Building resilience by connecting the dots

Christopher J. Lowe* and Malcolm Macdonald.†

Scottish Centre of Excellence in Satellite Applications, Level 4, Technology and Innovation Centre, University of Strathclyde, 99 George Street, Glasgow, G1 1RD

Abstract

Satellites typically operate in isolation from their orbiting neighbours, leaving them susceptible to even the most minor of failures. Loss of a payload, radio or critical supporting sub-system could render the platform useless, an unfavourable situation for mission stakeholders. There is however a partial solution through the addition of inter-satellite networking, which offers not only value in terms of general performance, but added resilience to failure in the form of degraded operations. While a traditional platform exhibits two fundamental states: *operational* (which includes the collection and dissemination of data) and *failed*, a network-capable platform (i.e. one with an inter-satellite communication capability) exhibits six states, each reached through a unique combination of sub-system failures. The result of this added resilience is a reduction in the likelihood of the satellite reaching a fully-failed state, at the burden of higher financial cost and complexity.

* Research Fellow, christopher.lowe@strath.ac.uk
† Director, malcolm.macdonald.102@strath.ac.uk

1. Introduction

Failures can, and do, occur on space systems, sometimes with serious consequences. Be it at a component, sub-system or platform-level, anomalies in space are often difficult to resolve, especially if they occur on hardware. Here, the impact of including an inter-satellite networking (ISN) capability, and how it can affect a platform's resilience in the face of failures, is investigated.

ISN can be defined as the ability to communicate with orbital neighbours, such that data can be transferred via inter-satellite communication links [1]. While the concept of ISN is not new, the investigation of its effect on a system's resilience (i.e. its ability to withstand performance degradation as a result of failure) has not received much attention. In what follows, methods for the assessment of failure effects are introduced and a model that captures the addition of ISN, and its impact on system resilience, is presented. Discrete-time Markov chains (Section 2) are used to describe the transition between states, where a platform containing ISN features additional degraded states not present on non-ISL capable platforms. These additional states enable the platform to reside in a functional state for longer, thus adding value to a mission despite partial on-board failures (Section 3).

2. Background

A point design, i.e. one that is rigid in its characteristics and operation, may be optimal in the nominal environment for which it was developed, but may be left exposed in the real world in which uncertainties exist. Consideration of uncertainties as part of the design process, in particular from the point of view of failure (reliability engineering), has been practiced since the early 1900s [2]. However, studies into the effects resulting from arbitrary uncertain events taking place has only recently come to the fore [3]–[5]. For clarity, uncertainty can be separated into two forms; i) aleatoric – that which cannot be known *a priori*, such as minor fluctuations in atmospheric conditions or the time of future solar activity, and ii) epistemic – that which occurs due to lack of knowledge, such as the centre of mass of an instrument or the cost of a launch. Here, while both are acknowledged as important and valid, only the former (aleatoric) will be considered. Irrespective of the type of aleatoric uncertainty (e.g. failure, market change, environmental change, technology obsolescence), the over-arching effect is generally a drop in system utility (performance), which we hope to measure in order to better understand utility over the system's lifetime. The change considered here is failure of on board systems, resulting in full or partial platform degradation.

Markov chains offer a convenient way of modelling the stochastic process of component/system failure and allow evaluation of expected state evolution over time. A Markov process is a stochastic process defined by a set of random variables $\{X(t), t \in T\}$, where each $X(t)$ is a random variable, or "state", defined on some probability space, at time t. T is the time horizon and can be discrete in nature, where $T = \{0, 1, 2, \dots\}$ forms a discrete time Markov chain (DTMC), or continuous, where $T = \{t: 0 \leq t < \infty\}$ forms a continuous time Markov chain (CTMC). For the purposes of this work, the focus shall be on DTMCs due their applicability to systems for which the probability of

transitioning between states may be time-varying [6]. The Markov property is analogous to the memoryless property, such that transition from one state to another depends only on the current state, and not on states in which the system resided during some previous time. Formally, this is

$$\text{Prob}\{X_{n+1} = x_{n+1} | X_n = x_n, X_{n-1} = x_{n-1}, \dots, X_0 = x_0\} = \text{Prob}\{X_{n+1} = x_{n+1} | X_n = x_n\}, \quad 2.1$$

where x_i is the state of the system at time i. The above can be simplified to

$$p_{ij}(n) = \text{Prob}\{X_{n+1} = j | X_n = i\}. \quad 2.2$$

Given a discrete and finite set of states, the matrix $P(n)$, in which $p_{ij}(n)$ is the entry in the i^{th} row and j^{th} column, is called the transition probability matrix, and contains the probability of transitioning between any two states at time n. $P(n)$ is written as

$$P(n) = \begin{bmatrix} p_{11}(n) & p_{12}(n) & \cdots & p_{1K}(n) \\ p_{21}(n) & p_{22}(n) & \cdots & p_{2K}(n) \\ \vdots & \vdots & \ddots & \vdots \\ p_{K1}(n) & p_{K2}(n) & \cdots & p_{KK}(n) \end{bmatrix}, \quad 2.3$$

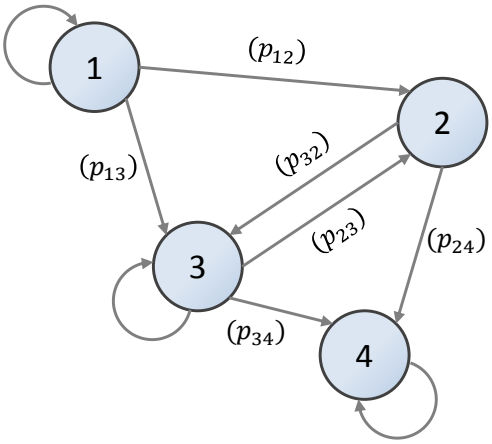
where K is the number of discrete states and thus $P(n) = \mathbb{R}^{K \times K}$. It follows that

$$\sum_{all\ j} p_{ij}(n) = 1, \forall i, \quad 2.4$$

where $i = j$ is a possible scenario and simply represents the absence of any transition from the current state. In a (time-) nonhomogeneous DTMC, $P(n)$ may vary with n, which is an important characteristic to capture for systems that suffer from either reliability decay over time, or a higher failure probability at the beginning of life. Analysis into the robustness/survivability achieved from swarm size [7] and system fractionation [8] has been conducted using a homogeneous DTMC, which results in this time-varying phenomena being missed. In [9] and [10] it is shown that this time-varying property is real for space systems and should be incorporated.

For systems that do not undergo servicing or repair, applicable to the majority of satellites, transition from a failed state back to an operational one is not possible, such that the 1st column in $P(n)$ is $[p_{11}(n) \quad 0 \quad 0 \quad 0]^T$ and the final row (fully failed) is $[0 \quad 0 \quad 0 \quad 1]$, which represents an *absorbing* state.

As an example, Figure 1 illustrates a typical Markov chain with four states, in which state 2 always transitions (no arc to itself) and state 4 is absorbing. It is recognised that satellite systems generally feature redundancy on most sub-systems, in particular on those that are mission critical. This is considered throughout this work such that the failure probabilities used are derived from actual flight data, where redundant systems, if present, would have failed.



$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & 0 \\ 0 & 0 & p_{23} & p_{24} \\ 0 & p_{32} & p_{33} & p_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Figure 1 – Example of a DTMC with an absorbing state

In order to identify the expected utility being offered by a system subject to stochastic failures over its lifetime, transient analysis should be conducted, which provides an answer to the question "what is the probability that our system will be in state x after m time steps, given that it is in state y now"?. The Chapman Kolmogorov equations offer such functionality, derived from the fact that the probability for a two-step transition from state i to k is

$$\text{Prob}\{X_{n+2} = k | X_{n+1} = j, X_n = i\} = p_{ij}p_{jk}, \forall j \in \{1, 2, \dots, K\},$$
$$\text{Prob}\{X_{n+2} = k | X_n = i\} = \sum_{all\ j} p_{ij}p_{jk}, \qquad 2.5$$

which is the ik^{th} element in P^2 for a homogenous DTMC, or the ik^{th} element in $P^{(n+2)}(n, n+1, n+2) = P(n+2)P(n+1)P(n)$ in a non-homogeneous DTMC. This can be generalised to obtain the Chapman Kolmogorov equation for non-homogeneous DTMCs, as

$$P^{(m)}(n, n+1, \dots, n+m-1) = P(n)P(n+1) \dots P(n+m-1). \qquad 2.6$$

Finally, given a row vector $\pi^{(n)} = \mathbb{R}^{1\times K}$ describing the probability of being in state $i \in \{1,2, \dots, K\}$ at some time n, the probability of being in state i at time $(m+n)$ is

$$\pi^{(m+n)} = \pi^{(n)}P(n)P(n+1) \dots P(n+m-1). \qquad 2.7$$

3. Failure on Space Systems

As discussed, aleatoric uncertainty is considered in this analysis, in particular with respect to system failure. Complete failure affects over 6% of satellites within their first 7 years and almost 9% of satellites within their first 12 years [9]. It is therefore not something that can be ignored, and both the likelihood of its arrival and its effects on the mission utility, should be understood. Exactly how a particular sub-system fails can be considered arbitrary for the purposes of this analysis, however it is important to understand the likelihood of failure occurring over a particular timeframe. It has been shown, in [9] and [11], that satellite sub-system failure can be approximated to a Weibull distribution, capturing the higher probability of failure at beginning of life that is neglected when using other functions, such as a linear regression. The reliability of a traditional space platform, or in other words its probability of residing in a non-failed state, is shown in Figure 2.

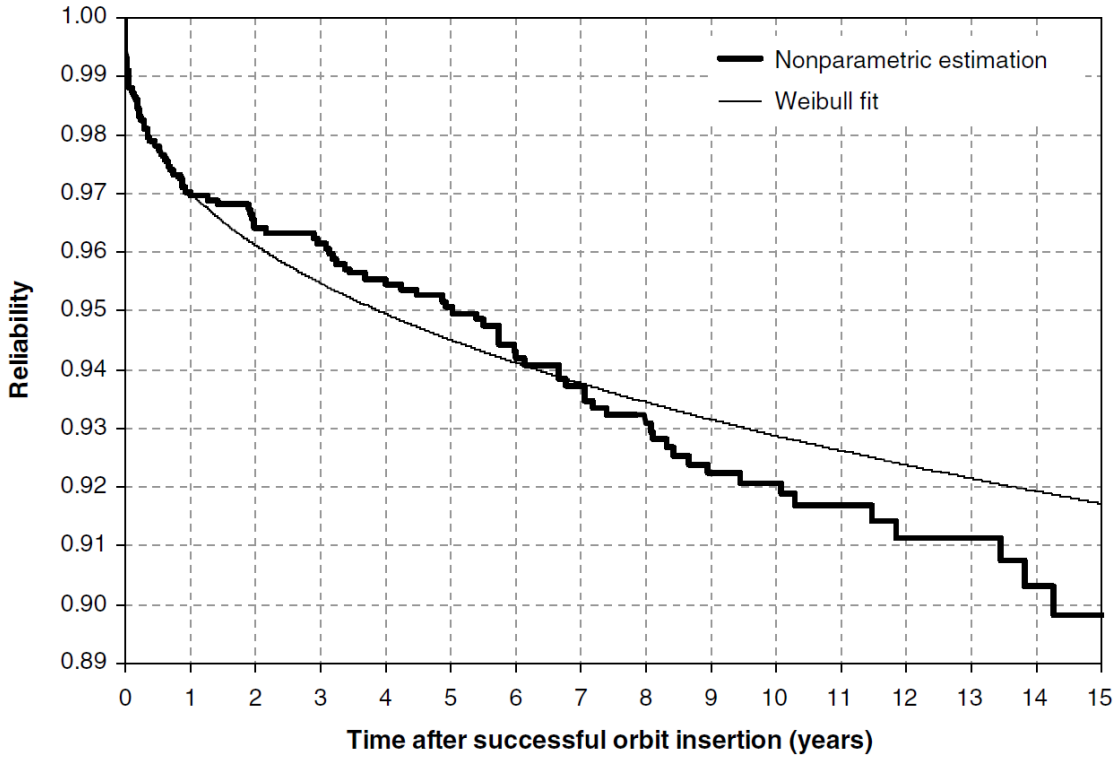


Figure 2 – Weibull distribution correlation to actual satellite reliability data [9]

The operation of a traditional, monolithic satellite can typically be reduced to that of data collection via its payload/s, followed by data delivery to the end user via its communication sub-system/s. Failure of either of these sub-systems, or a critical supporting element (e.g. power system or on-board computer), would result in this functionality not being possible and thus transition into a failed state. In reality, there exist various intermediate, partially-failed states following failure of

some non-critical component*, but for the purposes of this work, and to effectively illustrate the effect of a networking capability, these are omitted. For a satellite with an inter-satellite networking capability, there exists another function in addition to the nominal collect and deliver capabilities. This is its ability to relay information, which not only provides a higher nominal utility via increased functionality, but introduces four degraded operational states that are not available to a non-networked system. This is described in Figure 3, where the assumption is made that additional hardware is required for ISN, thus offering additional resilience†.

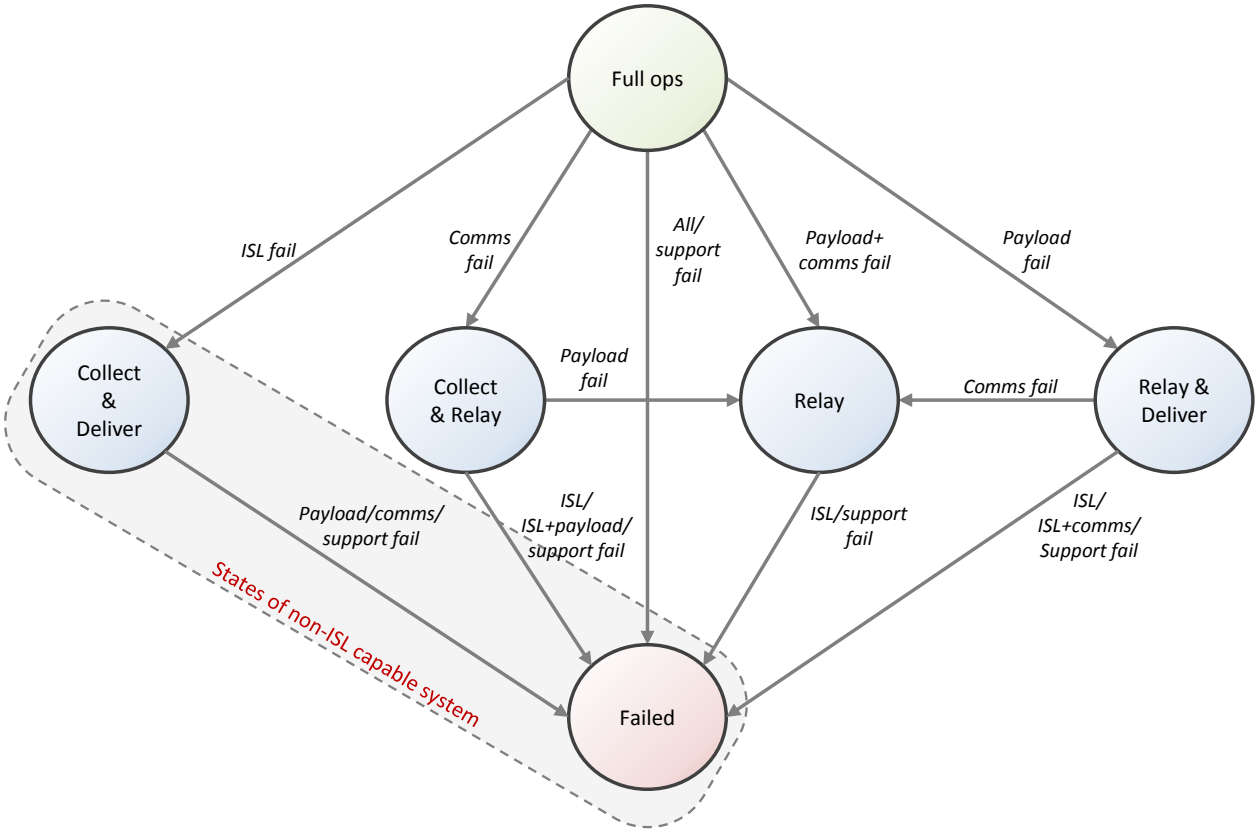


Figure 3 – State transition Markov chain diagram for a networked and non-networked system

Figure 3 shows the state possibilities being considered here, including the partially failed states and the potential transitions to and from each state.

* Consider, for example, failure of a string of solar cells or degraded reaction wheel mobility. The result would likely be a reduction in payload duty-cycle, such that some utility can still be maintained, but at a lower level than if fully operational

† While it is technically feasible that an ISN capability is implemented using the same hardware as used between the satellite and ground station, for the purposes of this work an ISL is assumed to require additional hardware. This is considered a reasonable assumption given the likelihood of differences in link frequency and potential pointing demands.

The information in Figure 3 is summarised in Table 1.

#	State	Transition		Comment
		from	to	
1	Full Ops	None	All	All systems operational
2	Collect & Deliver	1	6	ISL fail (full ops for traditional system)
3	Collect & Relay	1	5, 6	Communications fail
4	Relay & Deliver	1	5, 6	Payload fail
5	Relay	1, 3, 4	6	Payload & communications fail
6	Failed	All	None	1. Critical support system failure, or 2. Combined failure of ISL and either payload or comunications systems

Table 1 – State definition and transition (transitions from/to self not included)

Indeed, all partially- or fully-failed states could transition to a more operational state if a repair/replace service was available. This concern is not addressed in this work. The Markov chain in Figure 3 can also be represented as a probability state-transition matrix of the form

$$P(t_k) = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} & p_{15} & p_{16} \\ 0 & p_{22} & 0 & 0 & 0 & p_{26} \\ 0 & 0 & p_{33} & 0 & P_{35} & p_{36} \\ 0 & 0 & 0 & p_{44} & p_{45} & p_{46} \\ 0 & 0 & 0 & 0 & p_{55} & p_{56} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \qquad 3.1$$

where $P(t_k)$ is the transition matrix at time t_k, and p_{ij} is the probability that the system will transition from state i to state j during time period $t_{k+1} - t_k$. Note that the number of non-zero entries along the row direction indicates the number of states to which a state (defined by row number) can transition, and along the column direction the number of states from which a state (defined by column number) can be transitioned. Owing to the fact that the probability of transitioning between states, in the case of space platforms, is variable with time (time-nonhomogeneous), the discrete-time form of Markov chains is employed, whereby the matrix P may differ over the lifetime. Recall from Section 3, that the probability of being in each state at time t_k, given the likelihood of being in each initial state as defined by the vector $\pi^{(0)}$, is

$$\pi^{(t_k)} = \pi^{(0)} P(0) P(1) \dots P(t_k - 1). \qquad 3.2$$

For illustrative purposes, consider a satellite with failure probabilities for the payload, ISN system, communication system and critical support systems, of 5%, 4%, 3% and 5% respectively, in any

given year. The state probability plot is shown for both a non-networked system (Figure 4) and a networked system (Figure 5).

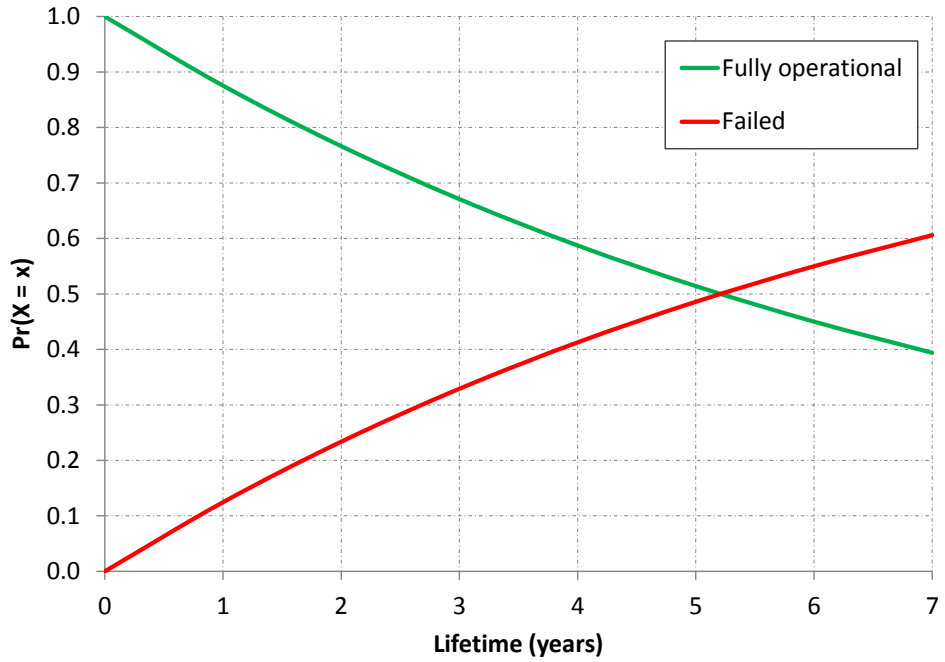


Figure 4 – State probability of a system without networking capability

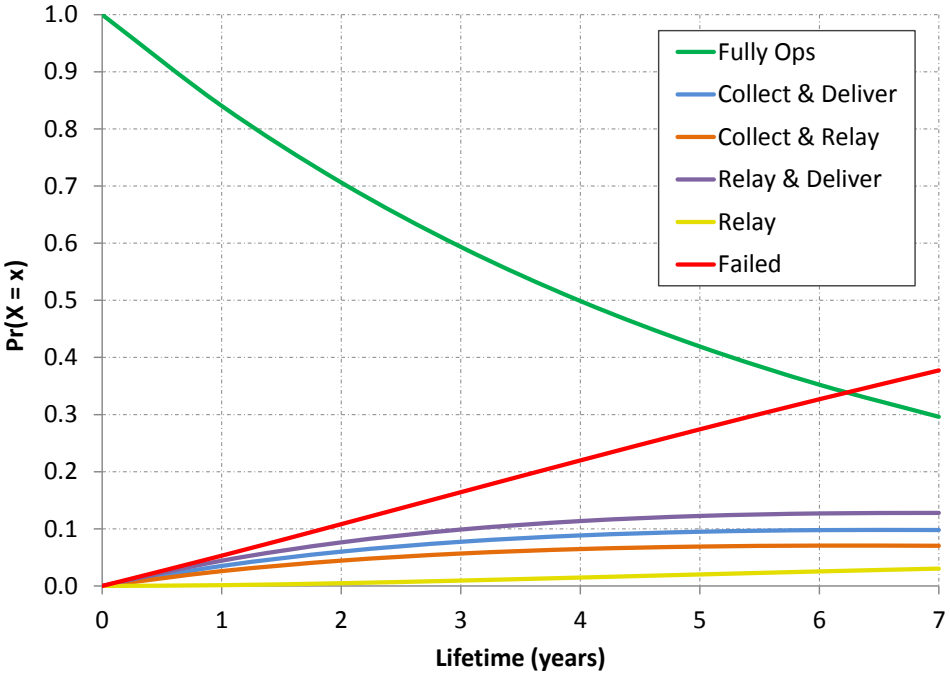


Figure 5 – State probability of a system with a networking capability

What Figure 4 and Figure 5 tell us is that, for this specific set of transition probabilities;

i. A network-capable system is equally likely to remain in either a fully operational state or a collect & deliver state, as a non-networked system is to remain in its fully operational state (which is, by definition, a collect & deliver state).
ii. A network-capable system is less likely to enter into a fully failed state, since there are intermediate, degraded states that are reachable. In this example, after 7 years, the networked system is ~23% less likely to be in a fully failed state, than a non-networked system.

While the failure likelihood of the above figures is perhaps exaggerated for illustrative effect, it is clear that there may be value in incorporating inter-satellite networking, with regards to lifetime utility, but it will come at a cost. Whether this additional financial and complexity burden is worth taking must be considered on a case-by-case basis.

4. Conclusions

The presence of an inter-satellite networking (ISN) capability on a space platform enables on-board functionality to be maintained following failures that would otherwise render a non-ISN capable platform unserviceable. These additional, partially-failed, states are enabled through the introduction of a "relay" functionality, i.e. the ability to transfer data to/from other satellites. While for a traditional, monolithic satellite, failure of the payload or downlink communication system would prevent useful operations taking place, an ISN-capable system would transition into a "relay & deliver" or "collect & relay" state, respectively. Furthermore, following failure of both the payload and communication system, the ISN-capable satellite could still offer a relay-service for other satellites with ISN capabilities.

It is clear that ISN not only offers additional capability, which could add value to a mission during nominal (fully-functional) operations, but it brings a higher likelihood that a platform would remain in a non-failed state during its lifetime.

5. References

[1] C. J. Lowe, "Methodologies for the Analysis of Value from Delay-Tolerant Inter-Satellite Networking," University of Strathclyde, 2017.

[2] J. H. Saleh and K. Marais, "Highlights from the early (and pre-) history of reliability engineering," *Reliab. Eng. Syst. Saf.*, vol. 91, no. 2, pp. 249–256, Feb. 2006.

[3] H. McManus, M. Richards, A. Ross, and D. Hastings, "A Framework for Incorporating 'ilities' in Tradespace Studies," in *AIAA Space*, 2007, pp. 1–14.

[4] J. Saleh, E. Lamassoure, and D. Hastings, "Space systems flexibility provided by on-orbit

servicing: Part 1," *J. Spacecr. …*, vol. 39, no. 4, 2002.

[5] E. Lamassoure, J. Saleh, and D. Hastings, "Space systems flexibility provided by on-orbit servicing: Part 2," *J. Spacecr. …*, vol. 39, no. 4, 2002.

[6] W. Stewart, "Section II - Markov Chains," in *Probability, Markov Chains, Queues, and Simulation: The Mathematical Basis of Performance Modeling*, 1st ed., Princeton University Press, 2009, pp. 191 – 375.

[7] C. D. Jilla, D. W. Miller, and R. J. Sedwick, "Application of Multidisciplinary Design Optimization Techniques to Distributed Satellite Systems," *J. Spacecr. Rockets*, vol. 37, no. 4, pp. 481–490, 2000.

[8] G. F. Dubos and J. H. Saleh, "Comparative cost and utility analysis of monolith and fractionated spacecraft using failure and replacement Markov models," *Acta Astronaut.*, vol. 68, no. 1–2, pp. 172–184, Jan. 2011.

[9] J.-F. Castet and J. H. Saleh, "Satellite Reliability: Statistical Data Analysis and Modeling," *J. Spacecr. Rockets*, vol. 46, no. 5, pp. 1065–1076, Sep. 2009.

[10] J.-F. Castet and J. H. Saleh, "On the concept of survivability, with application to spacecraft and space-based networks," *Reliab. Eng. Syst. Saf.*, vol. 99, pp. 123–138, Mar. 2012.

[11] J.-F. Castet and J. H. Saleh, "Satellite and satellite subsystems reliability: Statistical data analysis and modeling," *Reliab. Eng. Syst. Saf.*, vol. 94, no. 11, pp. 1718–1728, Nov. 2009.