed in time for example because of a long or complicated procedure, lack of autonomous action, or lack of available resources to complete the safety control. A safety control may also be implemented too quickly, in that for example a component may be replaced before it needs to be.
SAFE CONTROL ACTION IS PROVIDED TOO MUCH OR TOO LITTLE	Describes when the safety control was not adequate to protect against the source, or when the safety control provides too much protection against the source that it becomes detrimental to other parts of the habitat. For example, shielding may be inadequate to protect against a micrometeorite, or crew protection may not be enough to protect against radiation.
SAFE CONTROL ACTION CAUSES UNSAFE CONTROL ACTION	Describes when the safety control execution makes the current hazardous state worse, as in that safe control action ends up causing an overall unsafe control action. For example, a repair could be completed incorrectly. A good intentioned repair, or a safe control action, is completed incorrectly and the component or system performs worse than before, creating an overall unsafe control action
SAFE CONTROL ACTION CAUSES HAZARDOUS STATE	Describes when the safety control execution makes the current hazardous state worse, as in that safe control action ends up causing an overall unsafe control action. For example, a repair could be completed incorrectly. A good intentioned repair, or a safe control action, is completed incorrectly and the component or system performs worse than before, creating an overall unsafe control action
SAFE CONTROL ACTION IS APPLIED TO LONG OR STOPPED TOO SOON	Describes when the safety control is executed for an unnecessarily long period, or if the safety control is stopped prematurely in the event of a source of a disruption. For example, if the crew is relocated due to a dust storm and there is no indication of when the dust storm ends, the safety control will still be implemented and it will be implemented for too long. Conversely, if the crew exits the relocation area during the dust storm, that will not constitute a safe control action because it will have been stopped too soon.

These generic safety control flaws are used to develop and evaluate control effectiveness for each implementation strategy, as discussed in the next section.

3.3 Developing the Control Effectiveness Metric

Control effectiveness is a metric that we use to determine how well safety controls (and implementation strategies) address their target disruption or hazardous state. As a part of his Master's thesis work, Purdue alumnus Robert Kitching developed an initial definition of control effectiveness. The definition incorporates a secondary metric called implementation strategy effectiveness (ISE), as well as a criticality score. The ISE metric (Equation 1) is a set of four values that quantify a control's expected probability of success, availability, competence against a source of a disruption, and the expected time it would take to implement a control:

$$ISE = \{P_{perfect}, P_{available}, P_{competent}, t_{active}\} \quad (1)$$

The criticality score contains information on a control's susceptibility to flaws, specifically, how likely it is for a control to exhibit the flaw, and how severe the consequences are of that flaw occurring (Equation 2):

$$Criticality_{IS} = \sum_{i=1}^{n} Likelihood_{IS,i} * Severity_{IS,i} \quad (2)$$

Control effectiveness was then defined as the combination of ISE values and criticality for each implementation strategy. Most of this initial definition by Robert Kitching carried through to the development of a modified definition, which we present here. The main differences in the modified definition are the representation of a control's susceptibility to flaws and the activation time for a control.

Currently, control effectiveness (*CE*) is a metric comprised of four values that aid us in discerning how effective a safety control is at mitigating its target hazardous state or disruption. This data set is defined to help designers understand how likely it is for flaws to exist in both the design and execution of a safety control. Each value is based one or more of the generic safety control flaws

described in Table 2, and therefore, the susceptibility of each safety control to known control flaws influences the effectiveness of an implementation strategy. Associating each control effectiveness variable with one or more generic safety control flaws helps designers form a connection between unique factors in the safety control's design or execution that has the potential to make a control ineffective at accomplishing its goal. Therefore, the four values in the control effectiveness data structure must be considered together when selecting safety controls for implementation. The current definition of control effectiveness (*CE*) is shown in Table 3. The four variables are defined such that higher values are better.

Table 3. Definition of Control Effectiveness

$CE = \{P_{available}, P_{design}, P_{implementation}, M_{response}\}$			
$P_{available}$ **Probability of Availability (0 to 1)**	P_{design} **Probability of Competent Design (0 to 1)**	$P_{implementation}$ **Probability of Perfect Implementation (0 to 1)**	$M_{response} = 1 - (\frac{t_{sc,affect}}{t_{h,effect}})$ **Response Margin (0 to 1)**
Probability that the implementation strategy is available at the time of control	Probability that the implementation strategy will successfully control the source if it is perfectly implemented.	Probability that the implementation strategy will be implemented perfectly.	Measure of the combined time it takes the disruption to have cascading effects in other subsystems ($t_{h,effect}$) and the time it takes to activate the safety control implementation strategy ($t_{sc,affect}$) $t_{sc,affect} < t_{h,effect}$
Generic Safety Control Flaws that Inform the Estimation of Control Effectiveness			
SAFE CONTROL ACTION IS NOT PROVIDED	SAFE CONTROL ACTION IS PROVIDED TOO MUCH OR TOO LITTLE	SAFE CONTROL ACTION CAUSES UNSAFE CONTROL ACTION SAFE CONTROL ACTION CAUSES HAZARDOUS STATE	SAFE CONTROL ACTIONS IS PROVIDED TOO LATE OR TOO EARLY SAFE CONTROL ACTION IS APPLIED TOO LONG OR STOPPED TOO SOON

The three probabilities developed by Robert Kitching are the same three probabilities in the current *CE* data set. The response margin in the current definition is a new value of control effectiveness that incorporates both the activation time of an implementation strategy, as well as the time to

effect for the corresponding disruption or hazardous state. Time to effect has been directly included in the definition of control effectiveness based on feedback from NASA reviewers during the 2021 RETHi Annual Review. In addition to this change, the modified definition of control effectiveness no longer includes a criticality score for each implementation strategy. This is because the original definition presented redundancies when evaluating control effectiveness based on known control flaws. In the current definition, each control effectiveness value is based on one or more known control flaws, which aids in the evaluation of the probabilities and response margin. In this way, the likelihood and severity of each flaw occurring informs each control effectiveness value, rather than considering these values in isolation.

Probability of Availability

$P_{available}$ is a probability between 0 and 1 that quantifies whether an implementation strategy is available at the time the safety control is needed. Factors that influence this probability include the agent performing the safety control activity (i.e., robot vs. human), since the habitat may not be crewed 100% of the time. For example, in a dormant configuration, an implementation strategy requiring a human agent would automatically have $P_{available} = 0$, while in a crewed configuration, this availability would be higher. Conversely, a robot agent's availability in both crewed and dormant configuration would be the same. This suggests that a habitat supporting both human and robot agent implementations would increase availability in both the crewed and dormant configuration, while a habitat with only one implementation would make the control less available overall. In addition, a control requiring finite consumable resources might be available only once or a few times, compared to a control that uses reusable or renewable resources. In this case, the availability of a control requiring finite consumable resources would decrease throughout the operational life cycle of the habitat, while the control using reusable or renewable resources would maintain the same probability of availability over time (assuming no other wear and tear). Finally, a control's availability might also be influenced by surrounding activities in the habitat. For example, if several safety controls must be completed simultaneously and require the use of the same resources and/or agent, not all controls will be available when needed. Therefore, adequate safety control scheduling and resource distribution must be considered when designing and executing safety controls that overlap in implementation strategy requirements.

Probability of Competent Design

P_{design} is a probability between 0 and 1 that quantifies how well an implementation strategy will successfully control the source of a disruption or hazardous state, assuming it is perfectly implemented. This probability assumes that the safety control is implemented perfectly, and therefore focuses only on the design of the control mechanism and not how well it is implemented during operation. For example, if the metallic habitat structure is breached and needs to be repaired, two possible implementation strategies might be: (1) Human agent applies a flexible patch over the hole or (2) Human agent uses vacuum cementing to fill the hole. In this case, the first strategy would have a lower competency in design, as a flexible patch would not be an appropriate method of repairing a hole in a hard structure, as it is less likely to seal the breach adequately, or permanently. Repairing the hole with a similar hard material is more likely to permanently repair the breach and seal the structure, therefore resulting in a higher probability of competent design.

Probability of Perfect Implementation

$P_{implementation}$ is a probability between 0 and 1 that quantifies how likely it is for an implementation strategy to be implemented perfectly during operation. This probability differs from P_{design} in that it focuses only on the implementation of the control, and not how well it is designed to control the source of the disruption or hazardous state. Factors that influence this probability include the number of agents needed to complete the safety control. Multi-agent tasks are more susceptible to communication errors and inadequate coordination, which increases the likelihood of errors in implementation. In addition, complex or multi-step implementation strategies would be less likely to be implemented perfectly than a simple straightforward task. Finally, the type of agent completing the safety control is a factor, as human and robot agents have different skill sets and limitations. For example, a robot agent attempting to complete a safety control designed for a crew member would result in a lower probability of perfect implementation, compared to a human agent completing the same task. The same might be true for a human agent attempting to complete a task designed specifically for a robot agent.

Response Margin

$M_{response}$ is a value between 0 and 1 that measures the margin between the time it takes a disruption or hazardous state to have measurable effect on performance in the habitat ($t_{h,effect}$) and the time it takes to activate the safety control implementation strategy ($t_{sc,affect}$).

$$M_{response} = 1 - (\frac{t_{sc,affect}}{t_{h,effect}}) \quad (3)$$

This margin requires $t_{sc,affect} < t_{h,effect}$, so that the activation of the safety control occurs before performance degradation propagates to other habitat subsystems. If $t_{sc,affect} > t_{h,effect}$, this control effectiveness value should be labeled as "NO MARGIN" to indicate that the safety control, as designed, is not sufficient to address the disruption or hazardous state before other subsystems are affected. However, note that "NO MARGIN" does not mean that recovery of the system is not possible. It only indicates the need for the immediate activation of the mitigation safety control to address the propagated hazardous states, or the activation of an intervention control to keep the system in a safe state before the mitigation control can be completed. Therefore, response margin favors implementation strategies with smaller $t_{sc,affect}$ and larger $t_{h,effect}$.

The value of $t_{h,effect}$ is not influenced by the safety control implementation strategy, because the time it takes to see performance degradation in the habitat does not depend on the control, only the hazard. The value of $t_{sc,affect}$ does depend on the implementation strategy; it is influenced by factors that cause delay in the activation of the safety control. For example, multi-agent or complex multi-step implementation strategies will likely experience delays due to preparation time or coordination among controllers. In addition, implementation strategies that require agents to travel outside the habitat to external equipment would also experience a time delay due to preparation for exiting the habitat, as well as slower travel times. Controls that require resources or equipment would also require a time delay for a stop in inventory. Finally, we might also consider whether a safety control strategy requires time for approval from ground control. An implementation strategy that can be activated automatically would have a smaller response margin compared to one that needs permission for activation. Although a delay in activation is undesirable, such circumstances might exist for controls that have a potential downside that must be considered (e.g., supplies are

low and would be depleted or the implementation strategy could cause additional damage or an unsafe situation for crew members).

3.4 Developing Guiding Questions to Estimate Control Effectiveness

Evaluating each value in the control effectiveness data structure requires knowledge not only about past accident/incidents, but also current space system designs. Two mission designers working on the same team still might make different assumptions or conclusions about the potential success of a particular safety control based on their own expertise. Such differences would in turn lead to differences in the control effectiveness values assigned, and potential disagreements later in the operations phase. To help standardize the process of designing and evaluating our safety controls, we developed a set of "yes" or "no" guiding questions to answer for each safety control. These questions are based on the control flaw classification developed by Leveson (2004) and are intended to guide designers in thinking about each control's susceptibility to known control flaws. A "no" is preferred over a "yes" for all questions, as that would indicate the safety control is less likely to have the associated control flaw. An answer of "possibly" can also be indicated in answering questions that have the potential to be either "yes" or "no" depending on the specifics of a mission design. This "possibly" should be addressed later in the final design phase to determine whether the deployed design will result in the final "yes" or "no" answer.

Answering these guiding questions will not lead directly to numeric estimates of each control effectiveness value. There is still some subjectivity in the a priori estimates of control effectiveness assigned at this stage, and discussion between mission designers is encouraged to come to an agreement on the values assigned. However, the values determined here allow us to make preliminary selections of safety controls with high, low, or intermediate control effectiveness for implementation and evaluation in the MCVT. **Chapter 4: Validation of Control Effectiveness** discusses the use of these safety control sets in the validation of control effectiveness. Table 4 provides the questions and the relevant control flaws used to guide the assessment of control effectiveness for safety controls in the option space.

Table 4. Guiding Questions to Evaluate Control Effectiveness

Control Effectiveness Variable	**Guiding Questions**	**Relevant Control Flaw(s) from Leveson Classification**
$P_{available}$	Passive Controls Passive safety controls will always be available and therefore $P_{available} = 1$ for all passive safety controls	N/A
	Active Controls Does the control mechanism require equipment and/or material resources from inventory to compete the safety control (e.g., brushes, construction tools, fire extinguisher)?	Time lag
	Does the control mechanism require single-use resources (vs. reusable resources)?	Time lag, flaw(s) in the creation process
	Does the acting body have limited (or no) physical access to the affected area (e.g., safety control is implemented outside the habitat or completing the safety control requires maneuvering in tight spaces)?	Time lag, flaw(s) in the creation process
P_{design}	Passive Controls Is the control mechanism achieved through a design choice prior to system deployment and construction (e.g., structural materials, wall thickness, or equipment installation location)?	Flaw(s) in the creation process
	Is the control mechanism built into the habitat architecture upon deployment (e.g., erecting additional protective layers over the structure or installing thermal protective layers)?	Flaw(s) in the creation process, incorrect installation
	Does the control mechanism require regular maintenance to ensure continued control over relevant safety constraint?	Process changes, incorrect modification or adaptation
	Active Controls Is the control mechanism implemented regularly over the operational lifetime of the system to re-establish control over the relevant safety constraint?	Process changes, incorrect modification or adaptation
	Does the acting body have physical (or design) limitations that may affect the completion of the safety control goal with the available equipment and/or resources in the habitat?	Flaw(s) in the creation process, inadequate actuator operation

Table 4. Continued

	Does the acting body have physical (or design) limitations that may affect interactions with relevant habitat surfaces and therefore hinder the completion of the safety control?	Flaw(s) in the creation process, inadequate actuator operation
	Is the acting body attempting to complete a safety control for which it was not originally designed (vs. adapting its original functionality to accomplish this new safety control goal)?	Flaw(s) in the creation process, incorrect modification or adaptation, incorrect actuator operation
$P_{implementation}$	<u>Passive Control</u> N/A	N/A
	<u>Active Control</u> Does the control mechanism require more than one acting body to complete the safety control?	Inadequate coordination among controllers and decision makers, process models inconsistent, incomplete, or incorrect (lack of linkup, communication flaw, time lag
	If yes to the question above, do the acting bodies need to collaborate (or function as a team) to complete the safety control goal (vs. independently completing identical tasks in different areas)?	Inadequate coordination among controllers and decision makers, process models inconsistent, incomplete, or incorrect (lack of linkup, communication flaw, time lag
	Does one acting body operate or control additional acting bodies in the completion of the safety control goal (e.g., human agent operating a robot agent)?	Inadequate coordination among controllers and decision makers, process models inconsistent, incomplete, or incorrect (lack of linkup, communication flaw, time lag
	Human Agent as Acting Body Does the acting body complete the safety control at regular time intervals during habitat operation (e.g., for a regular maintenance activity)?	Process changes, incorrect modification or adaptation
	Does the acting body have to travel outside the habitat to complete the safety control?	Time lag
	Does the acting body require training on Earth to be able to complete the safety control goal (e.g., specialized engineering or construction skills)?	Flaw(s) in creation process, inadequate actuator operation

Table 4. Continued

	Does the acting body require specific expertise to complete the safety control?	Flaw(s) in creation process, inadequate actuator operation
	Does the acting body require an instruction manual to complete the safety control?	Flaw(s) in creation process, inadequate actuator operation, process models inconsistent, incomplete, or incorrect (lack of linkup)
	Does the acing body interact with automated systems during the completion of the safety control?	Process models inconsistent, incomplete, or incorrect (lack of linkup)
	Is the acting body taking commands from another acting body or supervisor throughout the completion of the safety control?	Inadequate coordination among controllers and decision makers, process models inconsistent, incomplete, or incorrect (lack of linkup, communication flaw, inadequate or missing feedback
	Is the acting body required to visually assess the affected area throughout the completion of the safety control?	Inadequate or missing feedback
	Is the acting body required to analyzed sensor data and feedback throughout the completion of the safety control?	Inadequate or missing feedback, Process models inconsistent, incomplete, or incorrect (lack of linkup)
	If yes to the question above, are the necessary sensors functioning and sending data at regular intervals?	Inadequate or missing feedback
	Robot Agent as Acting Body Does the acting body complete the safety control at regular time intervals during habitat operation (e.g., for a regular maintenance activity)?	Process changes, incorrect modifications or adaptation
	Is the acting body pre-programmed to complete the safety control goal (vs. planning actions/routes in real time)?	Process models inconsistent, incomplete, or incorrect (lack of linkup)
	If yes to the question above, does the acting body require regular software updates (or habitat status updates) to continue to correctly complete the safety control goal?	Process models inconsistent, incomplete, or incorrect (lack of linkup)

Table 4. Continued

	Does the acting body have to travel outside the habitat to complete the safety control?	Time lag
	Does the acing body interact with automated systems during the completion of the safety control?	Process models inconsistent, incomplete, or incorrect (lack of linkup)
	Is the acting body required to visually assess the affected area throughout the completion of the safety control?	Inadequate or missing feedback
	Is the acting body required to analyzed sensor data and feedback throughout the completion of the safety control?	Inadequate or missing feedback, Process models inconsistent, incomplete, or incorrect (lack of linkup)
	If yes to the question above, are the necessary sensors functioning and sending data at regular intervals?	Inadequate or missing feedback
	Automation as Acting Body Does the acting body complete the safety control goal at a regular time interval during habitat operation?	Process changes, incorrect modification or adaptation
	Does the acting body complete the safety control goal in response to sensor data and feedback on habitat performance during operation?	Inadequate or missing feedback
	If yes to the question above, does the acting body require regular software updates (or habitat status updates) to continue to correctly complete the safety control goal?	Process models inconsistent, incomplete, or incorrect (lack of linkup)
	Does the acting body require regular monitoring from an additional acting body to verify output during the completion of the safety control goal?	Process models inconsistent, incomplete, or incorrect (lack of linkup)
	Does the acting body have an override or off-switch that allows another acting body to manually take over completion of the safety control goal?	Process models inconsistent, incomplete, or incorrect (lack of linkup)
$M_{response}$	<u>Passive Control</u> N/A	N/A

Table 4. Continued

	Active Control The value of $t_{sc,affect}$ will be influenced by the answers to the questions above associated with the control flaw "Time Lag". No additional guiding questions are provided for this control effectiveness parameter.	Time Lag

3.4.1 Color-Coding Control Effectiveness

We want to use control effectiveness to down select safety controls that will create habitat architectures with high resilience. Control effectiveness has been defined so that those controls with high control effectiveness should lead to architectures with high resilience. To help us select those controls with high, low, or intermediate control effectiveness, we created a color-coded visualization based on the values in the CE data set. The specific color-mapping may be modified according to a user's risk preference, and the color-coded visualization is intended to aid designers in their evaluation of the control effectiveness data. Table 5 shows an example of a control effectiveness color-coding scheme.

Table 5. Example Control Effectiveness Color-Coding Scheme

Control Effectiveness Variable	**Units**	**Range**
$P_{available}$	Probability (0 to 1)	$0.7 \leq P_{available} \leq 1$
		$0.3 < P_{available} \leq 0.7$
		$0 < P_{available} \leq 0.3$
P_{design}	Probability (0 to 1)	$0.7 \leq P_{design} \leq 1$
		$0.3 < P_{design} \leq 0.7$
		$0 < P_{design} \leq 1$
$P_{implementation}$	Probability (0 to 1)	$0.7 \leq P_{implementation} \leq 1$
		$0.3 < P_{implementation} \leq 0.7$
		$0 < P_{implementation} \leq 0.3$
$M_{response}$	Margin (0 to 1)	$0.7 \leq M_{response} \leq 1$
		$0.3 < M_{response} \leq 0.7$
		$0 < M_{response} \leq 0.3$

Using a color-coding scheme such as the one in Table 5 gives a visual representation of how effective each safety control may be. "Green" probabilities or response margin indicates higher control effectiveness, while "red" indicates lower control effectiveness. The cut-off for each color can be modified to reflect the desired risk averseness. For example, a more risk averse color-coding scheme would move the thresholds higher to indicate a narrower window of risk acceptance for the safety controls, as shown in Table 6.

Table 6. Example "Risk Averse" Control Effectiveness Color-Coding Scheme

Control Effectiveness Variable	Units	Range
$P_{available}$	Probability (0 to 1)	$0.85 \leq P_{available} \leq 1$
		$0.45 < P_{available} \leq 0.85$
		$0 < P_{available} \leq 0.45$
P_{design}	Probability (0 to 1)	$0.85 \leq P_{design} \leq 1$
		$0.45 < P_{design} \leq 0.85$
		$0 < P_{design} \leq 0.45$
$P_{implementation}$	Probability (0 to 1)	$0.85 \leq P_{implementation} \leq 1$
		$0.45 < P_{implementation} \leq 0.85$
		$0 < P_{implementation} \leq 0.45$
$M_{response}$	Margin (0 to 1)	$0.85 \leq M_{response} \leq 1$
		$0.45 < M_{response} \leq 0.85$
		$0 < M_{response} \leq 0.45$

By organizing our safety controls into a table and assigning the color based on the chosen color-coding scheme, designers can then select desired controls for implementation in the MCVT to observe how the system responds with different sets of safety controls. In the following section, we apply this control effectiveness method to the same disruption and hazardous state example used in **Section 2.2** to describe the State and Trigger Model.

3.5 Application: Assessing Safety Controls for Example Disruption and Hazardous State Scenario

In this section, we use the control effectiveness metric to evaluate the available safety controls to address the dust accumulation hazardous state discussed in **Section 2.2** to depict the State and Trigger Model (Figure 5). This example scenario includes one or more preventive, mitigative, and

interventive safety controls from the database. The first step in evaluating the effectiveness of these control strategies is to identify several implementation strategies that describe how each safety control goal is achieved. The safety controls and corresponding implementation strategies for this scenario are shown in Table 7.

Table 7. Safety Control Implementation Strategies for Dust Accumulation Hazardous State

Hazardous State: Degradation in solar power generation due to dust accumulation on solar PV arrays			
Prevention Controls	**Implementation 1**	**Implementation 2**	**Implementation 3**
Ability to cover solar PV arrays (EXTRA PROTECTION FROM SOURCE)	Human agent secures cover over solar PV arrays	Robot agent secures cover over solar PV arrays	Built-in covers are activated to roll out over solar PV arrays
Ability to angle solar PV arrays away from incoming dust (COMPONENT CORRECTS FOR SOURCE)	Human agent manually rotates solar PV arrays using mechanical lever	Human agent remotely rotates solar PV arrays through system command in habitat control	
Mitigation Controls	**Implementation 1**	**Implementation 2**	**Implementation 3**
Ability to remove dust from solar PV arrays (REMOVE SOURCE FROM COMPONENT)	Human agent brushes dust from solar PV arrays	Robot agent brushes dust from solar PV arrays	Built-in brush automatically removes dust from solar PV arrays
Intervention Controls	**Implementation 1**	**Implementation 2**	**Implementation 3**
Ability to use backup nuclear power system (REDUNDANT COMPONENT SYSTEM)	Human agent activates nuclear power generation system in addition to solar power generation	Automatic activation of nuclear power generation in addition to solar power generation	
Ability to backup battery power (REDUNDANT COMPONENT SYSTEM)	Human agent activates battery power in addition to solar power generation	Automatic activation of battery power in addition to solar power generation	

To evaluate control effectiveness for the controls in our example, we must now consider each implementation strategy individually and answer the provided guiding questions. In answering these questions, we must make several assumptions for consistency in evaluating each safety control. The assumptions discussed here will carry through not only for this example, but for the

evaluation of all safety controls in MCVT v6, which is discussed more in **Chapter 4: Developing a Control Effectiveness Validation Plan**.

We first assume that a crew would only be present for 70% of the operational lifetime of the habitat, meaning implementation strategies requiring a human agent automatically have a maximum probability of availability of 0.7. This value may be lower for a particular implementation strategy due to factors such as performance limitations or scheduling conflicts. We also make the assumption that robot agents will be present for 100% of the operational lifetime; however, availability may be affected by similar factors as human agent.

In addition, when estimating the response margins, we must run the MCVT with no safety controls to obtain estimates of $t_{h,effect}$. The value of $t_{h,effect}$ in the context of this disruption and hazardous state was the time it took for a 10% drop in the solar power output due to dust coverage. The identification of $t_{h,effect}$ for all other safety controls should be the time it takes to have performance degrade past the operational threshold (set point) of the health management system. These thresholds are provided in **Chapter 4** for all safety controls in the MCVT. For this example, we found $t_{h,effect} = 25.114$ s. Note that the values of $t_{h,effect}$ in the MCVT will be very small, as the MCVT currently only supports simulations on the order of minutes due to computational limitations. Therefore, activation times ($t_{sc,affect}$) must be appropriately scaled to enable activation and full recovery for each control within 120 seconds. However, estimating activation time realistically is still necessary before scaling and computing the final response margin. To do so, we use the layout of the MCVT to obtain distance estimates from inventory (starting location of the agent) to all locations in the habitat. In this example, the distance to solar PV arrays is 59.8 m. In addition, we must make several other assumptions for consistency in estimating activation time based on the agent type and equipment required for each implementation strategy:

1. A human agent will be assigned a travel speed of 1.39 m/s (Choi, 2014), and a robot agent will be assigned a travel speed of 0.05 m/s (NASA, n.d.).
2. If an implementation strategy requires equipment or resources in the habitat inventory, a fixed 5-minute time is added to the human agent activation time, and a fixed 10-minute time is added to the robot agent activation time to account for equipment pick up.

3. If an implementation strategy requires an agent to travel outside, a fixed 10-minute time is added to the human agent activation time to account for preparation.
4. If an implementation strategy uses automated system detection and response, a fixed 2-minute time is used for the automation activation time.

These assumptions allow us to consistently estimate the response margin for adequate comparison of this control effectiveness dimension for all implementation strategies. The final control effectiveness values identified for each implementation strategy are provided in Figures 7–9, followed by a summary of the main findings.

Hazardous State: Degradation in solar power generation due to dust accumulation on solar PV arrays												
Prevention Controls	**Implementation 1**				**Implementation 2**				**Implementation 3**			
Ability to cover solar PV arrays	Ability for human agent to secure cover over solar PV arrays				Ability for robot agent to secure cover over solar PV arrays				Built-in covers are activated to roll out over solar PV arrays			
	$P_{available}$ 0.5	P_{design} 0.9	$P_{implementation}$ 0.7	$M_{response}$ 0.4261	$P_{available}$ 0.7	P_{design} 0.9	$P_{implementation}$ 0.3	$M_{response}$ NO MARGIN	$P_{available}$ 0.95	P_{design} 0.9	$P_{implementation}$ 0.95	$M_{response}$ 0.97
Ability to angle solar PV arrays away from incoming dust	Human agent rotates solar PV arrays using mechanical lever				Human agent remotely rotates solar PV array through system command in habitat control							
	$P_{available}$ 0.5	P_{design} 0.8	$P_{implementation}$ 0.9	$M_{response}$ 0.5928	$P_{available}$ 0.9	P_{design} 0.8	$P_{implementation}$ 0.95	$M_{response}$ 0.97				

Figure 7. Control Effectiveness Results for Prevention Safety Controls

Hazardous State: Degradation in solar power generation due to dust accumulation on solar PV arrays												
Mitigation Controls	**Implementation 1**				**Implementation 2**				**Implementation 3**			
Ability to remove dust from solar PV arrays	Human agent brushes dust from solar PV arrays				Robot agent brushes dust from solar PV arrays				Built-in brush automatically removes dust from solar PV arrays			
	$P_{available}$ 0.5	P_{design} 0.9	$P_{implementation}$ 0.7	$M_{response}$ 0.4261	$P_{available}$ 0.7	P_{design} 0.9	$P_{implementation}$ 0.4	$M_{response}$ NO MARGIN	$P_{available}$ 0.9	P_{design} 0.9	$P_{implementation}$ 0.9	$M_{response}$ 0.97

Figure 8. Control Effectiveness Results for Mitigation Safety Controls

Hazardous State: Degradation in solar power generation due to dust accumulation on solar PV arrays								
Intervention Controls	**Implementation 1**				**Implementation 2**			
Ability to sure backup nuclear power system	Human agent activates nuclear power generation system in addition to solar power generation				Automatic activation of nuclear power generation in addition to solar power generation			
	$P_{available}$ 0.5	P_{design} 0.9	$P_{implementation}$ 0.9	$M_{response}$ 0.83	$P_{available}$ 0.95	P_{design} 0.9	$P_{implementation}$ 0.95	$M_{response}$ 0.97
Ability to use backup battery power	Human agent activates battery power in addition to solar power generation				Automatic activation of battery power in addition to solar power generation			
	$P_{available}$ 0.5	P_{design} 0.9	$P_{implementation}$ 0.9	$M_{response}$ 0.83	$P_{available}$ 0.95	P_{design} 0.9	$P_{implementation}$ 0.95	$M_{response}$ 0.97

Figure 9. Control Effectiveness Results for Intervention Safety Controls

From the results in Figures 7, 8, and 9, we can now answer the following three questions for this example scenario: (1) How should we implement each safety control? (2) How might each safety control/implementation strategy be flawed? (3) Based on all available safety controls and considered implementation strategies, what control has the potential to be the most effective?

In Figure 7, we have two available prevention controls: *Ability to cover solar PV arrays* and *Ability to angle solar PV arrays away from incoming dust*. Based on the control effectiveness values for each, the implementation strategies which rely on automation have the highest control effectiveness. These are: *Built-in covers are activated to roll out over solar PV arrays* and *Human agent remotely rotates solar PV arrays through system command in habitat control*. These controls have all green control effectiveness values, indicating that their availability, design, execution, and response margins are less likely to be flawed than the other implementation strategies. This is largely because these strategies do not require an agent to travel outside the habitat to manually complete the tasks, which significantly reduces the activation time and avoids potential challenges or dangers for crew members or autonomous robots. Based on the control effectiveness of the agent dependent strategies, we can conclude that the robot agent strategy has the lowest control effectiveness and has the most susceptibility to control flaws. The human agent strategy then has intermediate control effectiveness, because it is better than the robot agent, but still worse than the automated implementation. Therefore, we can conclude the automation implementation strategies may be the most effective and represent the highest control effectiveness options for these prevention safety controls.

In Figure 8, we have one available mitigation control: *Ability to remove dust from solar PV arrays*. Once again, the implementation strategy that relies on automation has the highest control effectiveness: *Ability for built-in brush to remove dust from solar PV arrays*. This implementation also has all green control effectiveness values, while the other two strategies are lacking in one or more of the control effectiveness values. As with the prevention control, the response margins of the implementations relying on human or robot agents are lower, with the robot agent having no response margin at all. This means the robot cannot respond quickly enough to avoid at least a 10% drop in solar power output. Robots in general cannot travel as quickly as humans, especially when there are obstacles. However, human agents must suit up for outdoor activities, meaning they

also have a low response margin. In addition to the lower response margin, the implementation of this safety control is difficult for a robot agent because the solar PV arrays are raised off the ground, and assuming the brush is long enough to reach the entire surface, manipulating such an object would be difficult for a robot agent. Therefore, like the prevention controls, we can conclude that the robot agent strategy has the lowest control effectiveness and is the most susceptible to control flaws. The human agent strategy then has intermediate control effectiveness, as it performs better than the robot agent but still worse than the automated implementation. Therefore, the automated implementation strategy with built-in brush on the solar PV arrays has the potential to be the most effective and represents the highest control effectiveness option for this mitigation control.

In Figure 9, we have two available intervention controls: *Ability to use backup nuclear power generation system* and *Ability to use backup battery power.* Based on the control effectiveness values for the two implementations, we once again find that the automated implementation strategies offer higher control effectiveness compared to the human agent implementation strategy. These are: *Automatic activation of nuclear power generation system in addition to solar power generation* and *Automatic activation of battery power in addition to solar power generation.* Although all response margins for these implementations and the human agent implementation are green, we see that the response margin with a human agent is lower compared to an automated activation, which is to be expected. In addition, since we assumed that a human agent is not present 100% of the operational lifetime, the availability of these implementation strategies is lower compared to the automated strategy, which is always available, save for complete power outages or system faults. Therefore, the automated implementation strategies have the potential to be most effective and represent the highest control effectiveness options for these intervention controls.

In **Chapter 4: Developing a Control Effectiveness Validation Plan**, we discuss the development of the MCVT to allow for the evaluation of different control implementation strategies with varying control effectiveness. The identification of high, low, or intermediate control effectiveness implementation strategies (as discussed in this example) will be important for the selection of safety controls for implementation in the MCVT, and ultimately, for understanding whether control effectiveness is appropriately defined to select high control effectiveness safety controls that will lead to resilient architectures.

4. DEVELOPING A CONTROL EFFECTIVENESS VALIDATION PLAN

The validation of control effectiveness is essential in determining whether the definition of control effectiveness presented in **Chapter 3** is appropriate to make selections of safety controls that will lead to desired habitat resilience. The validation of this metric is also a necessary step in validating our overall control-theoretic approach to resilience. This chapter focuses on the development and implementation of a validation procedure for control effectiveness, which maps to *Step 4: Safety Control Assessment* in our control-theoretic approach.

4.1 The Four Step Validation Cycle

Our control effectiveness validation approach is iterative and relies on the MCVT for the simulation of habitat response under various environmental and operational hazards. Therefore, a primary part of developing and implementing the validation procedure includes the development of disruption scenarios that simulate a wide range of disruption, hazardous state, and safety control combinations from the database. In particular, the MCVT must support functionalities that will allow us to study habitat response under sets of safety controls with varying control effectiveness values. As such, in parallel with developing these scenarios, the functional requirements for MCVT subsystems must be regularly expanded to incorporate capabilities to model new disruptions, hazardous states, and safety controls under consideration. Therefore, the control effectiveness validation procedure is not only iterative based on the correlation between control effectiveness and resilience, but also on the MCVT expansion process, which is intertwined with each of the four steps in the validation procedure shown in Figure 10. In the following sections, we discuss each step in the validation procedure, as well as the MCVT development tasks that must be completed as a part of each step.

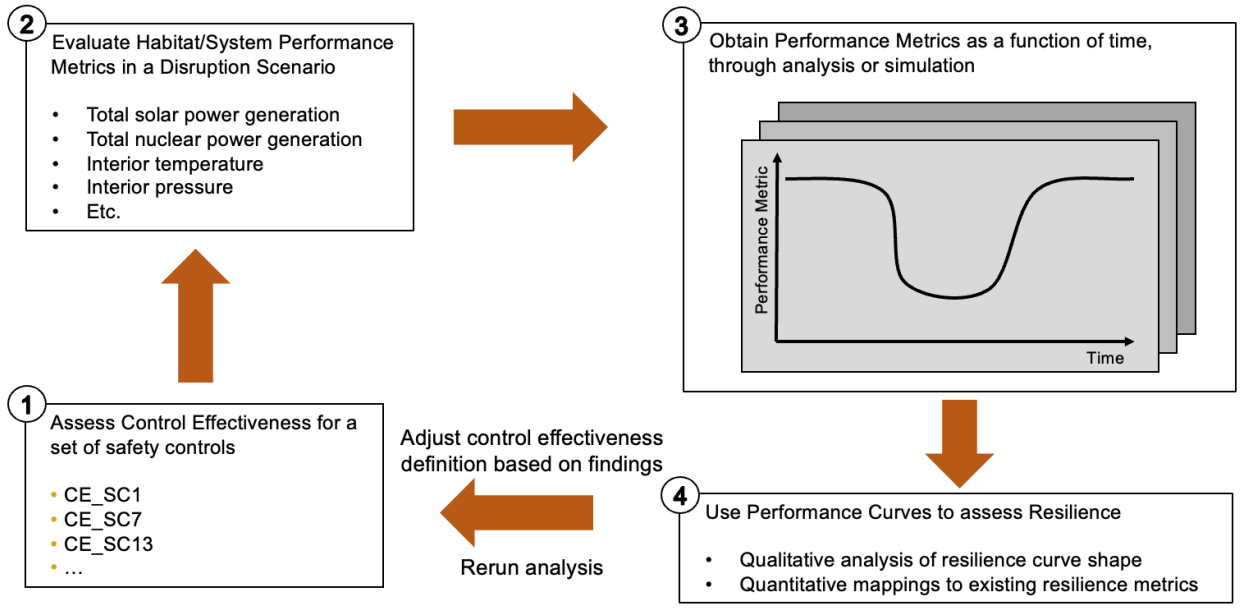

Figure 10. Control Effectiveness Validation Procedure

4.1.1 Assess Control Effectiveness for a Set of Safety Controls

Step 1 is to assess control effectiveness for sets of safety controls that are modeled in the MCVT. The process of evaluating control effectiveness and assigning an appropriate color-coding scheme has been discussed in **Chapter 3**. The first task of this step in the validation procedure is identifying and developing disruption scenarios that are representative of the disruption, hazardous state, and safety control combinations we are interested in studying in the MCVT to answer research questions. Since the MCVT platform is intended to answer research questions in all three RETHi research thrusts, these disruption scenarios were brainstormed and developed with faculty across all three thrusts. The main requirement for the disruption scenarios in the context of this thesis is that the disruption scenarios must be complex enough to allow us to explore a wide range of safety controls with varying control effectiveness values. This will allow us to determine whether high control effectiveness safety controls (and sets of safety controls) are appropriate to create high resilience habitat architectures.

Figure 11 shows the general structure of disruption scenarios built from combinations of disruptions, hazardous states, and safety controls in the database. The sequence of events that comprises a disruption scenario is as follows:

1. A disruption occurs
2. The effects of the disruption cause one or more hazardous states
3. Hazardous states can be resolved by safety controls
4. Safety controls improve subsystem performance, resolving the hazardous state, and returning the system to a nominal state.

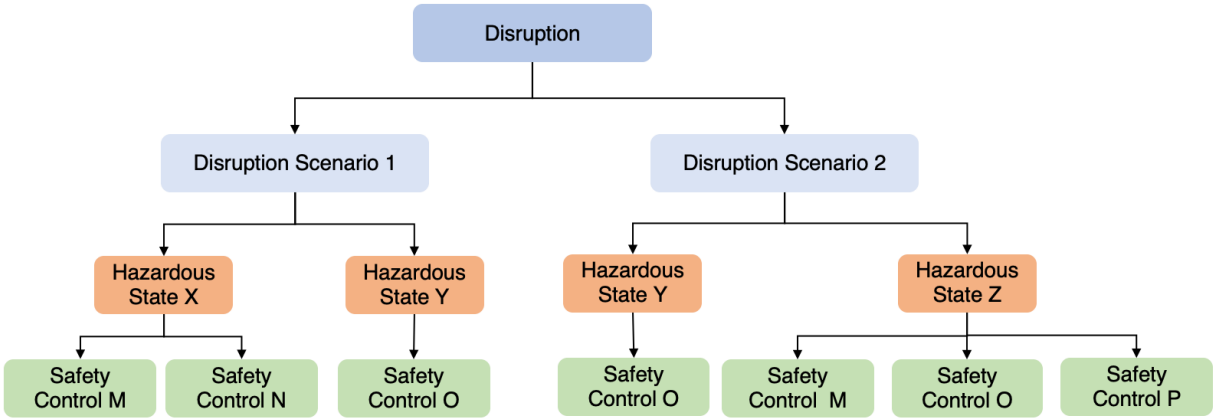

Figure 11. General Structure of Events for a Disruption Scenario

Note that in Figure 11, a single disruption can have one or more disruption scenarios. For example, a meteorite impact is a disruption, but the impact location can cause different disruption scenarios because the combinations of hazardous states resulting from an impact on the structure vs. on the solar PV arrays are different. In addition, the same hazardous state may occur in multiple disruption scenarios (hazardous state Y in Figure 11), and safety controls can overlap between multiple hazardous states in any disruption scenario (safety control M and O in Figure 11).

As previously mentioned, Purdue alumnus, Robert Kitching, completed a hazard assessment as a part of Steps 1–3 in our control-theoretic approach. The results of that assessment, in combination with RETHi interests across all three research thrusts, allowed us to identify six disruptions that offer a wide range of disruption scenarios to model in the MCVT. These disruptions are: (1) Meteorite Impact, (2) Moonquake, (3), Fire in Habitat Interior Environment, (4) Nominal Dust Accumulation on Exterior Habitat Systems, (5), Launch/Landing Dust Accumulation on Exterior Habitat Systems, and (6) Communication and Sensor Network Failure. For these six disruptions,

there are a total of 16 disruption scenarios that include various combinations of 117 hazardous states (97 of which are failures of different installed sensors), and 25 safety controls to address the hazardous states.

The six disruptions modeled in MCVT v6 can originate in one or more habitat subsystems and impact one or more secondary subsystems to cause hazardous states. To visualize the physical interdependencies in the MCVT and how disruptions propagate in our disruption scenarios, we created a general disruption propagation matrix to show the subsystem(s) in which each disruption can originate and the subsystem(s) into which effects can propagate (Figure 12).

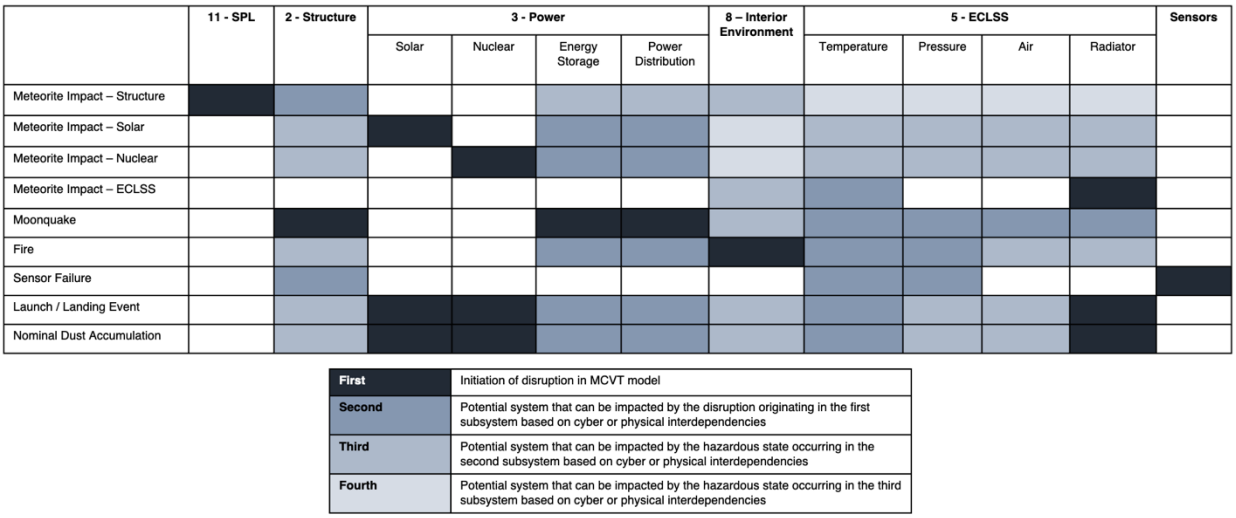

Figure 12. General Disruption Propagation Matrix for MCVT v6

Figure 12 depicts the propagation of disruptions down through four levels of cascading effects, where each lower-level effect is caused by one or more of the effects in the level directly above. As previously stated, the MCVT can model up to 16 unique disruption scenarios. The meteorite impact disruption has 4 disruption scenarios, as shown in Figure 12. A user can set an impact event at the structure, external solar power generation system, external nuclear power generation system, or external ECLSS thermal control system (radiator panels). The moonquake disruption has one scenario, as the moonquake can only originate in a fixed set of subsystems to cause the same hazardous states. The fire disruption has three scenarios, which depend on the start location of the fire in the interior environment. Note that in Figure 12 there is only one row for the fire disruption

because all fire events originate in the interior environment in the MCVT. However, there is the potential for causing three unique scenarios based on whether the fire spreads to the internal power storage and distribution equipment, or ECLSS thermal and pressure control systems. The sensor failure disruption has one scenario, as the same sensor failure and hazardous state can occur for all installed sensors. The launch/landing disruption has one scenario because the location of this disruption is fixed and causes the same hazardous states. Finally, the nominal dust accumulation disruption has six scenarios because users can control whether dust accumulates on the solar PV arrays, nuclear radiator panels, and/or ECLSS radiator panels. Dust can accumulate on one, two, or all three of these surfaces in different combinations. Here we consider dust accumulation on all three surfaces as a redundant scenario that is captured in the effects from a meteorite impact. Therefore, nominal dust accumulation provides six more unique disruption scenarios.

Table 8 lists all 117 hazardous states and 25 safety controls associated with the six disruptions modeled in MCVT v6 that combine in different ways to form the 16 disruption scenarios discussed above.

Table 8. Disruptions, Hazardous States, and Safety Controls Modeled in MCVT v6

Disruption/Failure	**Hazardous State**	**Safety Control(s)**
Launch/Landing Event	Solar PV arrays are covered by dust (**HS100**)	Ability to remove dust from solar PV arrays (**SC798**)
		Ability to activate secondary power generation system (**SC11**)
		Ability to activate battery power as power generation source (**SC822**)
	Nuclear radiator panels are covered by dust (**HS213**)	Ability to remove dust from nuclear radiator panels (**SC799**)
		Ability to activate secondary power generation system (**SC11**)
		Ability to activate battery power as power generation source (**SC822**)
	ECLSS radiator panels are covered by dust (**HS216**)	Ability to remove dust from ECLSS radiator panels (**SC785**)
Dust External to Habitat	Solar PV arrays are covered by dust (**HS100**)	Ability to remove dust from solar PV arrays (**SC798**)
		Ability to activate secondary power generation system (**SC11**)
		Ability to activate battery power as power generation source (**SC822**)
	Nuclear radiator panels are covered by dust (**HS213**)	Ability to remove dust from nuclear radiator panels (**SC799**)

Table 8. Continued

		Ability to activate secondary power generation system (**SC11**)
		Ability to activate battery power as power generation source (**SC822**)
	ECLSS radiator panels are covered by dust (**HS216**)	Ability to remove dust from ECLSS radiator panels (**SC785**)
	Paint degradation on ECLSS radiator panels (**HS217**)	Ability to repair paint damage on ECLSS radiator panels (**SC786**)
Meteorite Impact	Habitat structural mechanical layer is breached (**HS127**)	Ability to repair the structural mechanical layer (**SC795**)
		Ability to regulate temperature of interior environment (**SC823**)
		Ability to regulate pressure of interior environment (**SC824**)
	Habitat structural protective layer is breached (**HS38**)	Ability to repair the structural protective layer (**SC796**)
	Solar PV arrays are damaged (**HS41**)	Ability to replace solar PV arrays (**SC799**)
	Nuclear radiator panels are damaged (**HS134**)	Ability to replace nuclear radiator panels (**SC801**)
	Solar PV arrays are covered by dust (**HS100**)	Ability to remove dust from solar PV arrays (**SC798**)
		Ability to activate secondary power generation system (**SC11**)
		Ability to activate battery power as power generation source (**SC822**)
	Nuclear radiator panels are covered by dust (**HS213**)	Ability to remove dust from nuclear radiator panels (**SC799**)
		Ability to activate secondary power generation system (**SC11**)
		Ability to activate battery power as power generation source (**SC822**)
	ECLSS radiator panels are covered by dust (**HS216**)	Ability to remove dust from ECLSS radiator panels (**SC785**)
	Solar power distribution converters are damaged (**HS227**)	Ability to repair individual power converters (**SC802**)
	Solar power distribution main generation bus is damaged (**HS228**)	Ability to repair main power generation bus (**SC803**)
	Energy storage system is damaged (**HS35**)	Ability to repair battery cells (**SC355**)
Moonquake	Smart power distribution converters are damaged (**HS227**)	Ability to repair individual power converters (**SC802**)
	Smart power distribution main bus is damaged (**HS228**)	Ability to repair main power generation bus (**SC803**)
	Energy storage system is damaged (**HS35**)	Ability to repair battery cells (**SC355**)
Fire Internal to the Habitat	Open fire in interior environment (**HS71**)	Ability to extinguish active fire in interior environment (**SC797**)
		Ability to regulate temperature of interior environment (**SC823**)
		Ability to regulate pressure of interior environment (**SC824**)

Table 8. Continued

	ECLSS air tank has a leak (**HS219**)	Ability to repair piping between the air tank in ECLSS pressure system and interior environment (**SC788**)
	ECLSS compressor performance is decreased (**HS222**)	Ability to repair the compressor in ECLSS thermal system (**SC791**)
Communication and Sensor Network Failure	Sensor(s) in Subsystem X* experience sensor drift	Ability to repair sensor drift in sensor(s) in Subsystem X**
	Sensor(s) in Subsystem X* fail	Ability to repair failed sensor(s) in Subsystem X**
	Sensor(s) in Subsystem X* experience simultaneous drift and failure	Ability to repair sensor(s) in Subsystem X**
Failures that can result from damage levels of other subsystems	ECLSS fan has buildup of dust in filter (**HS218**)	Ability to remove dust from fan in ECLSS pressure system (**SC787**)
	Failures from Interior Environment damage index	
	ECLSS air supply valve is malfunctioning (**HS220**)	Ability to repair the air supply valve in ECLSS pressure system (**SC789**)
	ECLSS evaporator has air side leak (**HS223**)	Ability to repair the evaporator air side leak in ECLSS thermal system (**SC792**)
	Failures from power storage damage index	
	ECLSS pressure system consumes excess power (**HS221**)	Ability to repair power consumption fault in ECLSS pressure system (**SC790**)
	ECLSS thermal system consumes excess power (**HS225**)	Ability to repair power consumption fault in ECLSS thermal system (**SC794**)
	Failures from power distribution damage index	
	ECLSS heater performance is decreased (**HS224**)	Ability to repair the heater in ECLSS thermal system (**SC793**)

** 97 subsystem sensors that can be affected are: ECLSS Exterior – temperature sensor (HS229-HS231, ECLSS Pressure – flowmeter (HS244-HS246), ECLSS Thermal – temperature sensor (HS229-HS230), Structural Mechanical/SPL – accelerometer (HS241-HS243), Interior Environment – temperature sensor (HS229-HS231), Power Generation – power meter (HS232-HS234) and temperature sensor (HS229-HS231), Power Storage – power meter (HS232-HS234), temperature sensor (HS229-HS231), current sensor (HS235-HS237), and charge reader (HS238-HS240)*

*** 97 Subsystems sensors that can be repaired are: ECLSS Exterior – temperature sensor (SC804-SC806), ECLSS Pressure – flowmeter (SC819-SC821), ECLSS Thermal – temperature sensor (SC804-SC806), Structural Mechanical/SPL – accelerometer (SC816-SC818), Interior Environment – temperature sensor (SC804-SC806), Power Generation – power meter (SC807-SC809) and temperature sensor (SC804-SC806), Power Storage – power meter (SC807-SC809), temperature sensor (SC804-SC806), current sensor (SC810-SC812), and charge reader (SC813-SC815)*

After finalizing all disruption scenarios modeled in MCVT v6, the next task in completing Step 1 of this validation procedure is assessing control effectiveness for all the safety controls listed in Table 8. To do so, we have to create possible implementation strategies for all controls. Implementation strategies can be developed to reflect changes in who or what completes the control action (human or robot) and how they complete the control action (change in method or tools). This will influence the value of control effectiveness for the control activity. In developing these strategies and answering the guiding questions shown in **Chapter 3**, it was necessary to discuss particular controls with subsystem modelers in the Resilience Thrust to gather more information on how the control activities may be accomplished in a hypothetical deep space habitat. In addition, discussions with the Robotics Thrust provided context for capabilities and limitations for a robot agent completing control activities. Therefore, the implementation strategies developed for available controls in MCVT v6 are shown in Table 9 and reflect appropriate potential methods of completing the control activities in a deep space habitat.

Table 9. Implementation Strategies for Disruption Scenario Safety Controls

Safety Control	**Implementation Strategy 1**	**Implementation Strategy 2**	**Implementation Strategy 3**
Ability to remove dust from solar PV arrays	Human agent brushes dust from solar PV arrays	Robot agent brushes dust from solar PV arrays	Built-in brush removes dust from solar PV arrays
Ability to remove dust from nuclear radiator panels	Human agent brushes dust from nuclear radiator panels	Robot agent brushes dust from nuclear radiator panels	Built-in rush automatically removes dust from nuclear radiator panels
Ability to activate secondary power generation system	Smart power distribution automatically activates secondary generation system		
Ability to activate battery power	Smart power distribution automatically draws power from battery storage		
Ability to remove dust from ECLSS radiator panels	Human agent brushes dust from ECLSS radiator panels	Robot agent brushes dust from ECLSS radiator panels	Built-in rush automatically removes dust from ECLSS radiator panels
Ability to repair paint damage on ECLSS radiator panels	Human agent applies additional paint layer to ECLSS radiator panels	Robot agent applies additional paint layer to ECLSS radiator panels	
Ability to repair breach in structural mechanical layer	Human agent uses vacuum cementing to fill breach in structure	Human agent applies flexible patch over breach in structure	Robot agent applies flexible patch over breach in structure

Table 9. Continued

Ability to repair breach in structural protective layer	Human agent replaces and compacts regolith to fill breach in structural protective layer	Robot agent replaces and compacts regolith to fill breach in structural protective layer	
Ability to regulate the temperature of the interior environment	Temperature control system automatically heats/cools the interior environment		
Ability to regulate the pressure of the interior environment	Pressure control system automatically increases/decreases pressure of interior environment		
Ability to replace individual power converters	Human agent replaces power converter	Robot agent replaces power converter	
Ability to replace main power generation bus	Human agent replaces power generation bus	Robot agent replaces power generation bus	
Ability to replace solar PV arrays	Human agent replaces solar PV arrays	Robot agent replaces solar PV arrays	
Ability to replace nuclear radiator panels	Human agent replaces nuclear radiator panels	Robot agent replaces nuclear radiator panels	
Ability to replace energy storing units	Human agent replaces energy storing units	Robot agent replaces energy storing units	
Ability to extinguish active fire in interior environment	Human agent uses a fire extinguisher to put out fire	Human agent uses fire blanket to put out fire	Robot agent uses fire blanket to put out fire
Ability to remove dust from fan in ECLSS pressure system	Human agent removes dust from fan filter in ECLSS pressure system	Robot agent removes dust from fan filter in ECLSS pressure system	
Ability to repair piping between the air tank in ECLSS pressure system and interior environment	Human agent applies flexible patch over leak in air tank of ECLSS pressure system	Human agent applies hardening putty over leak in air tank of ECLSS pressure system	Human agent replaces and re solders a new pipe
Ability to repair the air supply valve in ECLSS pressure system	Human agent replaces the air supply valve in ECLSS pressure system	Robot agent replaces the air supply valve in ECLSS pressure system	
Ability to repair the compressor in ECLSS thermal system	Human agent replaces the compressor in ECLSS thermal system		
Ability to repair the evaporator in ECLSS thermal system	Human agent replaces the evaporator in ECLSS thermal system		
Ability to repair the heater in ECLSS thermal system	Human agent replaces the heater in ECLSS thermal system		

Using the guiding questions and color-coded scheme discussed in **Chapter 3**, we assigned control effectiveness values to the safety controls in MCVT v6. Figures 13–32 present the control

effectiveness values for all safety control implementations in Table 9. These control effectiveness values allow us to identify individual high, low, or intermediate control effectiveness options and form sets of all high, all low, or mixed control effectiveness sets for implementation in the MCVT v6. The next step in the validation plan is then identifying relevant system/habitat performance metrics for each safety control that will allow us to characterize how systems and/or the habitat is performing relative to desired objectives. This is discussed in **Section 4.1.2. Identify Relevant Performance Metrics**.

Hazardous State: Solar PV arrays are covered by dust												
Mitigation Controls	**Implementation 1**				**Implementation 2**				**Implementation 3**			
Ability to remove dust from solar PV arrays	Human agent brushes dust from solar PV arrays				Robot agent brushes dust from solar PV arrays				Built-in brush automatically removes dust from solar PV arrays			
$P_{available}$ 0.5	P_{design} 0.9	$P_{implementation}$ 0.7	$M_{response}$ 0.804	$P_{available}$ 0.7	P_{design} 0.9	$P_{implementation}$ 0.4	$M_{response}$ 0.666	$P_{available}$ 0.9	P_{design} 0.9	$P_{implementation}$ 0.95	$M_{response}$ 0.994	

Figure 13. Control Effectiveness Values for SC798

Hazardous State: Nuclear radiator panels are covered by dust												
Mitigation Controls	**Implementation 1**				**Implementation 2**				**Implementation 3**			
Ability to remove dust from nuclear radiator panels	Human agent brushes dust from nuclear radiator panels				Robot agent brushes dust from nuclear radiator panels				Built-in brush automatically removes dust from nuclear radiator panels			
$P_{available}$ 0.5	P_{design} 0.9	$P_{implementation}$ 0.7	$M_{response}$ 0.783	$P_{available}$ 0.7	P_{design} 0.9	$P_{implementation}$ 0.4	$M_{response}$ 0.561	$P_{available}$ 0.9	P_{design} 0.9	$P_{implementation}$ 0.95	$M_{response}$ 0.993	

Figure 14. Control Effectiveness Values for SC800

Hazardous State: ECLSS radiator panels are covered by dust												
Mitigation Controls	**Implementation 1**				**Implementation 2**				**Implementation 3**			
Ability to remove dust from ECLSS radiator panels	Human agent brushes dust from ECLSS radiator panels				Robot agent brushes dust from ECLSS radiator panels				Built-in brush automatically removes dust from ECLSS radiator panels			
$P_{available}$ 0.5	P_{design} 0.9	$P_{implementation}$ 0.7	$M_{response}$ 0.069	$P_{available}$ 0.7	P_{design} 0.9	$P_{implementation}$ 0.4	$M_{response}$ NO MARGIN	$P_{available}$ 0.9	P_{design} 0.9	$P_{implementation}$ 0.95	$M_{response}$ 0.970	

Figure 15. Control Effectiveness Values for SC785

Hazardous State: Solar PV arrays are covered by dust Nuclear radiator panels are covered by dust Solar PV arrays are damaged Nuclear radiator panels are damaged												
Mitigation Controls	**Implementation 1**				**Implementation 2**				**Implementation 3**			
Ability to activate secondary power generation system	Smart power distribution system automatically activates secondary power generation system											
	$P_{available}$ 0.95	P_{design} 0.9	$P_{implementation}$ 0.95	$M_{response}$ 0.97								
Ability to activate battery power	Smart power distribution automatically draws power from battery storage											
	$P_{available}$ 0.95	P_{design} 0.9	$P_{implementation}$ 0.95	$M_{response}$ 0.97								

Figure 16. Control Effectiveness Values for SC11 and SC822

Hazardous State: Paint degradation on ECLSS radiator panels												
Mitigation Control(s)	**Implementation 1**				**Implementation 2**				**Implementation 3**			
Ability to repair paint damage on ECLSS radiator panels	Human agent applies additional paint layer to ECLSS radiator panels				Robot agent applies additional paint layer to ECLSS radiator panels							
	$P_{available}$ 0.4	P_{design} 0.95	$P_{implementation}$ 0.5	$M_{response}$ 0.069	$P_{available}$ 0.6	P_{design} 0.95	$P_{implementation}$ 0.3	$M_{response}$ NO MARGIN				

Figure 17. Control Effectiveness Values for SC786

Hazardous State: Habitat structural mechanical layer is breached												
Mitigation Control(s)	**Implementation 1**				**Implementation 2**				**Implementation 3**			
Ability to repair the structural mechanical layer	Human agent uses vacuum cementing to fill breach in structure				Human agent applies flexible patch to breach in structure				Robot agent applies flexible patch to breach in structure			
	$P_{available}$ 0.4	P_{design} 0.8	$P_{implementation}$ 0.5	$M_{response}$ NO MARGIN	$P_{available}$ 0.2	P_{design} 0.4	$P_{implementation}$ 0.9	$M_{response}$ NO MARGIN	$P_{available}$ 0.3	P_{design} 0.4	$P_{implementation}$ 0.7	$M_{response}$ NO MARGIN

Figure 18. Control Effectiveness Values for SC795

Hazardous State: Habitat structural protective layer is breached												
Mitigation Control(s)	**Implementation 1**				**Implementation 2**				**Implementation 3**			
Ability to repair the structural mechanical layer	Human agent replaces and compacts regolith to fill breach in structural protective layer				Robot agent replaces and compacts regolith to fill breach in structural protective layer							
	$P_{available}$ 0.6	P_{design} 0.9	$P_{implementation}$ 0.7	$M_{response}$ 1	$P_{available}$ 0.8	P_{design} 0.9	$P_{implementation}$ 0.3	$M_{response}$ 1				

Figure 19. Control Effectiveness Values for SC796

Hazardous State: Habitat structural mechanical layer is breached Fire in interior environment												
Mitigation Control(s)	**Implementation 1**								**Implementation 2**			**Implementation 3**
Ability to regulate the temperature of the interior environment	Temperature control system automatically heats/cools the interior environment											
	$P_{available}$ 0.95	P_{design} 0.95	$P_{implementation}$ 0.95	$M_{response}$ NO MARGIN								
Ability to regulate the pressure of the interior environment	Pressure control system automatically increases/decreases pressure of interior environment											
	$P_{available}$ 0.95	P_{design} 0.95	$P_{implementation}$ 0.95	$M_{response}$ NO MARGIN								

Figure 20. Control Effectiveness Values for SC823 and SC824

Hazardous State: Energy storage system is damaged									
Mitigation Control(s)	**Implementation 1**				**Implementation 2**				
Ability to replace energy storing units	Human agent replaces energy storing units				Robot agent replaces energy storing units				
	$P_{available}$ 0.4	P_{design} 0.9	$P_{implementation}$ 0.9	$M_{response}$ NO MARGIN	$P_{available}$ 0.5	P_{design} 0.9	$P_{implementation}$ 0.2	$M_{response}$ NO MARGIN	

Figure 21. Control Effectiveness Values for SC355

Hazardous State: Smart power distribution main bus is damaged									
Mitigation Control(s)	**Implementation 1**				**Implementation 2**				
Ability to replace energy storing units	Human agent replaces battery cells				Robot agent replaces battery cells				
	$P_{available}$ 0.4	P_{design} 0.9	$P_{implementation}$ 0.9	$M_{response}$ NO MARGIN	$P_{available}$ 0.5	P_{design} 0.9	$P_{implementation}$ 0.2	$M_{response}$ NO MARGIN	

Figure 22. Control Effectiveness Values for SC803

Hazardous State: Smart power distribution converters are damaged									
Mitigation Control(s)	**Implementation 1**				**Implementation 2**				
Ability to replace energy storing units	Human agent replaces battery cells				Robot agent replaces battery cells				
	$P_{available}$ 0.4	P_{design} 0.9	$P_{implementation}$ 0.9	$M_{response}$ NO MARGIN	$P_{available}$ 0.5	P_{design} 0.9	$P_{implementation}$ 0.2	$M_{response}$ NO MARGIN	

Figure 23. Control Effectiveness Values for SC802

Hazardous State: Solar PV arrays are damaged												
Mitigation Control(s)	**Implementation 1**				**Implementation 2**				**Implementation 3**			
Ability to replace solar PV arrays	Human agent replaces damaged solar PV arrays				Robot agent replaces damaged solar PV arrays							
	$P_{available}$ 0.3	P_{design} 0.9	$P_{implementation}$ 0.7	$M_{response}$ NO MARGIN	$P_{available}$ 0.4	P_{design} 0.9	$P_{implementation}$ 0.2	$M_{response}$ NO MARGIN				

Figure 24. Control Effectiveness Values for SC799

Hazardous State: Nuclear radiator panels are damaged												
Mitigation Control(s)	**Implementation 1**				**Implementation 2**				**Implementation 3**			
Ability to replace nuclear radiator panels	Human agent replaces damaged nuclear radiator panels				Robot agent replaces damaged nuclear radiator panels							
	$P_{available}$ 0.3	P_{design} 0.9	$P_{implementation}$ 0.6	$M_{response}$ NO MARGIN	$P_{available}$ 0.4	P_{design} 0.9	$P_{implementation}$ 0.1	$M_{response}$ NO MARGIN				

Figure 25. Control Effectiveness Values for SC801

Hazardous State: Fire in interior environment												
Mitigation Control(s)	**Implementation 1**				**Implementation 2**				**Implementation 3**			
Ability to extinguish active fire in interior environment	Human agent uses fire extinguisher to put out fire				Human agent uses fire blanket to put out fire				Robot agent uses fire blanket to put out fire			
	$P_{available}$ 0.3	P_{design} 0.8	$P_{implementation}$ 0.9	$M_{response}$ 0.0354	$P_{available}$ 0.4	P_{design} 0.6	$P_{implementation}$ 0.8	$M_{response}$ 0.0354	$P_{available}$ 0.5	P_{design} 0.6	$P_{implementation}$ 0.5	$M_{response}$ 0.149

Figure 26. Control Effectiveness Values for SC797

Hazardous State: ECLSS fan has buildup of dust in filter												
Mitigation Control(s)	**Implementation 1**				**Implementation 2**				**Implementation 3**			
Ability to remove dust from fan in ECLSS pressure system	Human agent removes dust from fan filter in ECLSS pressure system				Robot agent removes dust from fan filter in ECLSS pressure system							
	$P_{available}$ 0.5	P_{design} 0.8	$P_{implementation}$ 0.9	$M_{response}$ NO MARGIN	$P_{available}$ 0.7	P_{design} 0.8	$P_{implementation}$ 0.6	$M_{response}$ NO MARGIN				

Figure 27. Control Effectiveness Values for SC787

Hazardous State: ECLSS air tank has a leak												
Mitigation Control(s)	**Implementation 1**				**Implementation 2**				**Implementation 3**			
Ability to repair air tank in ECLSS pressure system	Human agent applies flexible patch over leak in air tank of ECLSS pressure system				Human agent applies hardening putty over leak in air tank of ECLSS pressure system				Human agent replaces and re solders a new pipe			
	P_{availa} 0.4	[No Title] 0.2	$P_{implementation}$ 0.9	$M_{response}$ NO MARGIN	$P_{available}$ 0.4	P_{design} 0.6	$P_{implementation}$ 0.8	$M_{response}$ NO MARGIN	$P_{available}$ 0.2	P_{design} 0.9	$P_{implementation}$ 0.3	$M_{response}$ NO MARGIN

Figure 28. Control Effectiveness Values for SC788

Hazardous State: ECLSS air supply valve is malfunctioning												
Mitigation Control(s)	**Implementation 1**				**Implementation 2**				**Implementation 3**			
Ability to repair the air supply valve in ECLSS pressure system	Human agent replaces the air supply valve in ECLSS pressure system				Robot agent replaces the air supply valve in ECLSS pressure system							
	$P_{available}$ 0.3	P_{design} 0.9	$P_{implementation}$ 0.8	$M_{response}$ NO MARGIN	$P_{available}$ 0.4	P_{design} 0.9	$P_{implementation}$ 0.3	$M_{response}$ NO MARGIN				

Figure 29. Control Effectiveness Values for SC789

Hazardous State: ECLSS compressor performance is decreased						
Mitigation Control(s)	**Implementation 1**				**Implementation 2**	**Implementation 3**
Ability to repair the compressor in ECLSS thermal system	Human agent replaces the compressor in ECLSS thermal system					
	$P_{available}$ 0.2	P_{design} 0.9	$P_{implementation}$ 0.3	$M_{response}$ NO MARGIN		

Figure 30. Control Effectiveness Values for SC791

Hazardous State: ECLSS evaporator has air side leak						
Mitigation Control(s)	**Implementation 1**				**Implementation 2**	**Implementation 3**
Ability to repair the evaporator in ECLSS thermal system	Human agent replaces the evaporator in ECLSS thermal system					
	$P_{available}$ 0.2	P_{design} 0.9	$P_{implementation}$ 0.3	$M_{response}$ NO MARGIN		

Figure 31. Control Effectiveness Values for SC792

Hazardous State: ECLSS heater performance is decreased						
Mitigation Control(s)	**Implementation 1**				**Implementation 2**	**Implementation 3**
Ability to repair the heater in ECLSS thermal system	Human agent replaces the heater in ECLSS thermal system					
	$P_{available}$ 0.2	P_{design} 0.9	$P_{implementation}$ 0.6	$M_{response}$ NO MARGIN		

Figure 32. Control Effectiveness Values for SC793

4.1.2 Identify Relevant Performance Metrics

Before running our chosen disruption and safety control scenarios in the MCVT, we must identify the relevant habitat/system performance metrics that will enable us to characterize how components, subsystems, or systems are performing relative to their desired objectives. To do so requires knowledge of where in the MCVT habitat architecture the disruptions, hazardous states, and safety controls exist (i.e., in what subsystems are they modeled), and how exactly they are modeled. Specifically, we need to understand the overall control process that enables each safety control to response to disruptions and hazardous states and keep the habitat in a safe operating state. Developing these control processes will allow for the identification of modeling requirements for individual MCVT subsystems.

The completion of this step in the control effectiveness validation procedure requires three tasks:

1. Develop an MCVT habitat diagram to represent the distribution of disruption events and safety controls embedding in the system
2. Use a Design Structure Matrix to understand the flow of information, energy, and matter in the simulation
3. Develop control process models to inform the requirements needed to model individual safety controls in the MCVT

To develop requirements for the safety controls discussed in Section **4.1.1. Assess Control Effectiveness for a Set of Safety Controls**, we must first understand where in the MCVT the disruption and safety controls apply. This will help allocate requirements to the appropriate model. To do this, we can develop an MCVT diagram with an abstract representation of the core systems, subsystems/component assemblies, and installed sensors in the simulation environment (Figure 33).

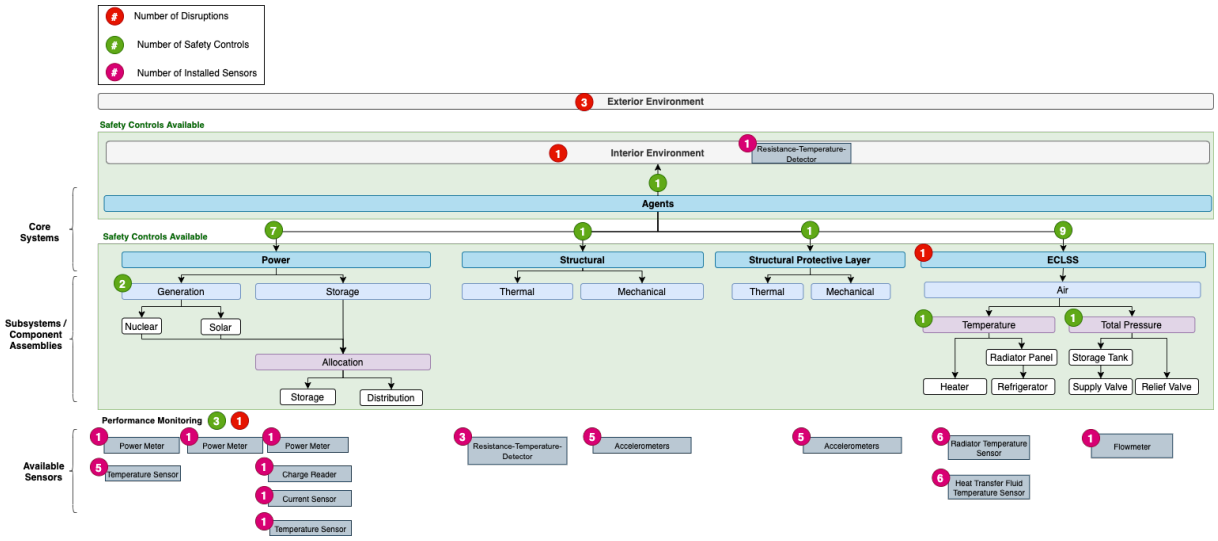

Figure 33. MCVT Architecture with Distribution of Disruptions, Safety Controls, and Installed Sensors

The red, green, and magenta circles in Figure 33 indicate the number of disruptions, safety controls, and installed sensors corresponding to each MCVT system, and/or subsystem/component assembly. Note that the sensors listed at the bottom of the diagram are vertically aligned with their relevant subsystem/component assembly. The disruption events in the simulation largely originate

from the exterior environment because they are environmental events which impact the system performance. Three exceptions are (1) Fire, which originates within the habitat interior environment, (2) ECLSS mechanical fan failure, caused by failure modes in other subsystems, and (3) Sensor network failures, which do not originate from the exterior environment, but within the sensor network itself. The safety controls available to response to disruptions are distributed throughout the core systems and subsystem/component assemblies. Note that most safety controls require agent intervention, such as cleaning dust from solar or nuclear power generation panels, or repairing a breach in the structure. Other safety controls within the subsystems represent automated functionalities that act to keep the system in a safe operating state. For example, these include the ability to regulate the interior environment temperature and pressure.

The next step in identifying performance metrics for each safety control is using a Design Structure Matrix to understand the cyber and physical interdependencies between subsystems in the MCVT. This will tell us what information is passed between subsystems, and what information is important to monitor for each safety control. The cyber and physical interdependencies between all subsystems has been compiled into an interactive Design Structure Matrix by RETHi PhD student at Purdue, Herta Montoya. This Design Structure Matrix is contained in the RETHi project documentation and will be used for the identification of safety control performance metrics in the MCVT.

The last step is to identify functional requirements for the control processes in our disruption scenarios. To do so, we have adapted the control process model developed by Leveson (2004) to represent the implementation of a single safety control in the MCVT. The reference control process model adapted from Leveson (2004) is shown in Figure 34.

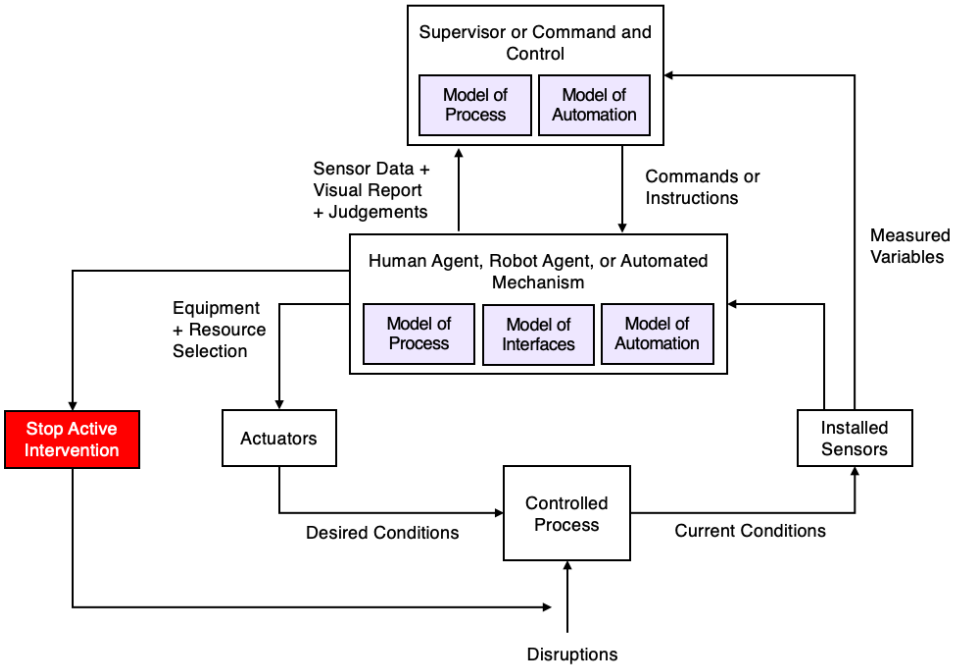

Figure 34. Control Process Model Adapted from Leveson (2004)

This diagram enables us to frame our safety controls in such a way that we can more easily identify information that is relevant in modeling safety controls in the MCVT and developing functional requirements. Specifically, we can create narratives for each controlled process by filling in the colored italics of the following template adapted from Leveson (2004):

The *controlled process* is executed by an *acting body* who…

(1) Obtains information (observes) the process state from *measured variables*, and

(2) Uses this information to *initiate action* by manipulating controlled variables to keep the process operating within predefined limits (*constraints*) or set points despite *disruptions* to the process

As an example of a safety control narrative, consider the safety control, *Ability to remove dust from solar PV arrays*, with the specific implementation of a human agent using a brush to remove dust. The control process narrative is as follows:

Managing the amount of dust on the solar PV array is executed by a *human agent* who observes degradation in system performance from measured *solar power output* and uses this information to *obtain a brush and remove dust from the solar PV arrays* to keep the solar output power above *90% of the nominal power output* despite *dust continually accumulating on the surface.*

The colored portions of the process narrative can then inform the modeling of this safety control in the MCVT. For this control, we can identify the following general modeling requirements for the control process:

1. Ability to model dust accumulation on the solar PV arrays
2. Ability to model degradation in solar power output due to dust accumulation on solar PV arrays
3. Ability to detect degradation in solar power output
4. Ability to model dust removal from the solar PV arrays
5. Ability to model the recovery of performance degradation when dust is removed from solar PV arrays

In addition, for the assessment of control effectiveness, MCVT users must be able to control certain aspects of the control process. Specifically, the following user requirements for the control process can be identified:

1. Ability for a user to set the dust accumulation rate on the solar PV arrays
2. Ability for a user to set the activation time of dust removal activities after hazard detection
3. Ability for a user to set the dust cleaning rate during active intervention

In addition to building the control process narrative and identifying functional requirements, we can adapt the control process model in Figure 34 to visually depict the control process and feedback loops of measured information. Figure 35 shows the control process for the same safety control, *Ability to remove dust from solar PV arrays*, with the specific implementation of a human agent using a brush to remove dust.

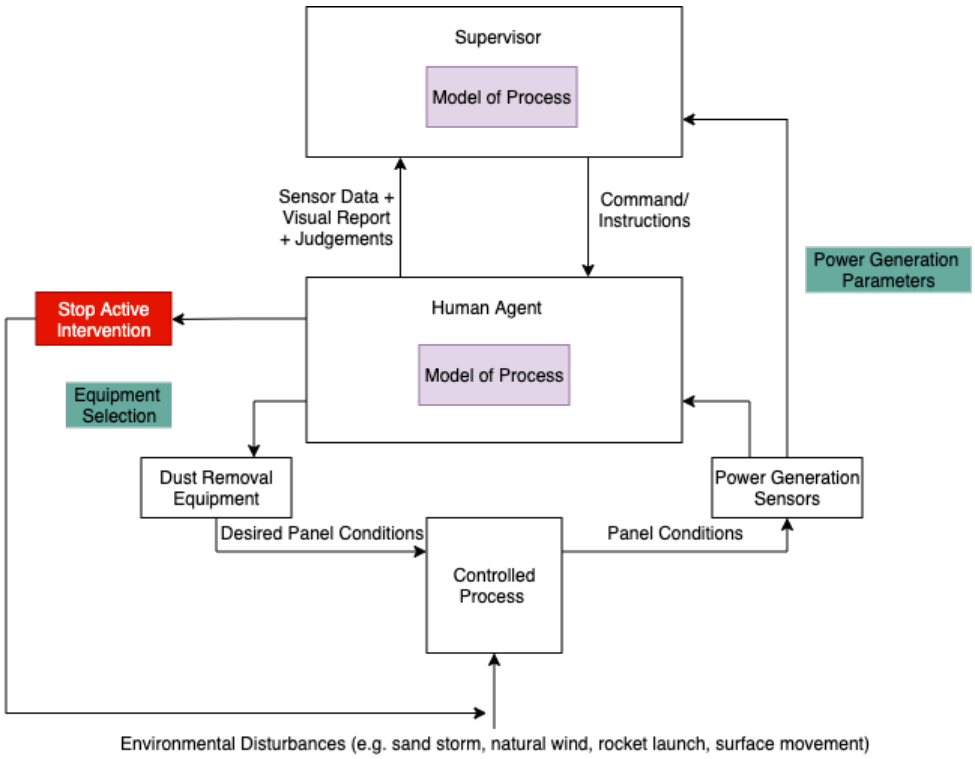

Figure 35. Control Process Model for Safety Control: *Ability to remove dust from solar PV arrays*

Starting at the bottom of Figure 35, we have our controlled process block, which in this case is to manage the amount of dust on the solar panels. The desired panel conditions are an input to the process, while the actual panel conditions during operation are the output of the process. The control process is managed by both a supervisor and active agent. Both of these roles may be done by a human agent during operation. In this diagram, the roles are isolated in two blocks to show they are unique roles in the control process. In this case, the supervisor monitors the measured variables, which are the power generation parameters (i.e., solar power output). Monitoring this variable over time enables the detection of deterioration in power generation. When power generation drops by 10% of the nominal value, the supervisor instructs the agent to use a brush to clean the solar PV array and execute the control process. The human agent is then responsible for monitoring the conditions of solar panels until all dust is removed, while the supervisor continually monitors power generation to ensure it returns to the nominal value. Upon completion of the

control process, the human agent would stop active intervention until conditions necessitate re-activation of the loop (i.e., more dust accumulates).

The procedure of developing control process diagrams and narratives was completed for the safety controls in MCVT v6. Doing so allowed for the identification of the functional requirements, performance metrics (measured variables), and operational limits (constraints) that inform which user-controlled variables, installed sensors, and FDD thresholds must be modeled in the MCVT for control effectiveness validation. These performance metrics will be used in Steps 3 and 4 of this control effectiveness validation plans, as each metric can be plotted as a function of time to generate resilience curves and conduct a resilience assessment. Table 10 highlights the performance metric and FDD threshold information for each safety control modeled in the MCVT.

Table 10. Control Process Information for Safety Controls in MCVT v6

Safety Control	**Controlled Process Description**	**Relevant Habitat Performance Metric(s)**	**Fault Metric / FDD Threshold**
Ability to remove dust from solar PV arrays	Control the amount of dust on solar PV arrays	1. Solar power output 2. Solar dust ratio 2. Solar irradiance	Binary indicator of status; fix if failed 0 – Healthy 1 – Failed
Ability to remove dust from nuclear radiator panels	Control the amount of dust on nuclear radiator panels	1. Nuclear power output 2. Nuclear dust ratio	Binary indicator of status; fix if broken 0 – Healthy 1 – Broken
Ability to activate secondary power generation system	Control the amount of power being generated	1. Power supplied to all subsystems 2. Solar power output 3. Nuclear power output	N/A – automated functionality based on total power being generated
Ability to activate battery power	Control the amount of power being distributed	1. Currently stored energy 2. Maximum available energy storage 3. Power supplied to all subsystems	N/A – automated functionality based on total power being generated
Ability to remove dust from ECLSS radiator panels	Control the amount of dust on ECLSS radiator panels	1. ECLSS radiator secondary loop fluid temperatures	% Of ECLSS radiator panel covered by dust Fix if greater than 20%
Ability to repair paint damage on ECLSS radiator panels	Control the amount of paint on surface of ECLSS radiator panels	1. ECLSS radiator secondary loop fluid temperatures	% Of ECLSS radiator panel with chipped paint Fix if greater than 20%

Table 10. Continued

Ability to repair breach in structural mechanical layer	Control the size of a breach in structural mechanical layer	1. Interior environment temperature 2. Interior environment pressure 3. Hole radius in structure 4. Total volume removed	Monitor total volume removed. Any value greater than zero requires fixing
Ability to repair breach in structural protective layer	Control the size of a breach in the structural protective layer	1. Interior environment temperature 2. Interior environment pressure 3. Volume of hole in regolith (per panel) 4. Number of SPL elements damaged 5. Damage level (per panel)	Scale to classify existence and size of breach; fix if 2, 3, 4, or 5 1 – Healthy 2 – ½ depth of volume removed for one panel 3 – Full depth of volume removed for one panel. Structure has been directly impacted with no perforation 4 – Full depth of volume removed for 2 panels. Small hole in structure 5 – Full depth of volume removed for 3 panels. Large hole in structure
Ability to regulate the temperature of the interior environment	Control the temperature of the interior environment	1. Interior environment temperature	Monitor temperature of interior environment. Values above or below the set point requires activation of heating or cooling
Ability to regulate the pressure of the interior environment	Control the pressure of the interior environment	1. Interior environment pressure	Monitor pressure of interior environment. Values above or below the set point requires activation of the air supply/relief valve
Ability to replace individual power converters	Control the power output from individual converts to corresponding subsystem	1. Power output of each converter (1 – 6)	Binary indicator of convertor status; fix if broken 0 – Healthy 1 – Broken
Ability to replace main power generation bus	Control the amount of available power for distribution to subsystems	1. Power output of the main generation bus	Binary indicator of generation bus status; fix if broken 0 – Healthy 1 – Broken

Table 10. Continued

Ability to replace solar PV arrays	Control the number of functioning solar PV arrays	1. Solar power output 2. Solar irradiance 3. Percentage of solar panels broken	Percent decrease in solar power output. Decrease of 10% or greater requires fixing
Ability to replace nuclear radiator panels	Control the number of functioning nuclear radiator panels	1. Nuclear power output 2. Percentage of nuclear radiator panels broken	Percent decrease in nuclear power output. Decrease of 10% or greater requires fixing
Ability to replace energy storing units	Control the number of functioning battery cells	1. Number of battery cells available 2. Currently stored energy 2. Maximum available energy storage	Percentage of healthy battery cells remaining. Fix if below 90%
Ability to extinguish active fire in interior environment	Control the size of a fire in the interior environment	1. Interior environment temperature 2. Interior environment pressure	Binary indicator of system health; fix if failed 0 – Healthy 5 – Failure
Ability to repair fan damage in ECLSS pressure system	Control the amount of dust on the ECLSS fan filter	1. Flow rate of the fan	Binary indicator of fan status; fix if broken 0 – Healthy 1 – Broken
Ability to repair piping between the air tank in ECLSS pressure system and interior environment	Control the size of a leak in the ECLSS air tank	1. Air tank pressure 2. Upstream flow rate 3. Downstream flow rate	Scale to classify existence and size of leak; fix if 3 or 4 0 – Healthy 1 – Inconsequential Damage 2 – Minor Damage 3 – Moderate Damage 4 – Major Damage
Ability to repair the air supply valve in ECLSS pressure system	Control the functionality of the air supply valve in ECLSS pressure system	1. Downstream flow rate	Binary indicator of valve health; fix if broken 0 – Healthy 1 – Broken
Ability to repair the compressor in ECLSS thermal system	Control the functionality of the compressor in ECLSS thermal system	1. Requested RPM for compressor 2. Actual RPM of compressor	Scale to classify existence and severity of compressor damage; fix if 3 or 4 0 – Healthy 1 – Inconsequential Damage 2 – Minor Damage 3 – Moderate Damage 4 – Major Damage

Table 10. Continued

Ability to repair the evaporator in ECLSS thermal system	Control the functionality of the evaporator in ECLSS thermal system	1. Upstream flow rate 2. Downstream flow rate	Scale to classify existence and severity of evaporator damage; fix if 3 or 4 0 – Healthy 1 – Inconsequential Damage 2 – Minor Damage 3 – Moderate Damage 4 – Major Damage
Ability to repair the heater in ECLSS thermal system	Control the functionality of the heater in ECLSS thermal system	1. Requested heat 2. Actual heat provided	Scale to classify existence and severity of heater damage 0 – Healthy 1 – Inconsequential Damage 2 – Minor Damage 3 – Moderate Damage 4 – Major Damage

4.1.3 Obtain Performance Metrics as a Function of Time

The Modular Coupled Virtual Testbed (MCVT)

The MCVT is the simulation platform used to evaluate our control effectiveness definition. The MCVT has gone through several development phases during the completion of this thesis, and the final MCVT v6 was used to execute this validation procedure. The MCVT v6 includes physics-based models with damageable/repairable subsystem properties, including a 3-dimensional world and the associated models, power systems, robotic agents, the pressure and thermal control aspects of the environmental control life support system (ECLSS), and the fault detection and health management (Dyke et al., 2022).

The MCVT provides the capability to explore techniques and algorithms needed to extract the necessary amount of actionable information for repair and recovery through monitoring and embedded intelligence (Dyke et al., 2022). In the case of our safety control evaluation, the MCVT provides us with the ability to implement appropriate defenses that can respond to hazards, system deterioration, and common faults in system components. These defenses are the available

prevention, intervention, and mitigation controls relevant in addressing the disruptions and failures outlined in Table 8. In addition, the simulation environment is designed to support the exploration of a wide range of safety controls with varying control effectiveness values to evaluate our definition of control effectiveness. To achieve these goals, the MCVT was framed in a system-of-systems context, where each component is a constituent system which can operate independently, while their combination establishes the extraterrestrial habitat as an integrated complex system (Dyke et al., 2022). Doing so delivers both anticipated (and potentially unanticipated) emergent behaviors, and allows for the emergence and propagation of performance effects due not only to the disruptions and failures, but also our implemented safety controls.

The subsystems included in the MCVT architecture can be broadly classified into three groups (Dyke et al., 2022):

1. **Electro-Mechanical Systems (EMS):** those subsystems which directly propagate the physics of the habitat in both its operational condition as well as in its various hazardous states (including a protective regolith layer, a structural system, an environmental control and life support system (ECLSS), and a power generation and distribution system)
2. **Health Management Systems (HMS):** those subsystems which primarily provide system evaluation and decision making, but also include a physical aspect (including an internal and interplanetary communication network, a command & control system, and a data repository)
3. **Agent Systems:** a single robotic agent that acts as the interface between the EMS and HMS subsystems, playing a significant role in each as it affects the physical changes in the habitat according to the direction of the HMS

The layout of these MCVT subsystems is important in evaluating the response margin of our control effectiveness metric (as discussed at the end of in Chapter 3). The layout provides us with distance estimates between the different subsystems, which will inform the activation time of safety controls involving an agent. Figure 36 shows the physical distribution of the MCVT architecture.

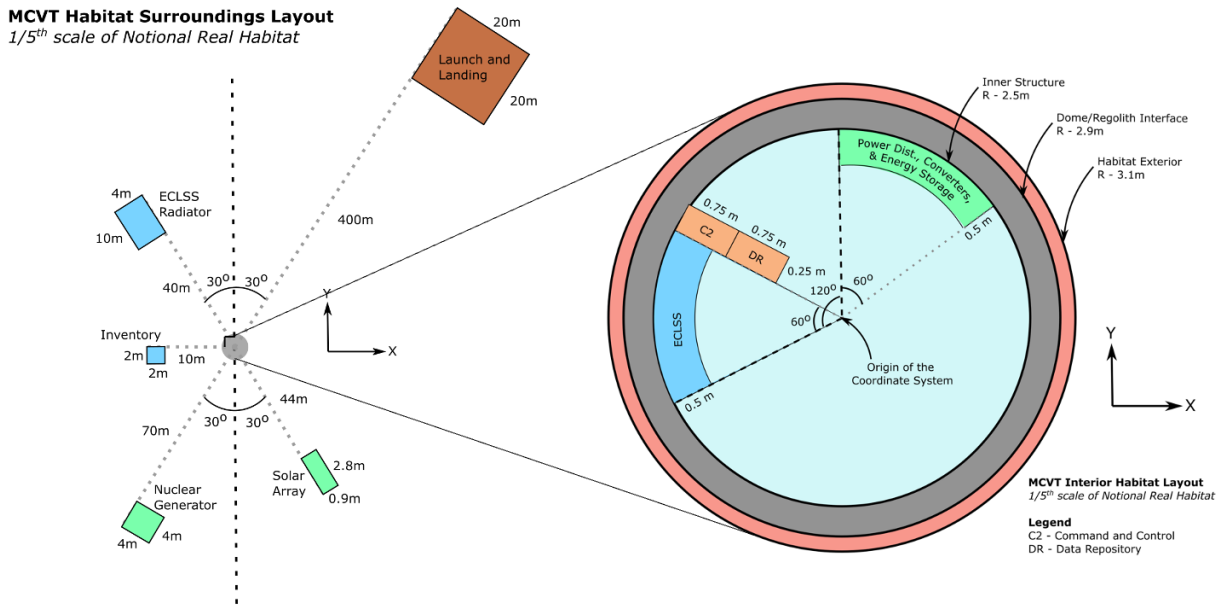

Figure 36. MCVT Layout of Habitat Subsystems (Dyke et al., 2022)

Modeling Variations in Control Effectiveness Values

To use the MCVT to evaluate the definition of control effectiveness, we must understand how each value in the control effectiveness metric informs the modeling of safety controls. More specifically, we must determine how to set up a simulation to run with different sets of controls with different control effectiveness values (i.e., with different implementation strategies). Based on the identification of functional requirements in Step 2 of the validation procedure, the user-controlled capabilities that allow for the simulation of different implementation strategies are:

1. Ability for a user to control the activation or deactivation of all safety controls ($P_{available} = 1\ or\ 0$)
2. Ability for a user to set the activation time of all safety controls (set by $t_{sc,affect}$)
3. Ability for a user to set the intervention rate and/or time of all safety controls (informed by values of P_{design} and $P_{implementation}$

MCVT v6 currently supports only requirements 1 and 3. Future versions of the MCVT should support requirement 2 and enable the evaluation of the control effectiveness response margin. This will be discussed further in the **Chapter 6**. Here we describe current capabilities of the MCVT and

how users can set up each simulation to depict safety control architectures to reflect the control effectiveness values of chosen safety controls and implementation strategies.

To model variations in $P_{available}$, the MCVT provides users with the ability to control the activation of each safety control. Each safety control has a user-controlled threshold that determines whether an agent may be scheduled to complete each safety control. The thresholds are informed by the fault metric/FDD thresholds listed in Table 10. For example, if the control process has four levels of severity and levels 3 and 4 will activate the control process, then a user can set a value of 5 to "deactivate" the control because the FDD threshold will never reach 5. Conversely, setting a value of 2.5 will allow levels 3 and 4 to activate intervention. In the binary cases of Table 10, a value of 2 would "deactivate" the control, while a value of 1 would activate it. Currently, the MCVT is only capable of exploring $P_{available} = 0\ or\ 1$. This limitation is due to computational inefficiencies, as the MCVT only supports short-term modeling on the order of minutes, and doing a probabilistic activation analysis would require modeling for long periods of time. This task is therefore a part of future work, which is discussed in **Chapter 6**.

To model variations in P_{design} and $P_{implementation}$, the MCVT provides users with the ability to control the repair rate of each repairable feature in the simulation (i.e., users can control how quickly the safety control goal is achieved). As previously mentioned, the MCVT only supports short-term modeling on the order of minutes. Therefore, the repair rates that must be set by a user are higher than is physically realistic, as they have been scaled for each safety control to enable a full performance recovery to nominal values at least once in a 120 second simulation time. To reflect the different implementation strategies of each safety control, the repair rate set in each simulation is multiplied by both P_{design} and $P_{implementation}$ for each strategy. By multiplying these probabilities, we can obtain a single value between 0 and 1 (X) which will be used to set the repair rate in the simulation as X% of the repair rate that enables a full performance recovery to nominal values at least once in a 120 second simulation time.

The MCVT cannot currently support the evaluation of $M_{response}$; specifically in the context of controlling the activation time of our safety controls. However, the MCVT still helps in the estimation of response margin for each safety control. The propagation of a disruptive event from

its point of origin in the habitat will provide an estimate of $t_{h,effect}$, as the simulation will allow us to determine the time it takes to detect a measurable degradation in performance in the habitat. This variable will then be consistent for all safety controls that address the same disruptive event, as it is indicative of how we defined the hazardous state and not the activated safety controls. Conversely, the value of $t_{sc,affect}$ is a property of each safety control and how quickly intervention can begin. This value can be determined for each safety control based on the assumptions discussed in **Chapter 3** and the MCVT layout shown in Figure 36.

4.1.4 Use Performance Curves to Assess Resilience

As previously discussed, resilience can be defined in many ways. For RETHi, resilience is defined as the ability of a system, process, or organization to react to, survive, and recovery from disruptions. To understand the resilience of a hypothetical habitat with varying safety control architectures, we can conduct a resilience assessment using performance data from the MCVT. Doing so will allow us to determine whether our control effectiveness metric helps us select appropriate safety controls for implementation that result in resilient space habitat architectures. The resilience assessment includes both qualitative and quantitative analyses. We can first qualitatively analyze the performance metric curves relevant for our safety controls to determine whether they exhibit the characteristic "resilient" shape depicted in the nominal performance curve shown in Figure 37.

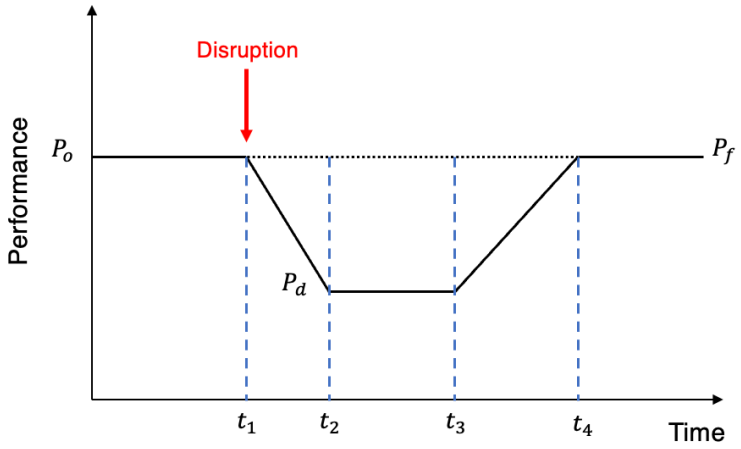

Figure 37. Nominal Resilience Curve

The high control effectiveness safety controls should result in performance curves that minimize both the system degradation from a disruption (survivability) and the time between the system's performance level at the time of disruption to its regained performance level (recoverability). Lower control effectiveness controls should *not* result in curves that have better survivability and recoverability compared to the high control effectiveness results. Such behavior would indicate an issue in how control effectiveness is defined, and re-evaluation would be necessary.

To understand whether one curve is "more resilient" than another, we will use a quantitative resilience assessment to map performance metric data for each safety control to existing resilience metrics from literature. Many different metrics have been defined in literature to quantify the resilience of a system, and therefore we are not redefining a new metric here. By conducting a literature review of resilience publications, we can identify wide-spread themes in existing metrics to apply in the control effectiveness validation procedure. In the next sections, we discuss the metrics identified in the literature review and the process of selecting metrics to use in the quantitative resilience assessment.

Resilience Metric Literature Review

For all resilience metric equations discussed in this section, refer back to Figure 37 for clarification on the variables used to formulate each metric.

Metric 1 (Re_1): Das et al. (2020)

The first metric considered is defined as the inverse of the time that a system is in a state of disruption. In this investigation, the time that the system is in a disrupted state is from t_1 to t_4:

$$Re_1 = \frac{1}{t_4 - t_1} \tag{4}$$

Metric 2 (Re_2): Henry & Ramirez-Marquez (2012)

The second metric is defined as the ratio of the recovered performance ($P(t_4)$) to the disrupted performance ($P(t_1)$

$$Re_2 = \frac{P(t_4)}{P(t_1)} \tag{5}$$

Metric 3 (Re_3): Bruneau et al. (2003)

The next metric uses integration to calculate the total performance lost by the system between the time of the disruption (t_1) and recovery (t_4):

$$Re_3 = \int_{t_1}^{t_4} (P(t_1) - P(t))dt \tag{6}$$

Metric 4 (Re_4): Ayyub et al. (2014)

The next metric is defined as:

$$Re_4 = \frac{t_1 + F(t_2 - t_1) + R(t_4 - t_3)}{t_4 - t_1}$$
$$F = \frac{\int_{t_1}^{t_2} P(t)dt}{\int_{t_1}^{t_2} P(t_1)dt} \qquad R = \frac{\int_{t_3}^{t_4} P(t)dt}{\int_{t_3}^{t_4} P(t_1)dt} \tag{7}$$

F and *R* in this equation are ratios of the actual performance of the system to the non-disrupted performance during the failure and recovery stages, respectively. This metric is larger when the there is less performance loss and when there is less time spent in a state of decreased performance. In Figure 39, *F* and *R* are the ratios of the striped sections to the orange sections.

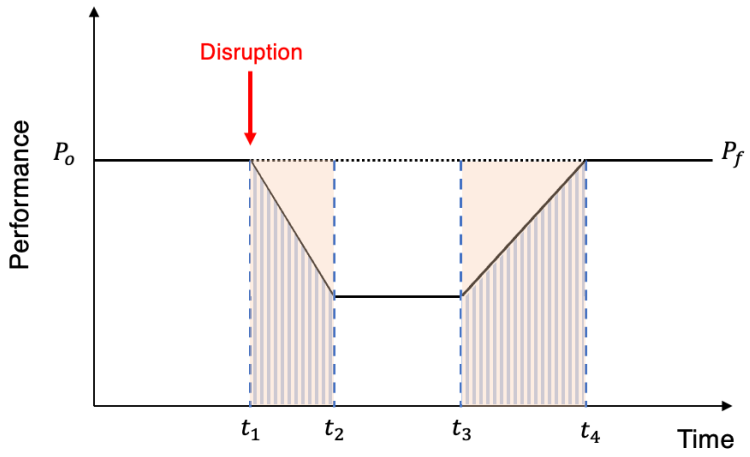

Figure 38. Visualization of Failure and Recovery Stages of the Baseline Resilience Curve

Metric 5 (Re_5): Henry & Ramirez-Marquez (2012)

The next metric improves on the authors' previously published metric (Re_2). This metric is the ratio of the increase in performance during recovery to the loss in performance following the disruption:

$$Re_5 = \frac{P(t_4) - min(P)}{P(t_1) - min(P)} \tag{8}$$

Metric 6 (Re_6): Yarveisy et al. (2020)

Metric 6 is comprised of a combination of three capacities that describe three portions of the baseline performance curve:

$$Re_6 = Ab + (Ad \cdot Res) - (Ab \cdot Ad \cdot Res) \tag{9}$$

Ab denotes the absorptive capacity of the system, or its ability to limit performance loss after a disruption. The coefficient C_{Ab} accounts for natural degradation of the system. It is assumed for this study that there is no natural degradation expected, as the run time of the MCVT is on the order of minutes. So, in computing this metric, $C_{Ab} = 1$, leaving Ab defined as the ratio of the minimum performance to the starting performance.

$$Ab = C_{Ab} \cdot \left(\frac{min(P)}{P(t_1)}\right) \tag{10}$$

Ad is a measure of the adaptive capacity of the system, or its ability to stabilize performance after a disruption.

$$Ad = 1 - \frac{t_3 - t_2}{t_4 - t_1} \tag{11}$$

Res is the restorative capacity of the system. This is the ability of the system to return to its original performance level. C_R in the equation below is a coefficient that accounts for natural degradation of the system, and is again assumed to be 1. C_T is the ratio of time spent not recovering to total time spent at a disrupted performance level.

$$Res = \frac{1}{90} \cdot tan^{-1}\left[\frac{P(t_4) - P(t_3)}{\frac{t_4 - t_3}{t_4 - t_1}}\right] \cdot C_T \cdot C_R \tag{12}$$

$$C_T = \frac{t_3 - t_1}{t_4 - t_1}$$

Metric 7 (Re_7): Cheng et al. (2020)

Like the metric defined by Yarveisy et al. (2020), this metric is comprised of capacities. In this case, Cheng et al. makes use of only the absorptive (*Ab*) and restorative capacity (*Res*) of the system.

Absorptive capacity refers to the ability of the system to absorb shocks, or limit performance loss due to a disruption. Restorative capacity is the ability of the system to recover from a loss. Each capacity is weighted by a coefficient, α or β. The sum of the two coefficients must be 1 but their values can be varied to emphasize the importance of the system's absorptive or restorative capacity over one another. For this study, we assume that the two capacities are equally important, so $\alpha = \beta = 0.5$.

$$Re_7 = \alpha(Ab) + \beta(Res) \tag{13}$$

Ab is the product of three values that describe different aspects of the performance curve. δ_d is the ratio of the actual performance of the system to the ideal performance in the absence of a disruption. σ_d is the ratio of the minimum performance to the original performance. ρ_d accounts for natural degradation and is again assumed to be 1, meaning there is no expected natural degradation.

$$Ab = \delta_d \sigma_d \rho_d$$
$$\delta_d = \frac{\int_{t_1}^{t_{min}} P(t)dt}{(t_{min} - t_1)P(t_1)} \qquad \sigma_d = \frac{min(P)}{P(t_1)} \tag{14}$$

Res is the product of the same three values as *Ab*, but characterizes the restorative stage of the curve, rather than the disrupted state. δ_r is the ratio of the actual performance of the system to the ideal performance in the absence of a disruption. σ_r is the ratio of the minimum performance to the original performance. ρ_r again, accounts for natural degradation and is assumed to be 1, meaning there is no expected natural degradation.

$$Res = \delta_r \sigma_r \rho_r$$
$$\delta_r = \frac{\int_{t_{min}}^{t_4} P(t)dt}{(t_2 - t_{min})P(t_1)} \qquad \sigma_d = \frac{P(t_4)}{P(t_1)} \tag{15}$$

Selecting Metrics for Quantitative Resilience Assessment

Not all metrics identified in the literature review return meaningful values (i.e., metric goes to infinity or does not exist (DNE)) for every performance curve shape. It is important to identify those metrics for which we can compute finite values for performance data that exhibits either the "standard" resilience shape shown in Figure 37, or some variation in the curve. For example, Figures 39–43 present ten different resilience curves that might be obtainable through different habitat architectures and selections of safety controls in the MCVT. These shapes were identified through the help of Purdue Masters student, Jacqueline Ulmer.

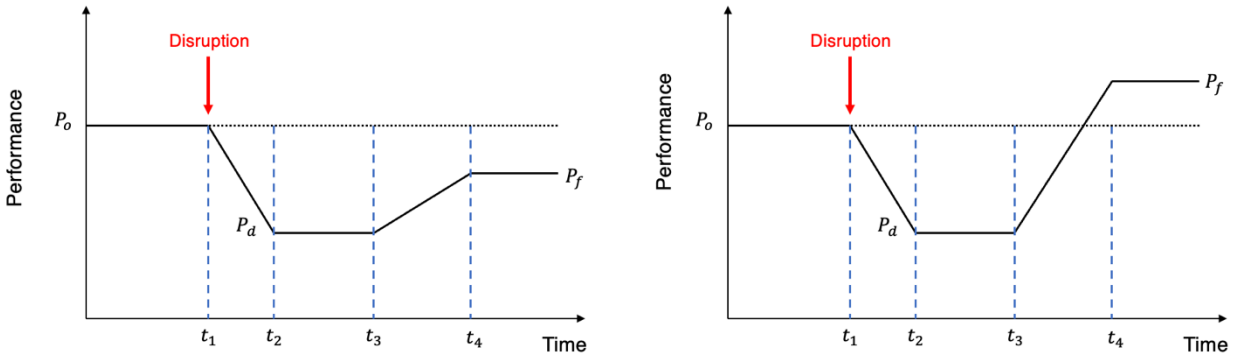

Figure 39. (Left) Performance Lost, (Right) Performance Gained

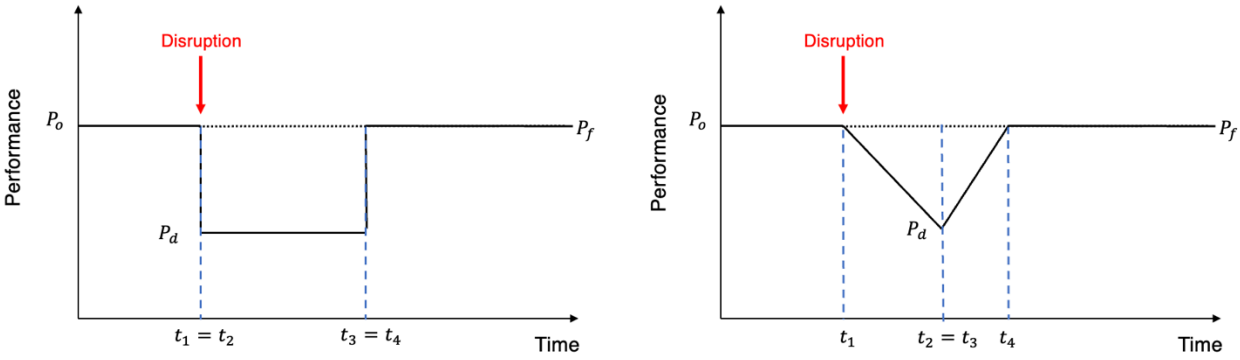

Figure 40. (Left) Bucket Shape, (Right) "V" Shape

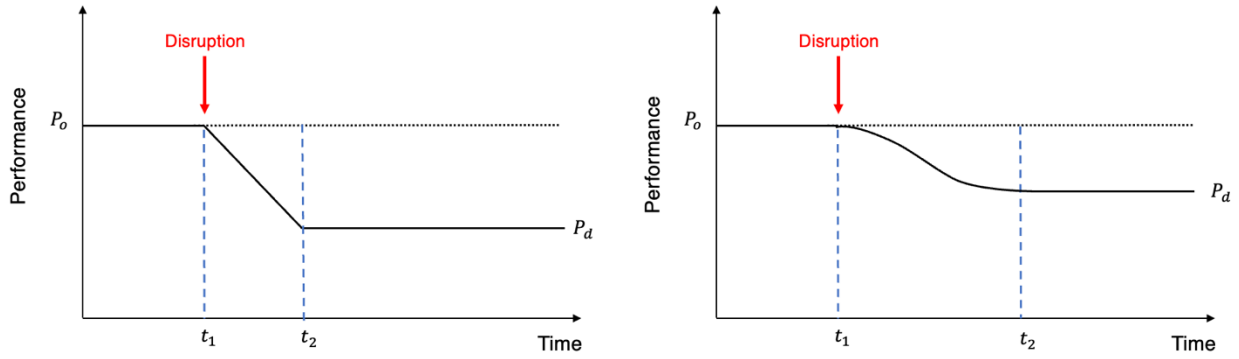

Figure 41. (Left) Rigid Curve with No Recovery, (Right) Smooth Curve with No Recovery

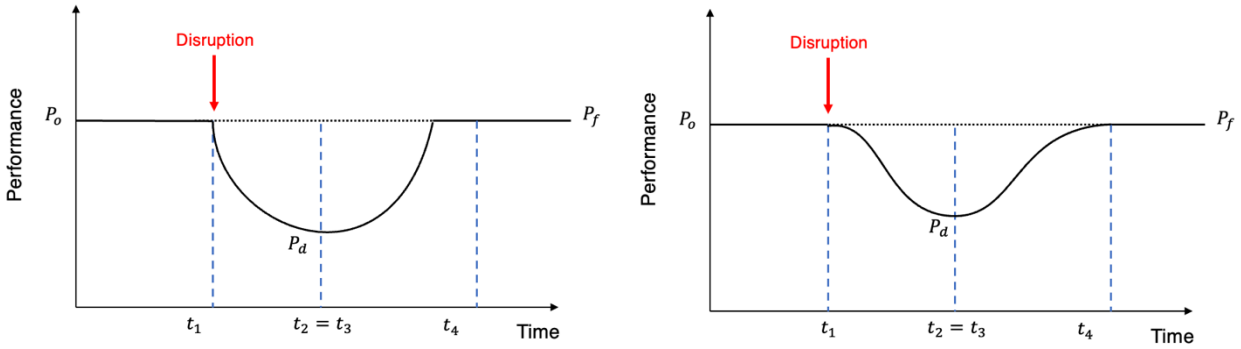

Figure 42. (Left) Rigid "U" Shape, (Right) "Scoop" Shape

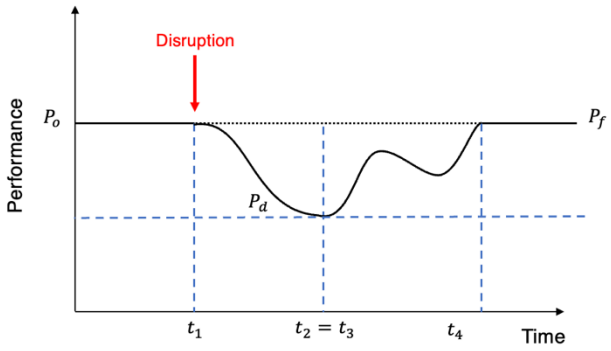

Figure 43. Multiple Minima

The seven metrics identified in the literature review were studied analytically to determine whether they support the computation of a meaningful value for each resilience curve shape in Figure 39–43. Tables 11–17 summarize the computation of each metric when applied to each performance curve shape. These results were completed by Purdue Masters student, Jacqueline Ulmer.

Table 11. Computation of Re_1 for Each Resilience Curve Shape

Shape	Value of Re_1	Works?	Explanation
Baseline	$0 < Re_1 < \infty$	Yes	
Baseline with Performance Lost/Gained	$0 < Re_1 < \infty$	Yes	
Bucket	$0 < Re_1 < \infty$	Yes	
V/U/Scoop	$0 < Re_1 < \infty$	Yes	
No Recovery, Smooth No Recovery	0	Yes*	t_4 does not exist
Multiple Minima	$0 < Re_1 < \infty$	Yes	

The Das et al. metric (Re_1) can be calculated for every shape that shows recovery. When there is no recovery, there is no value for t_4. Therefore, the denominator of the metric goes to ∞, and the value of Re_1 go to zero. Zero is a meaningful value for this metric because it is the smallest possible value obtainable and a curve with no recovery is the least resilient curve possible. However, because there is no t_4, using the equation directly to compute this metric (in the case of no recovery) will result in a value that does not exist. This is noted by the yellow shading and explanation in Table 11.

Table 12. Computation of Re_2 for Each Resilience Curve Shape

Shape	Value of Re_2	Works?	Explanation
Baseline	$1 < Re_2 < \infty$	Yes	
Baseline with Performance Lost/Gained	$1 < Re_2 < \infty$	Yes	
Bucket	$1 < Re_2 < \infty$	Yes	
V/U/Scoop	$1 < Re_2 < \infty$	Yes	
No Recovery, Smooth No Recovery	1	Yes*	t_4 does not exist
Multiple Minima	$1 < Re_2 < \infty$	Yes	

Like the Das et al. metric, the Henry & Ramirez-Marquez metric (Re_2) can be calculated directly for any curve that shows recovery. In the case of no recovery, there is no t_4 at which the recovered performance can be evaluated. To address this, using the value of the performance curve at the end of the simulation is appropriate to estimate performance at t_4. With no recovery, the final performance will be the same as the minimum performance, making $Re_2 = 1$. One is the smallest possible value of this metric and signifies that no recovery is the least resilient case.

Table 13. Computation of Re_3 for Each Resilience Curve Shape

Shape	Value of Re_3	Works?	Explanation
Baseline	$0 < Re_3 < \infty$	Yes	
Baseline with Performance Lost/Gained	$0 < Re_3 < \infty$	Yes	
Bucket	$0 < Re_3 < \infty$	Yes	
V/U/Scoop	$0 < Re_3 < \infty$	Yes	
No Recovery, Smooth No Recovery	∞	No	t_4 does not exist
Multiple Minima	$0 < Re_3 < \infty$	Yes	

The Bruneau et al metric (Re_3) is computable for all shapes that show recovery. When there is no recovery, the metric cannot be calculated because there is no t_4. Unlike the previous two metrics, there is no modification to enable the computation of the metric with no recovery. Integrating from t_1 to the end of the simulation time could result in a smaller value for the metric in a no recovery case, compared to the case with recovery. This would imply that the no recovery case was less resilient. However, computing the metric as-is is not possible because t_4 goes to infinity in the no recovery case. Therefore, Re_3 is not meaningful for the no recovery case, which is signified by the red shading in Table 13.

Table 14. Computation of Re_4 for Each Resilience Curve Shape

Shape	Value of Re_4	Works?	Explanation
Baseline	$0 < Re_4 < 1$	Yes	
Baseline with Performance Lost/Gained	$0 < Re_4 < 1$	Yes	
Bucket	$0 < Re_4 < \infty$	No	$t_1 = t_2$ and $t_3 = t_4$
V/U/Scoop	$0 < Re_4 < 1$	Yes	
No Recovery, Smooth No Recovery	∞	No	t_4 does not exist
Multiple Minima	$0 < Re_4 < 1$	Yes	

The Ayyub et al metric (Re_4) does not return meaningful values for the bucket shape or when there is no recovery. In the bucket shape, the values of *F* and *R* would both be zero, reducing the metric to $\frac{t_1}{t_4 - t_1}$, which is not meaningful unless t_1 is the same for every simulation. Even if this were the case, this metric would return the same value as the Das et al. metric. When there is no recovery, *R* would go to infinity, and therefore the metric would also go to infinity. Therefore, Re_4 is not meaningful for the bucket or no recovery cases, which is signified by the red shading in Table 14.

Table 15. Computation of Re_5 for Each Resilience Curve Shape

Shape	Value of Re_5	Works?	Explanation
Baseline	1	Yes	
Baseline with Performance Lost/Gained	$0 < Re_5 < \infty$	Yes	
Bucket	$0 < Re_5 < \infty$	Yes	
V/U/Scoop	$0 < Re_5 < \infty$	Yes	
No Recovery, Smooth No Recovery	0	Yes*	t_4 does not exist
Multiple Minima	$0 < Re_5 < \infty$	Yes	

The second Henry & Ramirez metric (Re_5) returns meaningful values for every shape when a slight modification is made for the no recovery case to account for the lack of a t_4. If the performance at the end of the simulation is used for $P(t_4)$, the metric returns a value of zero, which correctly implies that no recovery is the least resilient case.

Table 16. Computation of Re_6 for Each Resilience Curve Shape

Shape	Value of Re_6	Works?	Explanation
Baseline	$0 < Re_6 < 1$	Yes	
Baseline with Performance Lost/Gained	$0 < Re_6 < 1$	Yes	
Bucket	$0 < Re_6 < 1$	Yes	
V/U/Scoop	$0 < Re_6 < 1$	Yes	
No Recovery, Smooth No Recovery	DNE	No	t_3, t_4 do not exist
Multiple Minima	$0 < Re_6 < 1$	Yes	

The Yarveisy et al. metric (Re_6) returns meaningful values for every shape except when there is no recovery. In the no recovery case, t_3 and t_4 do not exist, and therefore computing values for adaptive and restorative capacities is impossible. There is no solution for this issue, as indicated by the red shading in Table 16.

Table 17. Computation of Re_7 for Each Resilience Curve Shape

Shape	Value of Re_7	Works?	Explanation
Baseline	$0 < Re_7 < 1$	No	Uses time of minimum performance
Baseline with Performance Lost/Gained	$0 < Re_7 < 1$	No	Uses time of minimum performance
Bucket	0	No	Uses time of minimum performance
V/U/Scoop	$0 < Re_7 < 1$	Yes	
No Recovery, Smooth No Recovery	DNE	No	t_3, t_4 do not exist
Multiple Minima	$0 < Re_7 < 1$	Yes	

For every shape except for "U", "V", "Scoop" and multiple minima, this Cheng et al. metric (Re_7) does not return meaningful values. For any shape that levels out at a stable minimum performance between t_2 and t_3, there is no single time at which we can identify minimum performance. This might lead to a variety of metric values depending on how available software functions locate the minimum of a constant line, making the metric unusable in those cases.

Table 18 summarizes the results discussed above. A green box refers to a metric that produces a meaningful result for that shape. A yellow box indicates that meaningful values can be extracted for the corresponding shape with a slight adjustment to the computation methodology. A red box means the metric does not produce a meaningful result for a curve of that shape.

Table 18. Summary of Analytic Computation of Resilience Metrics for Resilience Curve Shapes

	Baseline	Baseline with performance loss / gain	Bucket	"U", "V", and "Scoop"	No recovery both smooth and rigid	Multiple minima
Metric 1						
Metric 2						
Metric 3						
Metric 4						
Metric 5						
Metric 6						
Metric 7						

From Table 18, we can see that Metrics 1, 2, and 5 are most appropriate for use in a quantitative resilience assessment, as meaningful values can still be extracted for the no recovery case, whereas all other metrics will not be computable. Therefore, these three metrics will be applied in **Chapter 5. Control Effectiveness Validation Results** for Step 4 in the control effectiveness validation procedure.

5. CONTROL EFFECTIVENESS VALIDATION RESULTS

In this chapter, we present the current progress in completing the control effectiveness validation procedure for 18 user-controlled safety controls in MCVT v6. Although there are 25 total safety controls modeled in MCVT v6, 7 controls are automated functionalities in the simulation that are not user-controlled. Therefore, we are unable to evaluate implementation strategies for these controls, even though we can evaluate their control effectiveness. The other 18 safety controls are applicable in at least one of the 16 disruption scenarios discussed in **Chapter 4**, and therefore, in this chapter we present the results of activating each safety control implementation strategy individually in four relevant disruption scenarios.

5.1 MCVT v6 Current Capabilities and Limitations

The following capabilities are supported in MCVT v6 to assess control effectiveness for the 18 user-controlled safety controls:

1. Ability for a user to control the activation or deactivation of all safety controls ($P_{available} = 1\ or\ 0$)
2. Ability for a user to set the intervention rate of all safety controls (informed by values of P_{design} and $P_{implementation}$)

With these two capabilities, we can evaluate how different repair rates affect recovery capabilities with safety controls activated one at a time. Safety controls must be evaluated individually due to current limitations in the scheduling of agent intervention in MCVT v6. The simulation exhibits unanticipated behavior during the activation of sets of safety controls. Specifically, the agent often becomes "stuck" at one intervention, causing the same activity to repeat multiple times. This does not allow for the completion of any other control activity for the remainder of the simulation, and therefore multiple controls cannot be evaluated simultaneously. Therefore, the modeling and evaluation of sets of controls with all high, all low, or mixed control effectiveness values is reserved for the next iteration of the MCVT. This is discussed more in **Chapter 6**.

The assessment of control effectiveness and the identification of habitat/system performance metrics for each safety control has been discussed at length in **Chapter 4**. See Figures 13–32 and Table 10 for the relevant information obtained for Steps 1 and 2 in the control effectiveness validation procedure. Here we focus on the completion of Steps 3 and 4 in the validation procedure, which include using the MCVT to obtain performance metrics as a function of time and using that data to conduct a resilience assessment.

Here we present the simulation of 4 out of the 16 disruption scenarios that allow for the activation and evaluation of 15 user-controlled safety controls in the MCVT. The three other user-controlled safety controls that are not directly applicable in these four scenarios are: (1) *Ability to repair heater in ECLSS thermal control system*, (2) *Ability to repair air supply valve in ECLSS pressure system*, and (3) *Ability to repair evaporator in ECLSS thermal control system*. These safety controls apply for hazardous states that propagate from the damage levels in other subsystems, rather than from the disruption itself. Currently, the disruption scenarios simulated in MCVT v6 have yet to cause these hazardous states and require the activation of these controls. Further testing would be needed to identify the input parameters that lead to the hazardous states that activate these controls. This work could not be completed in the limited amount of time that the MCVT v6 was available for use. Therefore, this task is reserved for future exploration of the MCVT v6, which is discussed more in **Chapter 6**. The four disruption scenarios considered here to evaluate the other 15 safety controls are: (1) Meteorite impact on solar PV arrays, (2) Meteorite impact on nuclear radiator panels, (3) Meteorite impact on structure, and (4) Fire originating near the power storage and distribution systems. For each scenario, we present the performance data obtained through the MCVT, as well as the quantitative resilience assessment completed using the three resilience metrics identified at the end of **Chapter 4**.

5.2 Disruption Scenario 1: Meteorite Impact on Solar PV Arrays

Here we use the disruption scenario of an intensity level 5 meteorite impact hitting the solar PV arrays. The schematic for the hazardous states and safety controls in this disruption scenario is shown in Figure 44.

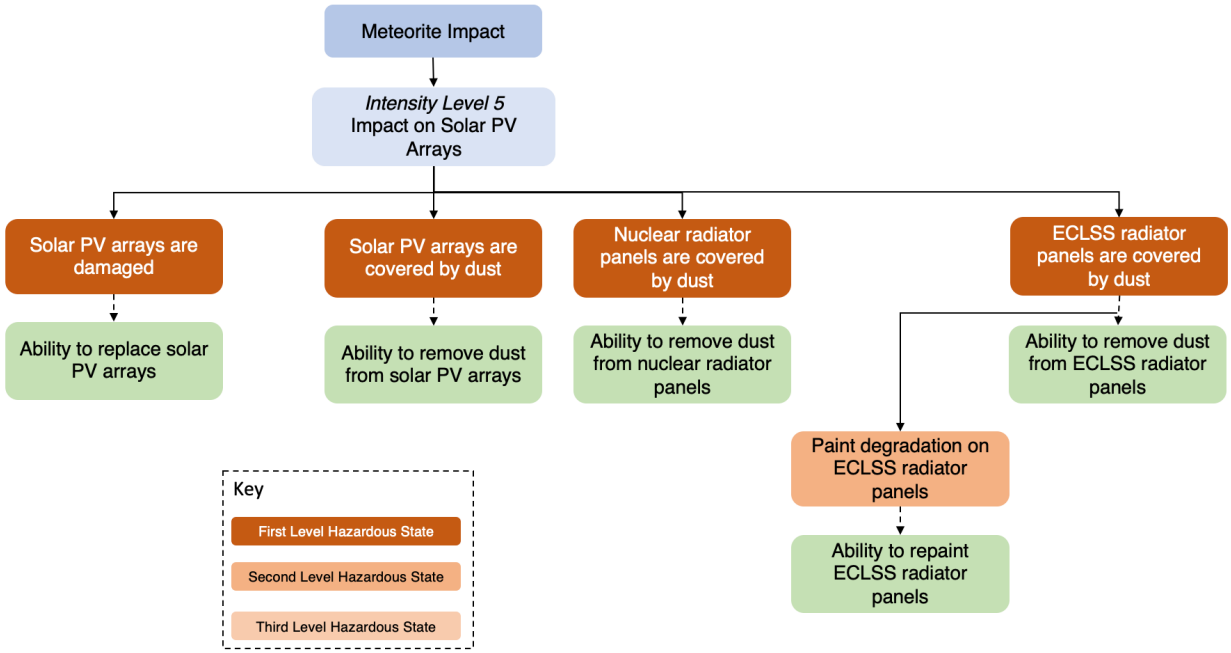

Figure 44. Schematic for Disruption Scenario 1: Meteorite Impact on Solar PV arrays

This disruption scenario includes five hazardous states and five safety controls that have two or more implementation strategies with varying control effectiveness values. Running the MCVT for this scenario allowed for the evaluation of two of these five safety controls: *Ability to remove dust from solar PV arrays* and *Ability to remove dust from nuclear radiator panels*.

Only two controls were evaluated due to unexpected behavior in the simulation for the other four controls. First, we were unable to activate the safety control *Ability to replace solar PV arrays.* The reason for this is not clear, because the safety control should have been activated based on observing over 10% drop in solar power output, which exceeds the FDD threshold. Instead, the safety control *Ability to replace energy storing units* activates four times during the simulation, rather than the control activated in the user interface. This behavior is especially anomalous because the results of this disruption do not indicate damage in the energy storage system or any drop in the number of remaining battery cells. Therefore, these results suggest an issue in the disruption propagation scheme and integrated health management system. This behavior was reported and will be addressed in the next iteration of the MCVT. In addition, the safety control *Ability to remove dust from ECLSS radiator panels* resulted in identical results for all three implementation strategies. No variation in the recovery was observed when varying the repair rate

for this safety control. It is unclear whether this behavior is due to an error in the ECLSS recovery model or if the different repair rates truly do not result in different recoveries for the system. The latter would suggest control effectiveness variations for this safety control do not result in significantly different resilience performance; however, not enough data was present to support this claim. Finally, the safety control *Ability to repaint ECLSS radiator panels* was not evaluated because its corresponding hazardous state was not observed. In the short simulation time, paint degradation due to dust build up was not present for the ECLSS radiator panels. This hazardous state would occur at longer simulation times, which the MCVT cannot currently support. Therefore, in the next sections we present the MCVT performance data gathered for the evaluation of the two safety controls *Ability to remove dust from solar PV arrays* and *Ability to remove dust from nuclear radiator panels*. Table 19 shows the organization of these controls into high, low, and intermediate categories.

Table 19. Categorization of High, Low, and Intermediate Safety Control Implementation Strategies for Disruption Scenario 1

Safety Control	**High control effectiveness implementation strategy**	**Intermediate control effectiveness implementation strategy**	**Low control effectiveness implementation strategy**
Ability to remove dust from solar PV arrays (**SC798**)	Built-in brush automatically removes dust from solar PV arrays	Human agent brushes dust from solar PV arrays	Robot agent brushes dust from solar PV arrays
Ability to remove dust from nuclear radiator panels (**SC800**)	Built-in brush automatically removes dust from nuclear radiator panels	Human agent brushes dust from nuclear radiator panels	Robot agent brushes dust from nuclear radiator panels

5.2.1 Obtain Performance Metrics as a Function of Time

As previously mentioned, the MCVT supports simulations on the order of minutes due to computational limitations. For all simulations shown here, the MCVT is run for a simulation time of 120 seconds. Therefore, nominal repair rates were identified for each safety control that would allow for a full recovery to maximum achievable performance at least once in the 120 second simulation. Then, the implementation repair rates are obtained by multiplying the nominal rate by the corresponding values of P_{design} and $P_{implementation}$ for each implementation strategy. These are the repair rates entered by the user for each MCVT simulation. The nominal repair rates and

implementation repair rates for this disruption scenario are shown in Table 20, followed by the results of activating each safety control individually.

Table 20. Safety Control Repair Rate Inputs for Agent Model in Disruption Scenario 1

Safety control	**Nominal Repair Rate**	**Implementation Strategy**	P_{design}	$P_{implementation}$	**Implementation Repair Rate**
Ability to remove dust from solar PV arrays (**SC798**)	20 $mg/cm^2/s$	Human agent brushes dust from solar PV arrays	0.9	0.7	17.1 $mg/cm^2/s$
		Robot agent brushes dust from solar PV arrays	0.9	0.4	12.6 $mg/cm^2/s$
		Brush automatically removes dust from solar PV arrays	0.9	0.95	7.2 $mg/cm^2/s$
Ability to remove dust from nuclear radiator panels (**SC800**)	10 $mg/cm^2/s$	Human agent brushes dust from nuclear radiator panels	0.9	0.7	8.55 $mg/cm^2/s$
		Robot agent brushes dust from nuclear radiator panels	0.9	0.4	6.3 $mg/cm^2/s$
		Brush automatically removes dust from ECLSS radiator panels	0.9	0.95	3.6 $mg/cm^2/s$

In the first three simulations for this disruption scenario, SC798 was activated in isolation and assigned the three different rates given in Table 20. Each run represented the activation of the corresponding implementation strategy, which subsequently caused variations in the recovery of the system after dust accumulates on the solar PV arrays. Figure 45 demonstrates the difference between resilience curves for these three safety control implementation strategies as a result of varying the repair rate based on the values of P_{design} and $P_{implementation}$ for each implementation strategy.

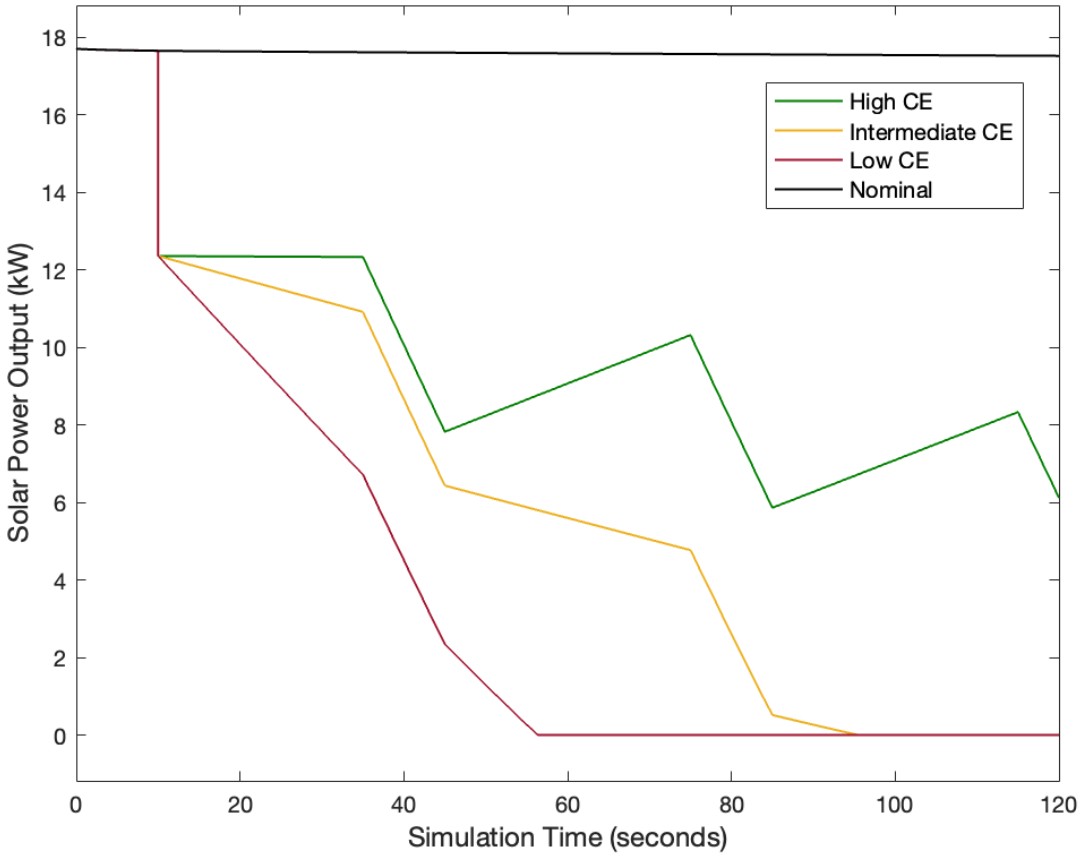

Figure 45. Solar Power Output for High, Low, and Intermediate Control Effectiveness Implementations of SC798 in Disruption Scenario 1

In this disruption scenario, the solar PV arrays are impacted by a meteorite. For all three cases, the impact resulted in an immediate drop in the solar power output, as 30% of the solar panels are damaged and nonfunctional. Since the safety control *Ability to replace solar PV arrays* cannot be activated in the simulation, the power output regained from cleaning dust is limited. This is noticeable in the green curve for the high CE case, as the solar power output flatlines and cannot exceed the level to which power output dropped due to panel damage. In order to return to the nominal level, the damaged solar PV arrays must be replaced in addition to cleaning the PV arrays.

Although regained performance is limited, we do see the expected differences in recovery curves based on control effectiveness. Dust accumulates on the solar PV arrays starting at 10 seconds and continues for the entire simulation. The agent activates three separate times, with the recovery time periods being separated by a delay that represents the agent traveling back to inventory before

returning to repeat the activity. The rate of cleaning for all three implementations is not high enough to keep the solar PV arrays clean throughout the simulation. In the high control effectiveness case, the first cleaning period has a low amount of dust, and therefore, the agent can remove all dust that has accumulated. This is shown by the flat line from 10 seconds to 45 seconds. However, as dust continually accumulates, we see solar power output decrease over time, as more dust accumulates than the agent can clean in each intervention period. Still, in the high control effectiveness case, the cleaning rate is high enough that solar power output is never lost completely. Conversely, in the low and intermediate cases, the solar PV arrays eventually become fully covered by dust, and solar power output drops to 0 kW. The intermediate control effectiveness case, however, allows the system to survive longer than the low control effectiveness case, as power output drops to 0 kW at a later time.

In addition to activating the agent to clean solar PV arrays, we ran the MCVT to simulate the three implementations for cleaning the nuclear radiator panels. Since the solar PV arrays are impacted by a meteorite in this scenario, nuclear power output does not have an immediate drop in the resilience curves, but rather a steady decrease due to dust accumulation. Like SC798, we do see the expected differences in recovery curves based on control effectiveness. This control differs, however, because the first agent intervention in the activation of this control is actually SC355: *Ability to repair energy storage units*. Since all safety controls were activated in isolation, the activation of this control is incorrect, because it was not only deactivated for these simulations, but no damage to the battery cells occurs due to this disruption. Because of this safety control activating, we see in Figure 46 that the nuclear power output is allowed to drop to 0 kW right away for all three cases, as the agent is delayed by control activities in the energy storage system.

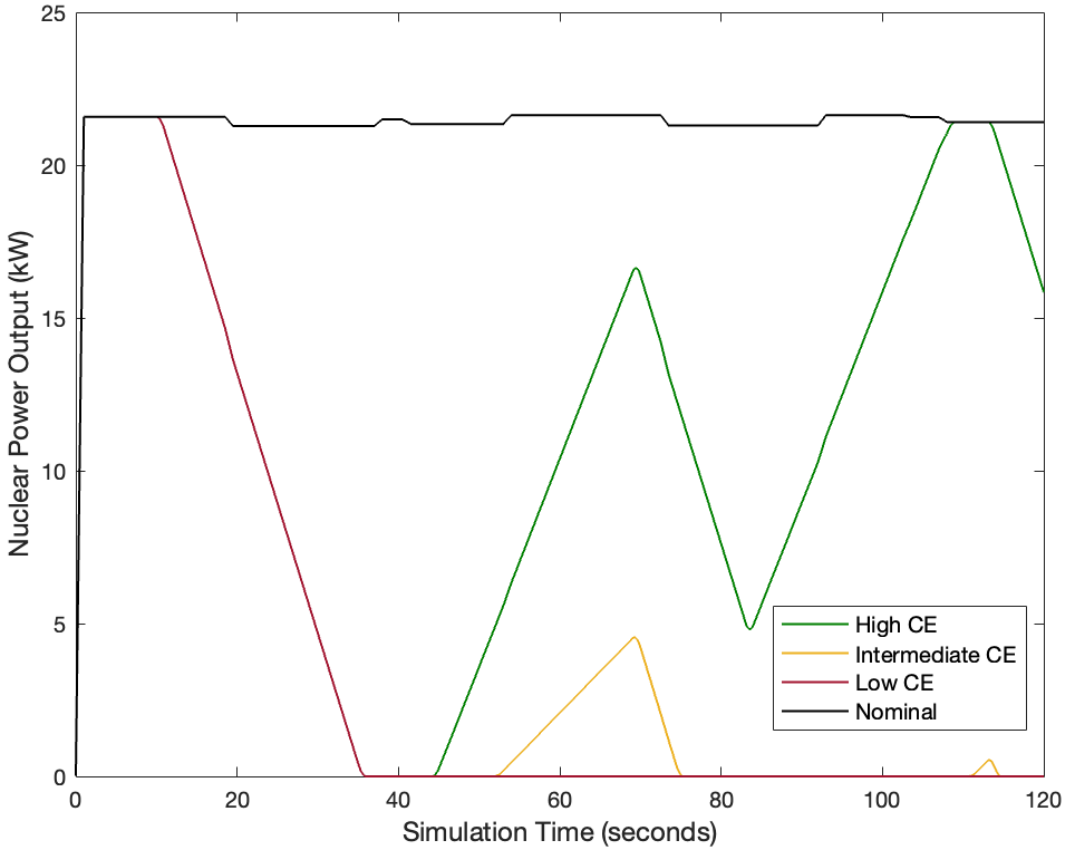

Figure 46. Nuclear Power Output for High, Low, and Intermediate Control Effectiveness Implementations of SC800 in Disruption Scenario 1

After the initial drop to 0 kW, the agent activates three separate times to clean the nuclear radiator panels, with the recovery time periods also being separated by a delay that represents the agent traveling back to inventory before returning to repeat the activity. As shown in Figure 46, the rate of cleaning for the high control effectiveness case is greater than the accumulation rate; therefore, nuclear power output returns to the nominal level after the agent completes the activity twice. Conversely, the low control effectiveness case does not exhibit any increase in the nuclear power output because the cleaning rate is much lower than the accumulation rate. Therefore, power output remains at 0 kW for the entire simulation. Finally, the intermediate control effectiveness case exhibits short-term recoveries in the power output, but the dust accumulation rate is still higher than the repair rate, and therefore the power output remains at 0 kW for most of the simulation time.

5.2.2 Use Performance Curves to Assess Resilience

The three resilience metrics identified in **Chapter 4** can be used to quantify habitat resilience for the two safety controls activated in this disruption scenario. The resilience metric values computed for each safety control and control effectiveness case are shown in Tables 21 and 22.

Table 21. Resilience Metric Values for SC798 in Disruption Scenario 1

Safety Control: Ability to remove dust from solar PV arrays (SC798)			
	Low Control Effectiveness	**Intermediate Control Effectiveness**	**High Control Effectiveness**
$Re_1 = \frac{1}{t_4 - t_1}$	0	0	0
$Re_2 = \frac{P(t_4)}{P(t_1)}$	0.528	0.872	1
$Re_5 = \frac{P(t_4) - min(P)}{P(t_1) - min(P)}$	0	0	0

Table 22. Resilience Metric Values for SC800 in Disruption Scenario 1

Safety Control: Ability to remove dust from nuclear radiator panels (SC800)			
	Low Control Effectiveness	**Intermediate Control Effectiveness**	**High Control Effectiveness**
$Re_1 = \frac{1}{t_4 - t_1}$	0	0.02	0.02
$Re_2 = \frac{P(t_4)}{P(t_1)}$	DNE	DNE	DNE
$Re_5 = \frac{P(t_4) - min(P)}{P(t_1) - min(P)}$	0	0.21	0.76

5.3 Disruption Scenario 2: Meteorite Impact on Nuclear Radiator Panels

Here we use the disruption scenario of an intensity level 5 meteorite impact hitting the nuclear radiator panels. Figure 47 shows the schematic for the hazardous states and safety controls for this disruption scenario.

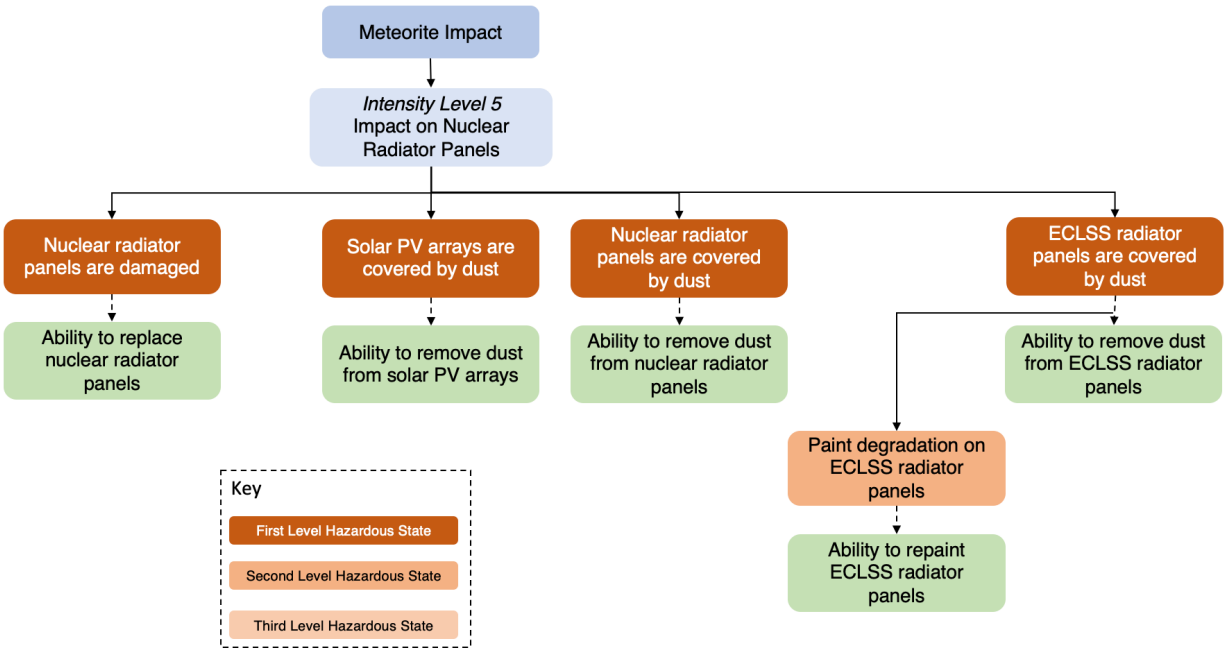

Figure 47. Schematic for Disruption Scenario 2: Meteorite Impact on Nuclear Radiator Panels

This disruption scenario includes five safety controls that have two or more implementation strategies with varying control effectiveness values. Just as for the first disruption scenario, running the MCVT for this scenario allowed for the evaluation of two of these five safety controls: *Ability to remove dust from solar PV arrays* and *Ability to remove dust from nuclear radiator panels*. Only these two controls were evaluated due to the same unexpected behavior in the simulation for the other three controls previously discussed. In particular, *Ability to replace nuclear radiator panels* never activated, even with over 10% drop in nuclear power output. This behavior was also reported and will be addressed in the next iteration of the MCVT. Therefore, in the next sections we present the MCVT performance data gathered for the evaluation of the same two safety controls *Ability to remove dust from solar PV arrays* and *Ability to remove dust from*

nuclear radiator panels for this disruption scenario. Table 23 shows the organization of these controls into high, low, and intermediate categories.

Table 23. Categorization of High, Low, and Intermediate Safety Control Implementation Strategies for Disruption Scenario 2

Safety Control	**High control effectiveness implementation strategy**	**Intermediate control effectiveness implementation strategy**	**Low control effectiveness implementation strategy**
Ability to remove dust from solar PV arrays (**SC798**)	Built-in brush automatically removes dust from solar PV arrays	Human agent brushes dust from solar PV arrays	Robot agent brushes dust from solar PV arrays
Ability to remove dust from nuclear radiator panels (**SC800**)	Built-in brush automatically removes dust from nuclear radiator panels	Human agent brushes dust from nuclear radiator panels	Robot agent brushes dust from nuclear radiator panels

5.3.1 Obtain Performance Metrics as a Function of Time

As previously mentioned, the MCVT supports simulations on the order of minutes due to computational limitations. For all simulations shown here, the MCVT is run for a simulation time of 120 seconds. Therefore, nominal repair rates were identified for each safety control that would allow for a full recovery to maximum achievable performance at least once in the 120 second simulation. Then, the implementation repair rates are obtained by multiplying the nominal rate by the corresponding values of P_{design} and $P_{implementation}$ for each implementation strategy. These are the repair rates entered by the user for each MCVT simulation. The nominal repair rates and implementation repair rates for this disruption scenario are shown in Table 24, followed by the results of activating each safety control individually.

Table 24. Safety Control Repair Rate Inputs for Agent Model in Disruption Scenario 2

Safety control	**Nominal Repair Rate**	**Implementation Strategy**	P_{design}	$P_{implementation}$	**Implementation Repair Rate**
Ability to remove dust from solar PV arrays (**SC798**)	$20\ mg/cm^2/s$	Human agent brushes dust from solar PV arrays	0.9	0.7	$17.1\ mg/cm^2/s$
		Robot agent brushes dust from solar PV arrays	0.9	0.4	$12.6\ mg/cm^2/s$
		Brush automatically removes dust from solar PV arrays	0.9	0.95	$7.2\ mg/cm^2/s$
Ability to remove dust from nuclear radiator panels (**SC800**)	$10\ mg/cm^2/s$	Human agent brushes dust from nuclear radiator panels	0.9	0.7	$8.55\ mg/cm^2/s$
		Robot agent brushes dust from nuclear radiator panels	0.9	0.4	$6.3\ mg/cm^2/s$
		Brush automatically removes dust from ECLSS radiator panels	0.9	0.95	$3.6\ mg/cm^2/s$

In the first three simulations for this disruption scenario, SC798 was activated in isolation and assigned the repair rates given in Table 24. Each run represented the activation of the corresponding implementation strategy, which subsequently caused variations in the recovery of the system after dust accumulates on the solar PV arrays. Figure 48 demonstrates the difference between resilience curves for these three safety control implementation strategies as a result of varying the repair rate based on the values of P_{design} and $P_{implementation}$ for each implementation strategy.

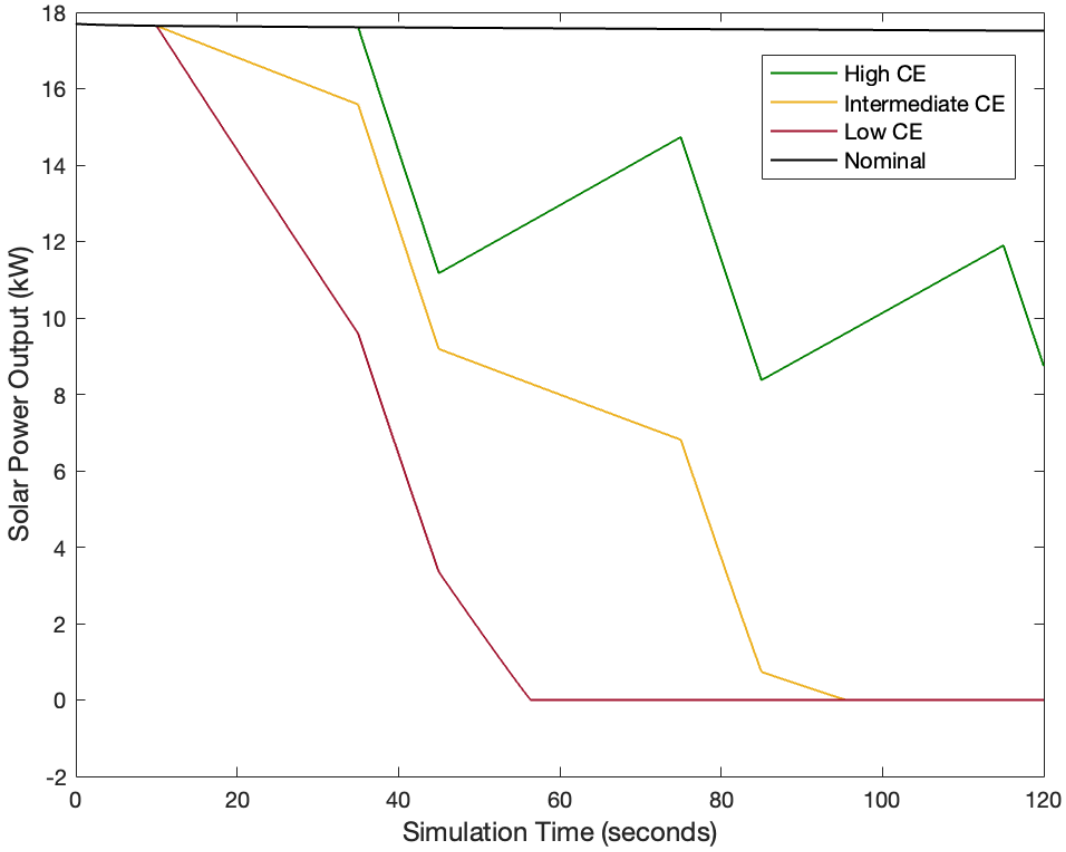

Figure 48. Solar Power Output for High, Low, and Intermediate Control Effectiveness Implementations of SC798 in Disruption Scenario 2

In this disruption scenario, the nuclear radiator panels are now impacted by a meteorite. Therefore, solar power output no longer has an immediate drop in the resilience curves, but rather a steady decrease due to dust accumulation. Like the first disruption scenario, we again see the expected differences in recovery curves based on control effectiveness. The agent activates three separate times to clean the solar PV arrays, with the recovery time periods again being separated by a delay that represents the agent traveling back to inventory before returning to repeat the activity. As seen in Figure 48, the rate of cleaning for the high control effectiveness case is greater than the accumulation rate; therefore, solar power output remains at the nominal level during the first agent intervention time period. After that first intervention, solar power output steadily drops over the rest of the simulation, however, the cleaning activity is fast enough to prevent the solar power output from dropping to 0 kW. In the lower and intermediate cases, we see that the cleaning rate

is lower than accumulation rate, as the solar power output starts to drop immediately after the accumulation starts at 10 seconds. These rates are slow enough that solar power output drops very quickly to 0 kW in the low control effectiveness case, while the intermediate case allows the system to survive longer and drop to 0 kW later in the simulation.

The nuclear radiator panels are impacted by a meteorite in this scenario, and therefore, the recovery curves for nuclear power output more closely match the solar power output case in the first disruption scenario. In all three control effectiveness cases here, there is an immediate drop in the nuclear power output, as 60% of the nuclear radiator panels are damaged and nonfunctional. Since the safety control *Ability to replace nuclear radiator panels* cannot be activated in the simulation, the power output regained from cleaning dust is limited. This is again noticeable in the green curve for the high CE case (Figure 49), as the nuclear power output flatlines at the end of the simulation and cannot exceed the level to which power output drops due to panel damage. In order to return to the nominal level, the damaged nuclear radiator panels must be replaced in addition to cleaned.

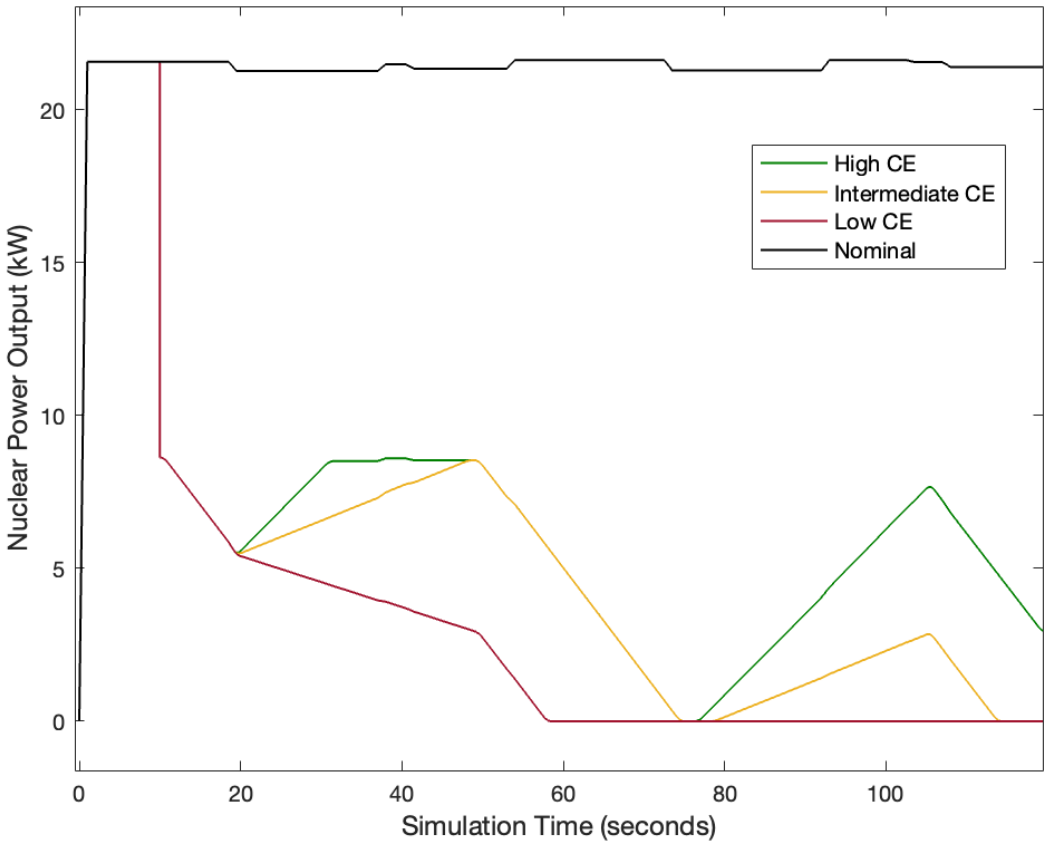

Figure 49. Nuclear Power Output for High, Low, and Intermediate Control Effectiveness Implementations of SC800 in Disruption Scenario 2

Although regained performance is limited, we again see the expected differences in recovery curves based on control effectiveness. Dust accumulates on the nuclear radiator starting at 10 seconds and continues for the entire simulation. Unlike the first disruption scenario, when attempting to clean the nuclear radiator panels, the first agent intervention is actually SC355: *Ability to repair energy storage units*. Since all safety controls were activated in isolation, the activation of this control is incorrect, because it was not only deactivated for these simulations, but no damage to the battery cells occurs due to this disruption. Because of this safety control activating, we see in Figure 46 that the nuclear power output is allowed to drop to 0 kW right away for all three cases, as the agent is delayed by control activities in the energy storage system.

The agent then activates three separate times, with the recovery time periods being separated by a delay that represents the agent traveling back to inventory before returning to repeat the activity.

The rate of cleaning for all three implementations is insufficient to keep the nuclear radiator panels clean throughout the simulation. As shown in Figure 49, the rate of cleaning for the high control effectiveness case is greater than the accumulation rate; therefore, nuclear power output flatlines at maximum achievable power output around 110 seconds. Conversely, the low control effectiveness does not exhibit any increase in the nuclear power output because the cleaning rate is much lower than the accumulation rate. Therefore, power output remains at 0 kW for the entire simulation. Finally, the intermediate control effectiveness case exhibits minimal recovery in the power output for short time periods, but accumulation rate is still higher than the repair rate, and therefore the power output also remains at 0 kW for most of the simulation time.

5.3.2 Use Performance Curves to Assess Resilience

The three resilience metrics identified in **Chapter 4** can be used to quantify habitat resilience for the two safety controls activated in this disruption scenario. The resilience metric values computed for each safety control and control effectiveness case are shown in Tables 25 and 26.

Table 25. Resilience Metric Values for SC798 in Disruption Scenario 2

Safety Control: Ability to remove dust from solar PV arrays (SC798)			
	Low Control Effectiveness	**Intermediate Control Effectiveness**	**High Control Effectiveness**
$Re_1 = \frac{1}{t_4 - t_1}$	0	0	0
$Re_2 = \frac{P(t_4)}{P(t_1)}$	0.545	0.858	1
$Re_5 = \frac{P(t_4) - min(P)}{P(t_1) - min(P)}$	0	0	0

Table 26. Resilience Metric Values for SC800 in Disruption Scenario 2

Safety Control: Ability to remove dust from nuclear radiator panels (SC800)			
	Low Control Effectiveness	**Intermediate Control Effectiveness**	**High Control Effectiveness**
$Re_1 = \frac{1}{t_4 - t_1}$	0	0.0256	0.05
$Re_2 = \frac{P(t_4)}{P(t_1)}$	0.33	1	1
$Re_5 = \frac{P(t_4) - min(P)}{P(t_1) - min(P)}$	0	1	1

5.4 Disruption Scenario 3: Meteorite Impact on Structure at Location 2

Here we use the disruption scenario of an intensity level 5 meteorite impact hitting the structural protective layer at location 2 (Figure 50).

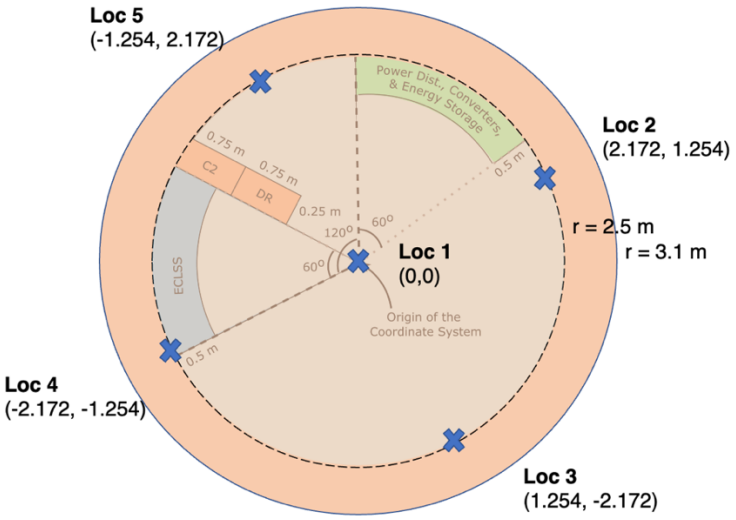

Figure 50. Meteorite Impact Locations on the Structural Protective Layer

The schematic for the hazardous states and safety controls in this disruption scenario are shown in Figure 51.

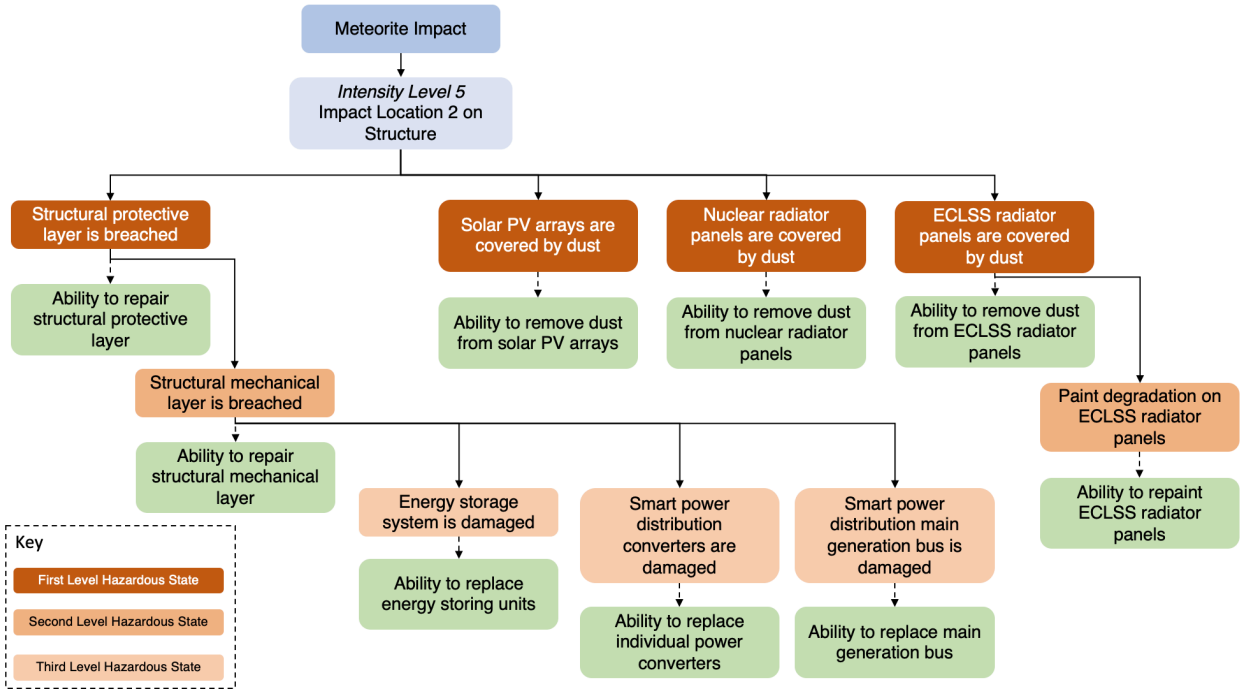

Figure 51. Schematic for Disruption Scenario 3: Meteorite Impact on Structure Location 2

This disruption scenario includes nine safety controls that have two or more implementation strategies with varying control effectiveness values. Running the MCVT for this scenario allowed for the evaluation of four of these nine safety controls: *Ability to remove dust from solar PV arrays*, *Ability to remove dust from nuclear radiator panels*, *Ability to repair structural mechanical layer*, and *Ability to repair energy storage units*. Only these four controls were evaluated due to unexpected behavior in the simulation. First, like with the first two disruption scenarios, the safety control *Ability to remove dust from ECLSS radiator panels* resulted in identical results for all three implementation strategies. No variation in the recovery was observed when varying the repair rate, and therefore we cannot evaluate the different control effectiveness implementation strategies. In addition, the safety control *Ability to repaint ECLSS radiator panels* was not evaluated because its corresponding hazardous state was not observed. The safety control *Ability to repair structural protective layer* did activate in the agent model of the simulation; however, no recovery was ever observed in the structural protective layer, no matter what value of repair rate was entered. This issue has been reported and will be addressed in future iterations of the MCVT. For the two power distribution controls, *Ability to replace individual power converters* and *Ability to replace main generation bus,* the agent did not activate because no damage state propagates to the power

distribution system. Therefore, there is no damage to repair, and these controls will not be activated in this disruption scenario. In the next sections we present the MCVT performance data gathered for the evaluation of the four safety controls *Ability to remove dust from solar PV arrays*, *Ability to remove dust from nuclear radiator panels*, *Ability to repair structural mechanical layer*, and *Ability to repair energy storage units*. Table 27 shows the organization of these controls into high, low, and intermediate categories.

Table 27. Categorization of High, Low, and Intermediate Safety Control Implementation Strategies for Disruption Scenario 3

Safety Control	**High control effectiveness implementation strategy**	**Intermediate control effectiveness implementation strategy**	**Low control effectiveness implementation strategy**
Ability to remove dust from solar PV arrays (**SC798**)	Built-in brush automatically removes dust from solar PV arrays	Human agent brushes dust from solar PV arrays	Robot agent brushes dust from solar PV arrays
Ability to remove dust from nuclear radiator panels (**SC800**)	Built-in brush automatically removes dust from nuclear radiator panels	Human agent brushes dust from nuclear radiator panels	Robot agent brushes dust from nuclear radiator panels
Ability to repair structural mechanical layer (**SC795**)	Human agent uses vacuum cementing to fill breach in structure	Human agent applies a flexible patch to breach in structure	Robot agent applies a flexible patch to breach in structure
Ability to remove dust from nuclear radiator panels (**SC800**)	Human agent replaces energy storing units		Robot agent replaces energy storing units

5.4.1 Obtain Performance Metrics as a Function of Time

As previously mentioned, the MCVT supports simulations on the order of minutes due to computational limitations. For all simulations shown here, the MCVT is run for a simulation time of 120 seconds. Therefore, nominal repair rates were identified for each safety control that would allow for a full recovery to maximum achievable performance at least once in the 120 second simulation. Then, the implementation repair rates are obtained by multiplying the nominal rate by the corresponding values of P_{design} and $P_{implementation}$ for each implementation strategy. These are the repair rates entered by the user for each MCVT simulation. The nominal repair rates and implementation repair rates for this disruption scenario are shown in Table 20, followed by the results of activating each safety control individually.

Table 28. Safety Control Repair Rate Inputs for Agent Model in Disruption Scenario 3

Safety control	**Nominal Repair Rate**	**Implementation Strategy**	P_{design}	$P_{implementation}$	**Implementation Repair Rate**
Ability to remove dust from solar PV arrays (**SC798**)	$20\ mg/cm^2/s$	Human agent brushes dust from solar PV arrays	0.9	0.7	$17.1\ mg/cm^2/s$
		Robot agent brushes dust from solar PV arrays	0.9	0.4	$12.6\ mg/cm^2/s$
		Brush automatically removes dust from solar PV arrays	0.9	0.95	$7.2\ mg/cm^2/s$
Ability to remove dust from nuclear radiator panels (**SC800**)	$10\ mg/cm^2/s$	Human agent brushes dust from nuclear radiator panels	0.9	0.7	$8.55\ mg/cm^2/s$
		Robot agent brushes dust from nuclear radiator panels	0.9	0.4	$6.3\ mg/cm^2/s$
		Brush automatically removes dust from ECLSS radiator panels	0.9	0.95	$3.6\ mg/cm^2/s$
Ability to repair breach in structural mechanical layer (**SC795**)	$0.003\ m^3/s$	Human agent uses vacuum cementing to fill breach in structure	0.8	0.5	$0.0012\ m^3/s$
		Human agent applies a flexible patch to breach in structure	0.4	0.9	$0.00108\ m^3/s$
		Robot agent applies a flexible patch to breach in structure	0.4	0.7	$0.00084\ m^3/s$
Ability to replace energy storing units (**SC355**)	$10\ 1/s$	Human agent replaces energy storing units	0.9	0.9	$8.1\ 1/s$
		Robot agent replaces energy storing units	0.9	0.2	$1.8\ 1/s$

For the meteorite disruption considered in this example, the meteorite impact location on the structural protective layer is held constant when activating all safety controls. The intensity level of this impact is Level 5, which results in a 0.05 m radius hole in the structural mechanical layer. Each of the four safety controls evaluated here were activated in isolation, with all other control activities deactivated.

The first safety control evaluated is the repair activity to fix the breach in the structure caused by the meteorite impact. Using the repair rates identified for each strategy in Table 28, Figure 52

shows the differences for the three implementations. For high and intermediate control effectiveness, the hole was fully repaired by the same time (97 seconds), however, the different repair rates caused by differences in control effectiveness affected how much of the hole was filled during each intervention. In these cases, the agent had to repeat the repair activity three times until complete recovery was achieved and the hole radius was 0 m. Conversely, the low control effectiveness case does not enable complete repair of the hole for this simulation. In this case, the agent activates a fourth time; however, the low control effectiveness repair rate is not sufficient to recover in the 120 second time frame.

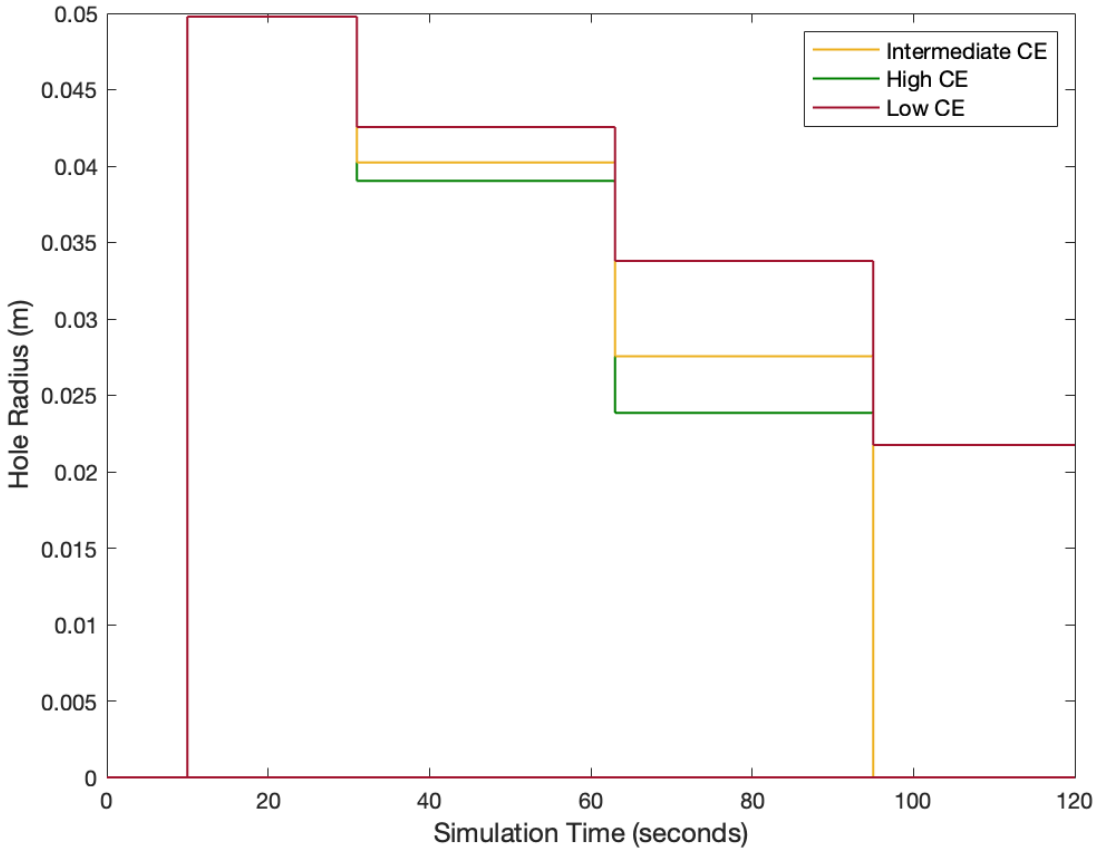

Figure 52. Hole Radius for High, Low, and Intermediate Control Effectiveness Implementations of SC795 in Disruption Scenario 3

Although the hole is being repaired by an agent, the MCVT does not exhibit recovery in interior environment temperature and pressure after the hole is fully repaired. This is due to limitations in

ECLSS, as the current model does not support re-pressurization after a breach. Improving that model to enable temperature and pressure recovery would be required to improve resilience of the system in this scenario.

The second safety control evaluated is the repair activity to replace energy storage units that have been damaged from a breach in the structural mechanical layer. This safety control had only two implementation strategies categorized as high and low based on the values of control effectiveness. Using the repair rates identified for each strategy in Table 28, Figure 53 shows the differences for the two implementations.

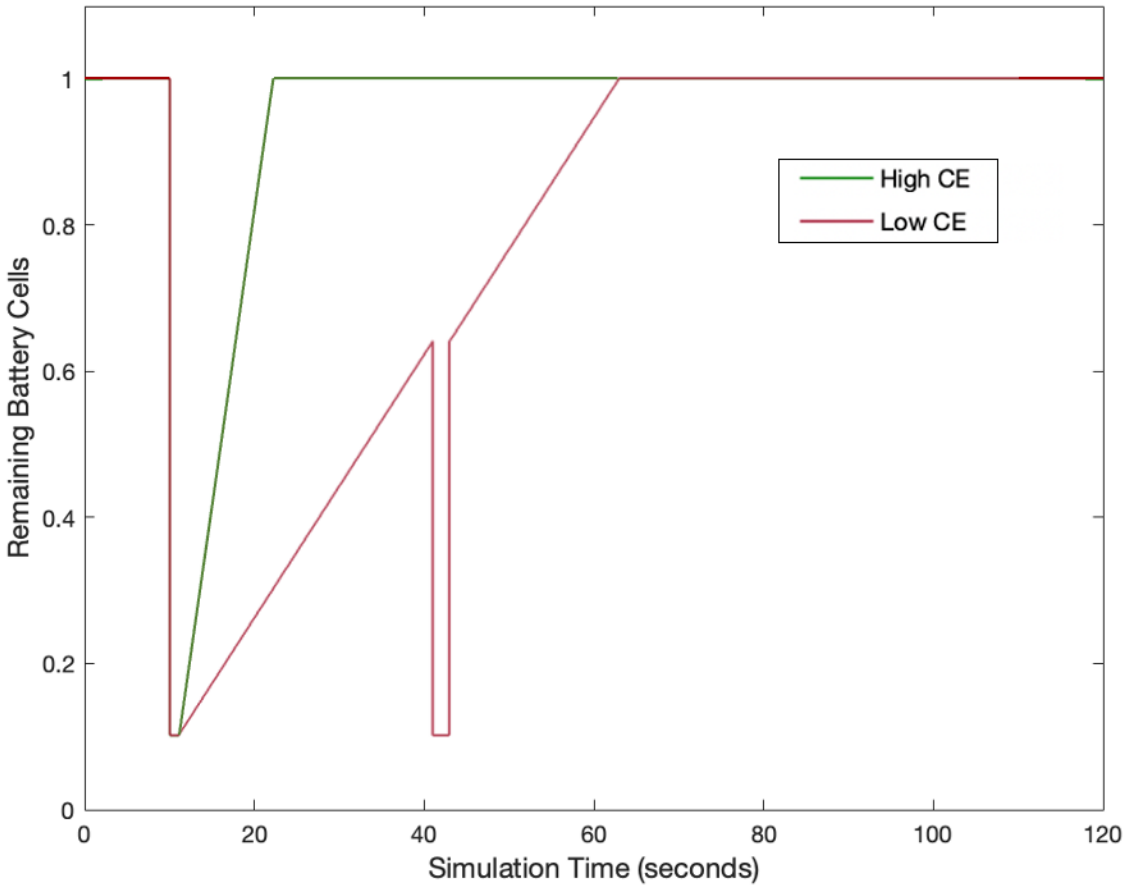

Figure 53. Remaining Battery Cells for High and Low Control Effectiveness Implementation of SC355 in Disruption Scenario 3

Figure 53 shows that both implementations for this control eventually result in the replacement of all damaged battery cells. However, the low control effectiveness implementation indicates the need for two repetitions of this repair activity before full recovery, while one iteration of the repair was sufficient to fully recover in the high control effectiveness case. In addition to looking at the remaining battery cells as a performance metric for this control, we can look at maximum energy storage as well. This performance metric presents interesting behavior that reveals the need for an additional control activity to enable the charging of the new battery cells installed from completing SC355 (Figure 54).

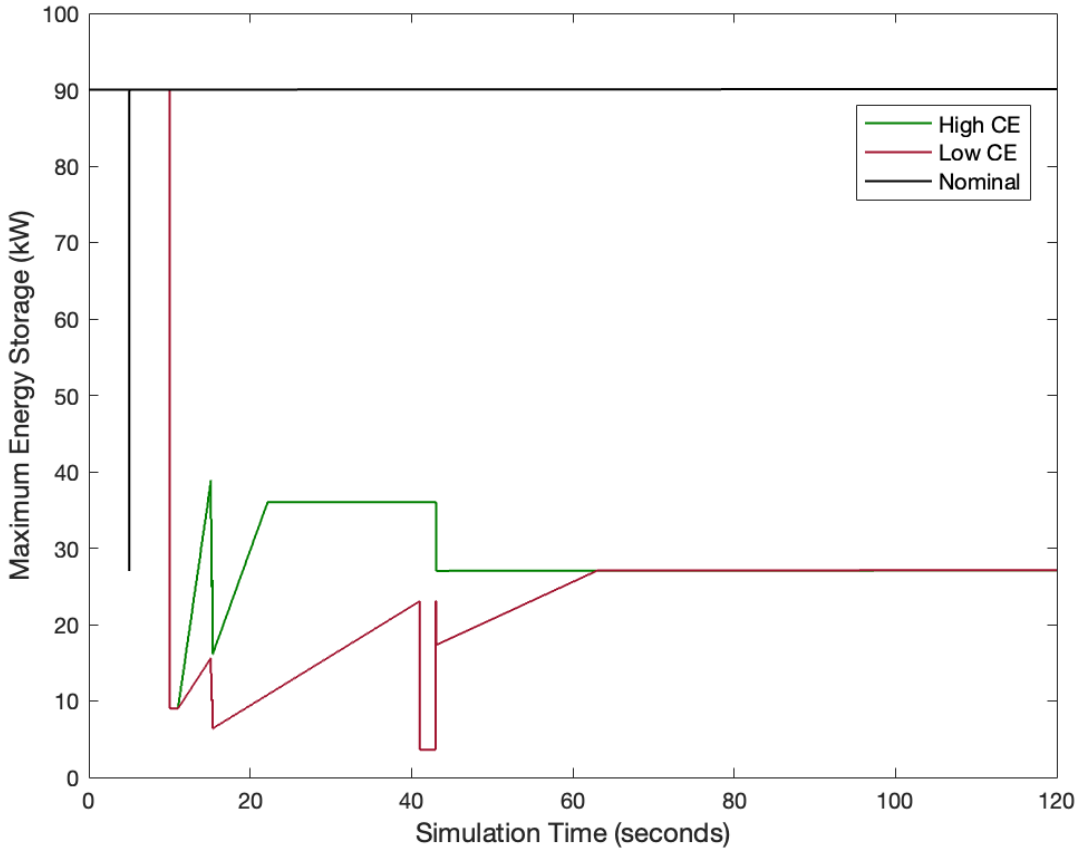

Figure 54. Maximum Energy Storage for High and Low Control Effectiveness Implementations of SC355 in Disruption Scenario 3

Figure 54 shows that although all battery cells have been replaced, maximum energy storage does not return to its nominal value because damage in the power distribution system has not been repaired by an agent, and therefore recharging is not possible. Based on the results from activating

only one safety control to replace battery cells, we can see that without recharging, replacement is not effective at creating a resilient power system. The completion of the secondary safety control would be required for the maximum energy storage system to return to the nominal value and improve overall resilience.

The third safety control evaluated is the repair activity to remove dust that has accumulated on the solar PV arrays. Using the repair rates identified for each strategy in Table 28, Figure 55 shows the differences for the three implementations.

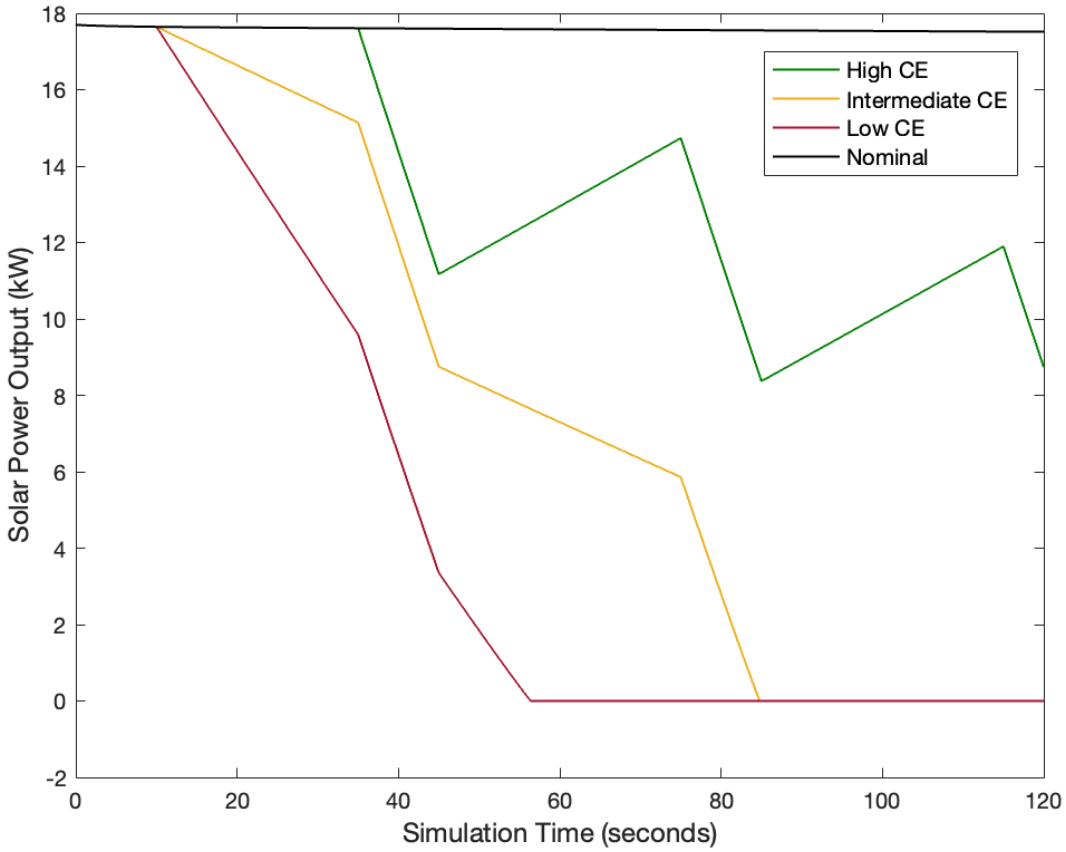

Figure 55. Solar Power Output for High, Low, and Intermediate Control Effectiveness Implementations of SC798 in Disruption Scenario 3

The agent activates three separate times to clean the solar PV arrays, with the recovery time periods again being separated by a delay that represents the agent traveling back to inventory before returning to repeat the activity. As shown in Figure 55, the rate of cleaning for the high control

effectiveness case is greater than the accumulation rate; therefore, solar power output remains at the nominal level during the first agent intervention time period. After that first intervention, solar power output steadily drops over the rest of the simulation; however, the cleaning activity is effective enough to prevent the solar power output from dropping to 0 kW. In the lower and intermediate cases, the cleaning rate is lower than accumulation rate because the solar power output starts to drop immediately after the accumulation starts at 10 seconds. These rates are insufficient to handle the coming dust, and solar power output drops very quickly to 0 kW in the low control effectiveness case, while the intermediate case allows the system to survive longer and drop to 0 kW later in the simulation.

The fourth safety control evaluated is the repair activity to remove dust that has accumulated on the nuclear radiator panels. Using the repair rates identified for each strategy in Table 28, Figure 56 shows the differences for the three implementations.

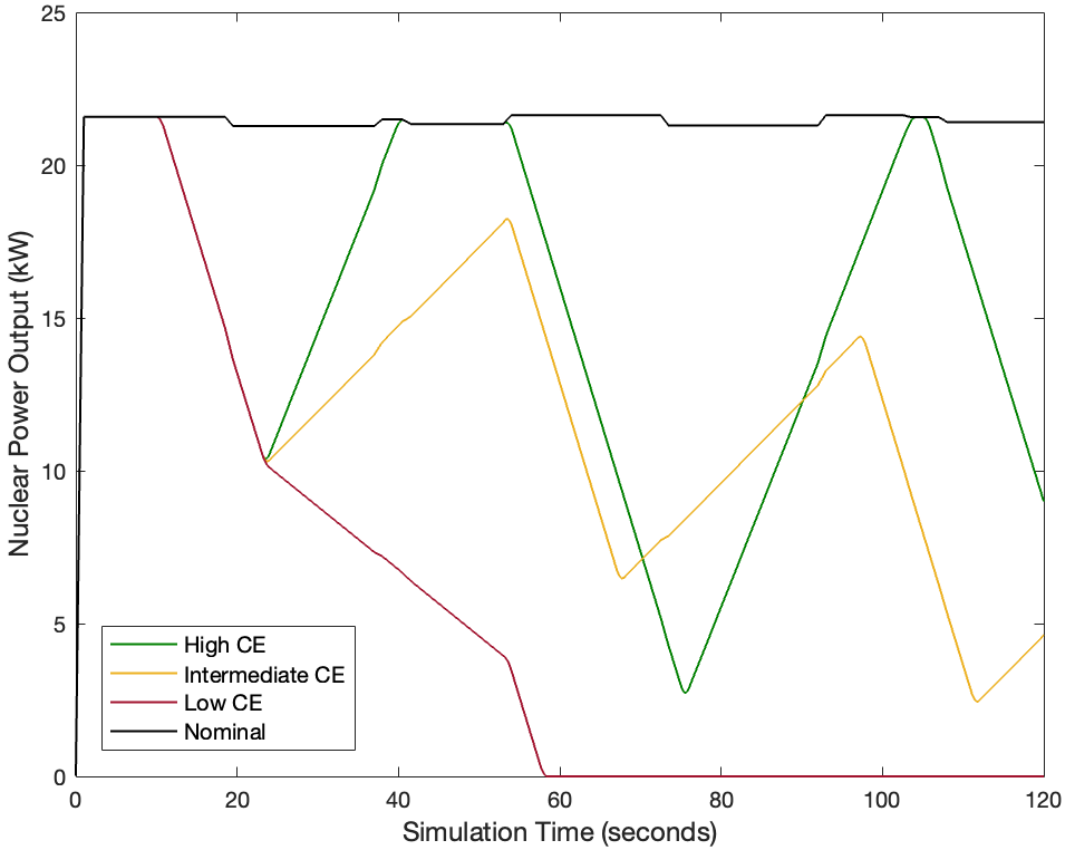

Figure 56. Nuclear Power Output for High, Low, and Intermediate Control Effectiveness Implementations of SC800 in Disruption Scenario 3

Figure 56 shows that dust accumulates on the nuclear radiator panels starting at 10 seconds and continues for the entire simulation. The rate of cleaning for the high control effectiveness safety control is greater than the accumulation rate; therefore, nuclear power generation returns to nominal levels each time the agent completes the cleaning. Conversely, for the intermediate control effectiveness case, the repair rate is not sufficient to enable full recovery; however, it is sufficient to keep the nuclear power output above 0 kW for the entire simulation. For this safety control, only the low control effectiveness case experiences full coverage of the nuclear radiator panel at 58 seconds and remains at 0 kW power output for the rest of the simulation.

5.4.2 Use Performance Curves to Assess Resilience

The three resilience metrics identified in Chapter 4 can be used to quantify habitat resilience for three safety controls activated in this disruption scenario. As previously mentioned, with the current habitat design in the MCVT, we are unable to recover temperature and pressure after a breach in the structure, even after the hole is repaired by an agent. Therefore, we can conclude that the system is not resilient to a breach event. Therefore, we do not map our resilience metrics to the temperature and pressure data here. However, our system is resilient to the energy storage damage and dust accumulation on solar PV arrays and nuclear radiator panels. The agent activities enable recovery in the identified habitat performance metrics, and therefore we map our resilience metrics here for these three cases. Since there are multiple agent interventions for some controls, we will evaluate resilience for the first intervention time period of each safety control. The resilience metric values computed for each safety control and control effectiveness case are shown in Tables 29–31.

Table 29. Resilience Metric Values for SC355 in Disruption Scenario 3

Safety Control: Ability to repair energy storage units (SC355)			
	Low Control Effectiveness	**Intermediate Control Effectiveness**	**High Control Effectiveness**
$Re_1 = \frac{1}{t_4 - t_1}$	0.03	N/A	0.03
$Re_2 = \frac{P(t_4)}{P(t_1)}$	6.364	N/A	10
$Re_5 = \frac{P(t_4) - min(P)}{P(t_1) - min(P)}$	0.596	N/A	1

Table 30. Resilience Metric Values for SC798 in Disruption Scenario 3

Safety Control: Ability to remove dust from solar PV arrays (SC798)			
	Low Control Effectiveness	**Intermediate Control Effectiveness**	**High Control Effectiveness**
$Re_1 = \frac{1}{t_4 - t_1}$	0	0	0
$Re_2 = \frac{P(t_4)}{P(t_1)}$	0.545	0.858	1
$Re_5 = \frac{P(t_4) - min(P)}{P(t_1) - min(P)}$	0	0	0

Table 31. Resilience Metric Values for SC800 in Disruption Scenario 3

Safety Control: Ability to remove dust from nuclear radiator panels (SC800)			
	Low Control Effectiveness	**Intermediate Control Effectiveness**	**High Control Effectiveness**
$Re_1 = \frac{1}{t_4 - t_1}$	0	0.023	0.033
$Re_2 = \frac{P(t_4)}{P(t_1)}$	0.176	0.843	0.995
$Re_5 = \frac{P(t_4) - min(P)}{P(t_1) - min(P)}$	0	0.700	0.991

5.5 Disruption Scenario 4: Fire Originating Near the Power Storage and Distribution Equipment

Here we use the disruption scenario of an intensity level 5 fire originating at Location 5 in the interior environment (Figure 57).

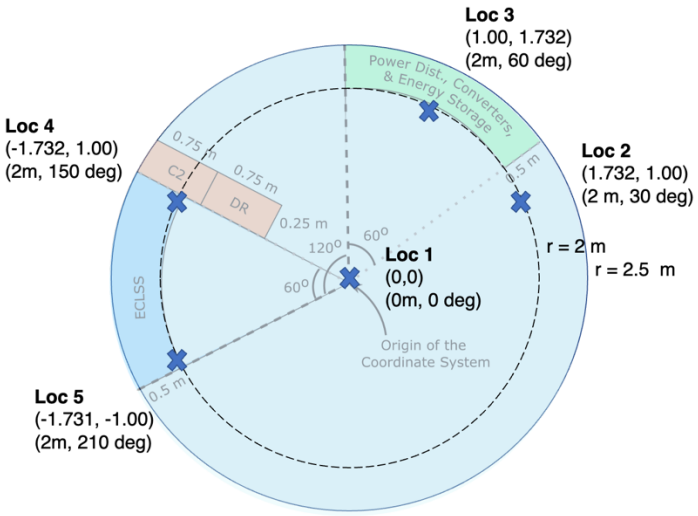

Figure 57. Fire Starting Locations in the Interior Environment of the MCVT

Figure 58 shows the schematic for the hazardous states and safety controls for this disruption scenario.

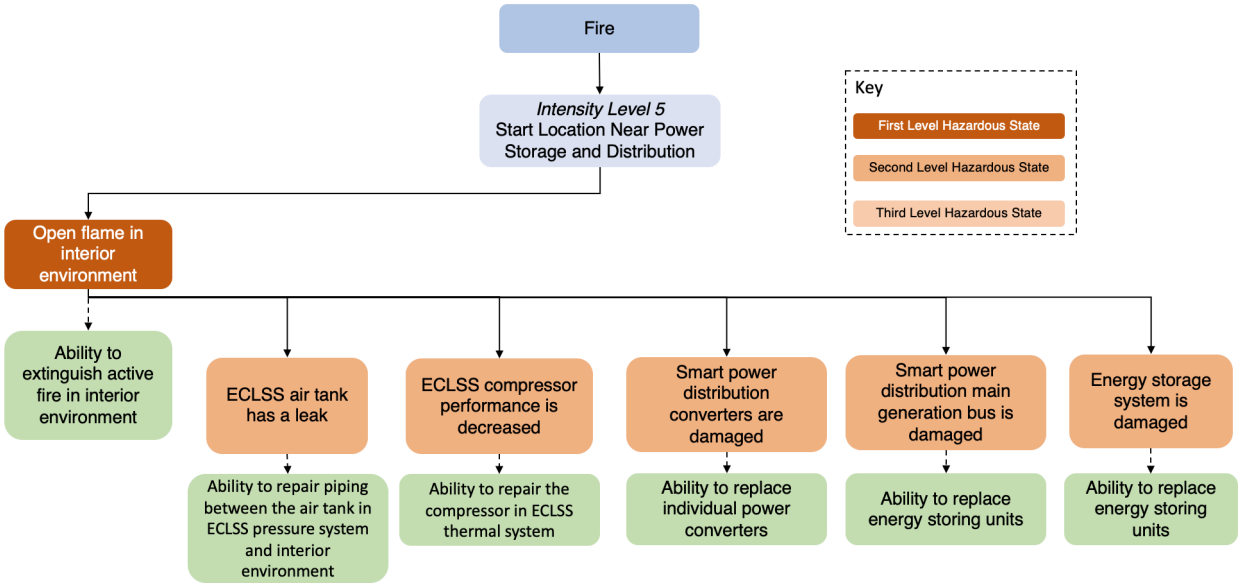

Figure 58. Schematic for Disruption Scenario 4: Fire Originating Near the Power Storage and Distribution Equipment

This disruption scenario includes six safety controls that have two or more implementation strategies with varying control effectiveness values. Running the MCVT for this scenario allowed

for the evaluation of one of these six safety controls: *Ability to extinguish active fire in interior environment*. This was the only safety control evaluated in this scenario because the five secondary hazardous states do not occur for this scenario. In particular, the three hazardous states for the energy storage and distribution systems have been suppressed due to issues with the damage level propagation in this system for this fire disruption. In addition, the ECLSS hazardous states simply are not present as a result of the fire, and therefore ECLSS does not require recovery activities. In the next sections we present the MCVT performance data gathered for the evaluation of the one safety control *Ability to extinguish active fire in interior environment*. Table 32 shows the organization of this control into high, low, and intermediate categories.

Table 32. Categorization of High, Low, and Intermediate Safety Control Implementation Strategies for Disruption Scenario 4

Safety Control	**High control effectiveness implementation strategy**	**Intermediate control effectiveness implementation strategy**	**Low control effectiveness implementation strategy**
Ability to extinguish active fire in interior environment (SC797)	Human agent uses a fire extinguisher to put out fire	Human agent uses fire blanket to put out fire	Robot agent uses fire blanket to put out fire

5.5.1 Obtain Performance Metrics as a Function of Time

As previously mentioned, the MCVT supports simulations on the order of minutes due to computational limitations. For all simulations shown here, the MCVT is run for a simulation time of 120 seconds. Therefore, nominal repair rates were identified for each safety control that would allow for a full recovery to maximum achievable performance at least once in the 120 second simulation. Then, the implementation repair rates are obtained by multiplying the nominal rate by the corresponding values of P_{design} and $P_{implementation}$ for each implementation strategy. These are the repair rates entered by the user for each MCVT simulation. The nominal repair rates and implementation repair rates for this disruption scenario are shown in Table 33, followed by the results of activating the safety control individually.

Table 33. Safety Control Repair Rate Inputs for Agent Model in Disruption Scenario 4

Safety control	**Nominal Repair Rate**	**Implementation Strategy**	P_{design}	$P_{implementation}$	**Implementation Repair Rate**
Ability to extinguish active fire in interior environment (**SC797**)	$0.05\ m/s$	Human agent uses a fire extinguisher to put out fire	0.8	0.9	$0.036\ m/s$
		Human agent uses fire blanket to put out fire	0.6	0.8	$0.024\ m/s$
		Robot agent uses fire blanket to put out fire	0.6	0.5	$0.015\ m/s$

The first metric that enables comparison of the implementation strategies for SC797 is the fire radius. Figure 59 shows the fire radius for the high, intermediate, and low control effectiveness cases. As seen in these figures, the higher control effectiveness strategy allows for a faster reduction in the fire radius, while the lower control effectiveness strategies have a slower reduction in the fire radius. These metrics, although interesting to observe for this safety control, do not enable the evaluation of control effectiveness and resilience at the habitat level. Therefore, we must also consider how the interior environment temperature and pressure respond to the increase and decrease in the fire radius; specifically, how quickly temperature and pressure return to normal after the fire is extinguished.

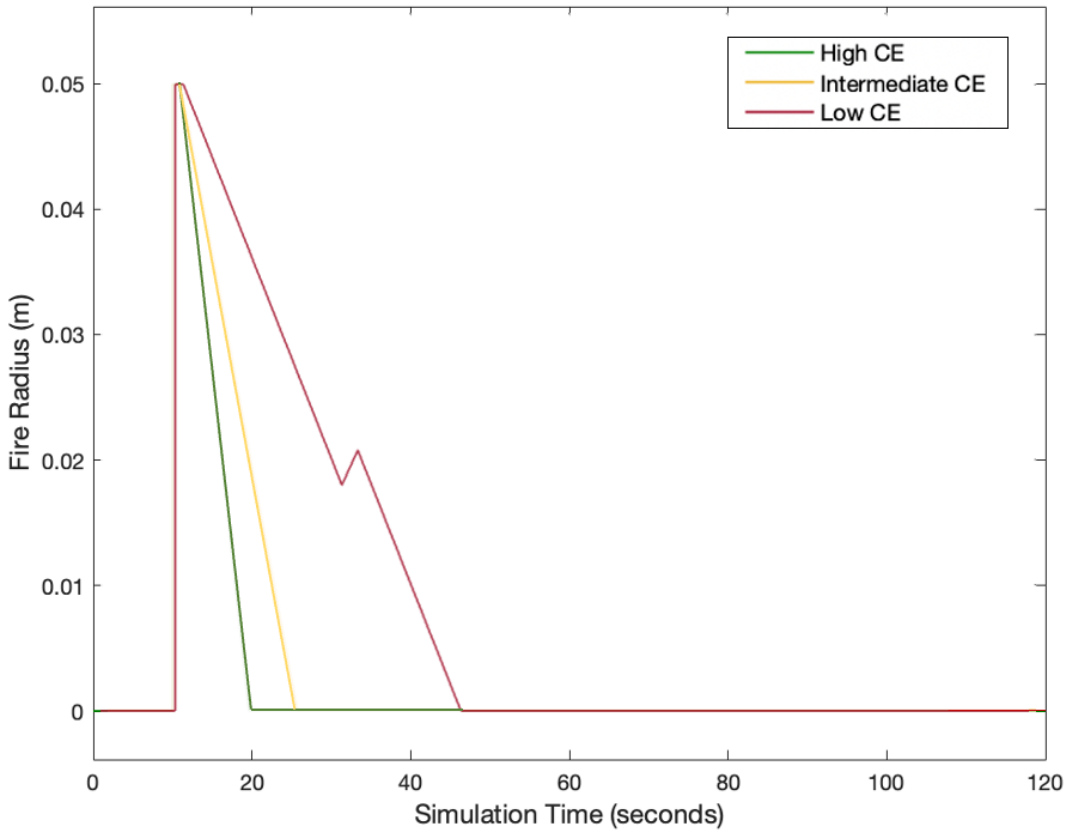

Figure 59. Fire Radius for High, Intermediate, and Low Control Effectiveness Implementation of SC797 in Disruption Scenario 4

Figures 60 and 61 present the recoveries in interior environment temperature and pressure after the activation of SC797. Both figures show that the high control effectiveness case allows for quickest recoveries in temperature and pressure, with the low control effectiveness case causing the temperature and pressure to rise for a much longer time period, as the fire is extinguished much more slowly. What's interesting in both Figures 60 and 61 is that no control effectiveness case shows a return to the nominal values shown in black. The temperature of the interior environment in all three cases restabilizes at values above the nominal temperature value, with each case stabilizing at a different temperature. Conversely, the pressure of the interior environment in all three cases restabilizes at identical levels slightly below the nominal pressure value. The cause of this behavior is not clear from the data gathered here, but the results do not suggest any major issues with the interior environment model, as stabilization of temperature and pressure after the fire most likely take a longer time than the 120 second simulation time frame.

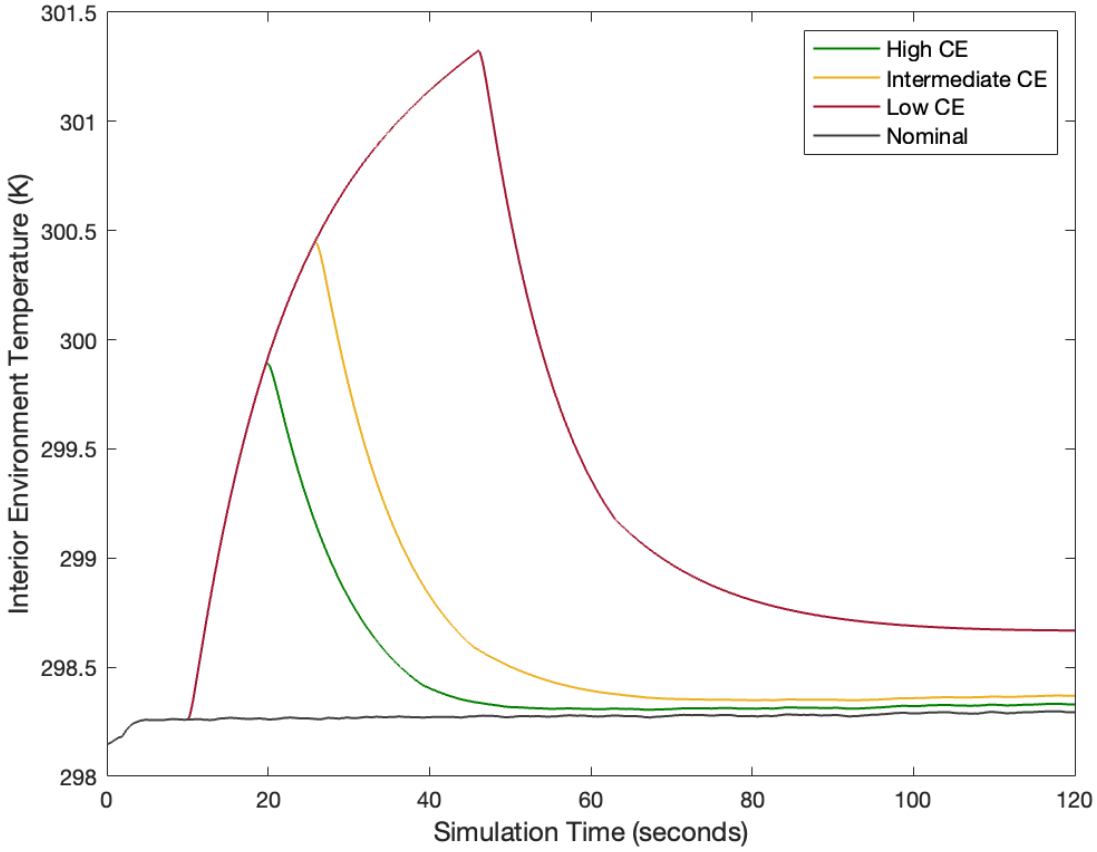

Figure 60. Interior Environment Temperature for High, Low, and Intermediate Control Effectiveness Implementations of SC797 in Disruption Scenario 4

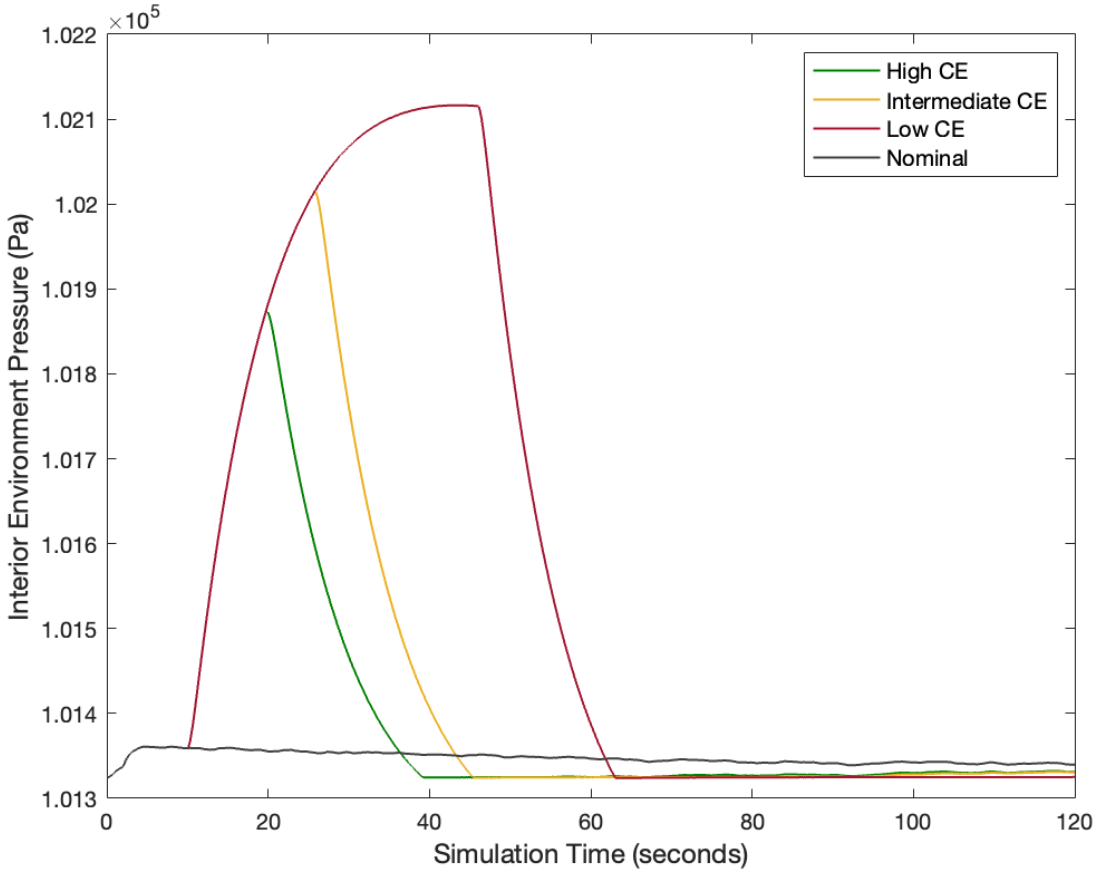

Figure 61. Interior Environment Pressure for High, Low, and Intermediate Control Effectiveness Implementations of SC797 in Disruption Scenario 4

5.5.2 Use Performance Curves to Assess Resilience

The three resilience metrics identified in **Chapter 4** can be used to quantify habitat resilience for the safety control activated in this disruption scenario. However, to do so requires a reformulation of the temperature and pressure performance metrics. This is because the performance curves in this case are inverted, and higher values for temperature and pressure indicate disrupted performance, while lower values indicate recovered performance. Therefore, to be consistent with the previously discussed performance curves and resilience metric values, we will use the following equations to define the performance metrics for temperature and pressure:

$$T_{metric} = \frac{1}{abs(T_{actual} - T_{desired})} \quad (16)$$

$$P_{metric} = \frac{1}{abs(P_{actual} - P_{desired})} \quad (17)$$

In these equations, $T_{desired} = 298.15\ K$ and $P_{desired} = 101325\ Pa$. These are the setpoints in the interior environment temperature and pressure. Using these equations, the resilience metric values were computed for each control effectiveness case for SC797. The results are shown in Tables 34 and 35.

Table 34. Resilience Metric Values for SC797 in Disruption Scenario 4 (Temperature Metric)

Safety Control: Ability to extinguish active fire in interior environment (SC797)			
	Low Control Effectiveness	**Intermediate Control Effectiveness**	**High Control Effectiveness**
$Re_1 = \frac{1}{t_4 - t_1}$	0.011	0.018	0.025
$Re_2 = \frac{P(t_4)}{P(t_1)}$	3.42	8.62	10.56
$Re_5 = \frac{P(t_4) - min(P)}{P(t_1) - min(P)}$	0	0	0

Table 35. Resilience Metric Values for SC797 in Disruption Scenario 4 (Pressure Metric)

Safety Control: Ability to extinguish active fire in interior environment (SC797)			
	Low Control Effectiveness	**Intermediate Control Effectiveness**	**High Control Effectiveness**
$Re_1 = \frac{1}{t_4 - t_1}$	0.019	0.028	0.033
$Re_2 = \frac{P(t_4)}{P(t_1)}$	34	34	34
$Re_5 = \frac{P(t_4) - min(P)}{P(t_1) - min(P)}$	0	0	0

6. CONCLUSIONS

6.1 Summary

In Chapter 1, we introduced the concept of resilience and the motivation of this thesis in the context of the Resilient Extra-Terrestrial Habitats Institute (RETHi). We outlined limitations in existing risk analysis and accident modeling techniques and the need for a new approach that will support the development of tools and technology to establish resilient deep space habitats.

In Chapter 2 we introduced the control-theoretic approach to risk analysis. We described how the approach is derived from the idea that safety is as a control problem. We defined the five steps in the control-theoretic approach, which aim to mitigate risk and keep the system operating in a region of safe behavior. We also introduced how our five-step approach maps to the system safety process and contributes to a larger design trade-off for making decisions. We then presented the state and trigger model used to depict the state-based model of our system to visualize sets of disruptions, hazardous states, accident states, safe states, and safety controls. We ended the chapter by summarizing the work completed by Purdue alumnus, Robert Kitching, for the first three steps in our control-theoretic approach.

In Chapter 3 we presented the development of a modified definition of control effectiveness based on the framework completed by Robert Kitching. We outlined the process of identifying implementation strategies for known safety controls, as well as generic safety control flaws based on the control flaw classification scheme developed by Leveson (2004). The definition of control effectiveness presented by Robert in his Master's thesis (Kitching, 2020) was then summarized, followed by the modifications made in this thesis work to eliminate redundancy and incorporate feedback from NASA reviewers during the 2021 RETHi annual review. We outlined a more repeatable procedure of evaluating control effectiveness through answering a set of guiding questions that relate to each value in control effectiveness and its related control flaws. The chapter ends with an application example in which we evaluate control effectiveness for a set of safety controls in a disruption and hazardous state scenario.

In Chapter 4 we presented the development of a control effectiveness validation plan which fulfills the fourth step in our control-theoretic approach: *Safety Control Assessment.* We first showed the four-step validation cycle, which relies on the MCVT for completion. Each step in the validation procedure was then discussed at length. For Step 1, we presented the development of disruption scenarios for MCVT study, as well as the results of control effectiveness evaluation for safety controls modeled in all disruption scenarios. For Step 2, we identified the relevant habitat/system performance metrics needed to evaluate habitat response due to disruption and safety control influence. This step involves the development of control process models to depict each safety control and its ability to constrain system behavior. The performance metrics and functional requirements identified for each control process were further clarified through the help of a Design Structure Matrix, and then allocated through the help of a habitat architecture diagram. For Step 3, we described the MCVT and presented the physical layout of subsystems in the simulation. The use of this platform in evaluating control effectiveness is outlined, with particular emphasis on the limitations of the platform in completing the validation procedure. Finally, for Step 4, we described the quantitative resilience assessment and identification of resilience metrics. Seven resilience metrics are presented from a literature review, followed by an analysis and selection of the most appropriate metrics for use in quantifying resilience for the MCVT.

In Chapter 5 we presented the current status in evaluating control effectiveness using MCVT v6. We describe the current capabilities and limitations of MCVT v6 and introduce the need for additional testing with future iterations of the MCVT. The rest of the chapter presents four disruption scenarios that were used to evaluate 5 safety controls which were activated individually in MCVT v6. These results for each scenario demonstrate the completion of Steps 3 and 4 in the control effectiveness validation plan.

6.2 Key Findings

This thesis presented the progress made in validating the control effectiveness metric, and therefore helped in validating our control-theoretic approach to resilient space habitat design. Robert Kitching, in his MSAA thesis, completed the first three steps in our control-theoretic approach and demonstrated that we can approach risk management from a controls perspective. The approach presented is grounded in system safety engineering and maps to the traditional risk

management framework, while incorporating new techniques in the identification and assessment of hazards and mitigation methods.

In this thesis, we further developed and improved the fourth step in our approach, *Step 4: Assessing Safety Controls*. Following the initial framework developed by Kitching, we presented a refined control effectiveness definition to assess how well safety controls address their target disruption or hazardous state. The modified definition relates the effectiveness of our safety controls to their susceptibility to known control flaws classified by Leveson (2004). We develop implementation strategies for our safety controls to investigate how a safety control achieves its control goal. For each implementation strategy, we consider certain probabilities of competent design, availability, and perfect implementation. In addition, we consider the response margin of each strategy, which incorporates both the time it may take to implement the control (activation time) and the time it takes to see performance degradation due to the corresponding hazardous state (time to effect). To aid in the evaluation of each control effectiveness value, we presented a series of guiding questions for each probability and response margin, which map to one or more of the known control flaws. Answering these questions for each control and corresponding implementation strategies can help standardize how a designer thinks about each control by establishing a consistent framework on which to compare each safety control.

To evaluate control effectiveness and determine whether it is an appropriate metric of selecting safety controls that will lead to desired resilience, we proposed a procedure for validating the definition presented in this thesis using the MCVT. We demonstrated how MCVT development proceeds in parallel with the four steps of the validation cycle to support the modeling of complex disruption scenarios and safety controls with different control effectiveness values. Control effectiveness validation is an iterative process which repeats based on the correlation between control effectiveness and resilience, as well as the expansion of the MCVT to explore new disruption scenarios. The validation procedure was developed to determine whether safety controls with established control effectiveness values lead to desired habitat resilience. Specifically, high control effectiveness safety controls should create high resilience architectures.

The thesis concluded with a discussion on the current status of evaluating control effectiveness using MCVT v6. Using four disruption scenarios that support the modeling of 15 user-controlled safety controls, we were able to evaluate five safety controls which are activated individually in their corresponding disruption scenario(s). The MCVT performance curves obtained for each disruption scenario demonstrated that when activated individually, high control effectiveness safety controls result in more resilient performance curves. This was not only observed qualitatively by characterizing the performance curve shapes, but also quantitatively by mapping to three resilience metrics. As presented in **Chapter 5**, for some safety controls, one or two resilience metrics were insufficient to quantitatively distinguish between safety control implementation strategies. For example, when computing Re_1, we often obtained values of zero for all control effectiveness cases. This primarily occurred when no recovery performance was obtained and the recovery time frame (t_4) went to infinity. In addition, for Re_2, we observed a case where this metric went to infinity. This was a result of the degraded performance equaling zero in the denominator. We also obtained values of one for this metric in several implementations of the same control. This indicated the recovered performance equaled the disrupted performance, and no positive recovery slope was ever achieved. Values of one were also obtained for Re_5, and this also occurred when the recovered performance equaled the disrupted performance, with identical minimum performance values. Re_5 also resulted in values of zero if the recovered performance equaled the minimum performance. These difficulties in computing the resilience metric values can be attributed to the deviation of the MCVT resilience curves from the "standard notional" curve shown in Figure 37. Although not every metric was sufficient for distinguishing the resilience achieved for each implementation strategy, for all safety controls evaluated using the MCVT v6, at least one metric produced resilience metric values that were proportional to control effectiveness. Specifically, we consistently observed that the resilience metric values increased with increasing control effectiveness, which is the correlation we expected. This means control effectiveness is appropriately defined to select individual safety controls that will lead to desired habitat resilience, demonstrating the validity of our control effectiveness approach.

6.3 Limitations and Potential Improvement

The first limitation of our control-theoretic process is that it still involves a certain amount of subjectivity in the evaluation of control effectiveness values for safety controls. When assessing

the probabilities of competent design, availability, and perfect implementation, we proposed a set of guiding questions that should be answered for all safety controls when assessing these values of control effectiveness. Although these guiding questions present a standardized framework to think about the controls in the context of particular control flaws, there is still subjectivity in the final values assigned to each probability. These values will be based on engineering judgement and experience; however, mission designers will have a consistent means of comparing control effectiveness values based on the answers to these guiding questions.

In addition to the assessment of control effectiveness, the primary limitation in this process is the dependency on the MCVT for the validation of control effectiveness. There are several limitations in the MCVT simulation that affect what control effectiveness values can be evaluated. First, as previously mentioned, the MCVT can only run simulations on the order of minutes. Therefore, we cannot evaluate values of $P_{available}$ between 0 and 1, as this requires long-term simulations on the order of weeks to months. To evaluate the values of $P_{available}$ and the affect this dimension has on resilience, we can make use of the RETHi's Control-Oriented Computational Dynamic Modeling (CDCM) platform. This is a lower fidelity simulation environment that can model the habitat for long time periods and support the probabilistic activation of safety controls.

In MCVT v6, users cannot currently set the activation time of the safety controls, which means the start times of all implementation strategies for a single control are identical. In reality, the activation time is set by the value of $t_{sc,affect}$ which is known for each control based on the known layout of the MCVT, estimated agent speeds, and required equipment/preparation time. For improvement of the control effectiveness validation results presented in **Chapter 5**, future iterations of the MCVT will include the ability for users to set activation time of safety controls. The validation plan can then be re-run with selections of high, low, and intermediate controls based on three of the control effectiveness values, P_{design}, $P_{implementation}$, and $M_{response}$, rather than just P_{design} and $P_{implementation}$. Doing so will provide more accurate results that allow us to conclude whether high control effectiveness safety controls lead to high resilience architectures.

The final limitation in MCVT v6 is that users cannot currently simultaneously activate more than one safety control and obtain meaningful results. The heuristic scheduler for the agent model

presents unexpected behavior when multiple controls are activated, which affects the completion of more than one control in a single simulation. Often, the agent repeats one control activity throughout the entire simulation, without moving onto the next activity. This behavior can be somewhat controlled by reducing the time an agent spends on each activity; however, doing so produces performance curves that are not meaningful for evaluating resilience with the entire set of controls. Therefore, modeling sets of safety controls with all high, all low, or mixed control effectiveness is not possible for MCVT v6, and a formal review and improvement of the heuristic scheduler is required to enable the activation of multiple safety controls in response to simultaneous hazardous states. This would allow for the identification of potential weaknesses in the safety structure that diminish resilience and the opportunity to enhance component, subsystem, or system design aspects that prove less resilient to disruptions than others. An additional issue with the heuristic scheduler identified in **Chapter 5** is the anomalous activation of safety controls that are deactivated from the user side. A thorough review of the damage level propagation and activation of agent interventions in MCVT v6 is necessary to eliminate those bugs that were identified in this first cycle of evaluating control effectiveness.

REFERENCES

Ayyub, Bilal M. “Systems Resilience for Multihazard Environments: Definition, Metrics, and Valuation for Decision Making.” *Risk analysis* 34.2 (2014): 340–355. Web.

Bahr, N. J. (2016). *System Safety Engineering and Risk Management.* Boca Raton: CRC Press, Taylor & Francis Group.

Bruneau, Michel et al. “A Framework to Quantitatively Assess and Enhance the Seismic Resilience of Communities.” *Earthquake spectra* 19.4 (2003): 733–752. Web.

Cheng, Congcong et al. “Improved Integrated Metric for Quantitative Assessment of Resilience.” *Advances in mechanical engineering* 12.2 (2020): 168781402090606–. Web.

Choi, C. Q. (2014, September 27). *Moonwalking astronauts can move surprisingly fast*. Space.com. Retrieved May 29, 2022, from https://www.space.com/27285-astronauts-moon-walking-speed-spacesuits.html

Das, Laya et al. “Measuring Smart Grid Resilience: Methods, Challenges and Opportunities.” *Renewable & sustainable energy reviews* 130 (2020): 109918–. Web.

Dyke, S. et al. (2018). RETHi Proposal Technical Narrative.

Dyke, S.J., Marais, K., Bilionis, I., Werfel, J. (2022). RETH institute Annual Report Appendix A Modular Coupled Virtual Testbed (Version 6.0), NASA Sharepoint.

Gullo, & Dixon, J. (2018). *Design for safety / edited by Louis J. Gullo, Jack Dixon.* (Gullo & J. Dixon, Eds.). Wiley.

Henry, Devanandham, and Jose Emmanuel Ramirez-Marquez. “Generic Metrics and Quantitative Approaches for System Resilience as a Function of Time.” *Reliability engineering & system safety* 99 (2012): 114–122. Web.

Kitching, R. (2020). *A Control-Theoretic Approach to the Resilient Design of Extra-Terrestrial Habitats* (Master’s thesis, Purdue University, West Lafayette, United States).

Leveson, N. (2004). A New Accident Model for Engineering Safer Systems. *Safety Science*, Vol. 42, No. 4, pp. 237-270.

Leveson. (2012). Engineering a Safer World: Systems Thinking Applied to Safety. The MIT Press. https://doi.org/10.7551.mitpress/8179.001.0001

NASA. (n.d.). *The Rover's wheels*. Mars Exploration Rovers. Retrieved May 29, 2022, from https://mars.nasa.gov/mer/mission/rover/wheels-and-legs/

Perrow, Charles. *Normal Accidents: Living with High-Risk Technologies*. Princeton University Press, 1999.

Rose, A., & Liao, S. (2005). Modeling regional economic resilience to disasters: A computable general equilibrium analysis of water service disruptions. *Journal of Regional Science*, 45(1): 75-112.

Uday, P., & Marais, K. (2015). Designing resilient systems-of-systems: A survey of metrics, methods, and challenges. *Systems Engineering*, 18 (5), 491-510.

Yarveisy, Rioshar, Chuan Gao, and Faisal Khan. “A Simple yet Robust Resilience Assessment Metrics.” *Reliability engineering & system safety* 197 (2020): 106810–. Web.