Exploring Design and Policy Options for Orbital Infrastructure Projects

by

Lieutenant Junior Grade Benjamin L. Putbrese, U.S. Navy

B.S. Aerospace Engineering
United States Naval Academy, 2013

Submitted to the Engineering Systems Division in Partial Fulfillment of the Requirements for the Degree of

Master of Science in Technology and Policy

at the

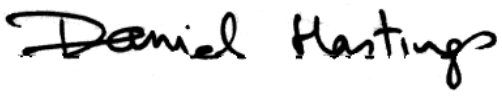
Massachusetts Institute of Technology

June 2015

©2015 Benjamin L. Putbrese. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

Signature of Author: _____
Engineering Systems Division
May 8, 2015



Certified by:_____
Daniel E. Hastings
Cecil and Ida Green Education Professor of Engineering Systems and Aeronautics and Astronautics
Thesis Supervisor

Accepted by: _____
Dava J. Newman
Professor of Aeronautics and Astronautics and Engineering Systems
Director, Technology and Policy Program

This work is sponsored by the United States Department of Defense under Contract FA8721-05-C-0002. Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the United States Government.

Exploring Design and Policy Options for Orbital Infrastructure Projects

by

Lieutenant Junior Grade Benjamin L. Putbrese, U.S. Navy

Submitted to the Engineering Systems Division
8 May 2015

In Partial Fulfillment of the Requirements for the Degree of

Master of Science in Technology and Policy

ABSTRACT

The space industry is currently at a significant inflection point. New economies are forming in low-Earth orbit (LEO), driven by miniaturization of technologies and the promise of lower launch costs, which should then allow many of these LEO systems to capitalize on designs incorporating smaller, shorter-lived spacecraft in highly-disaggregated constellations. Meanwhile, many spacecraft at geosynchronous orbit are continuing along a trend towards increasingly massive and longer-lasting satellites, and while they do represent some of the most exquisite, highest-performing satellites ever launched, some experts now feel that such trends are unsustainable and are beginning to place increasing strain on the underlying industry. To support current and future spacecraft, orbital infrastructures have been proposed as a means of providing on-orbit services to customer spacecraft and guiding space architectures towards more sustainable paradigms. In LEO, an infrastructure of communications and data relay spacecraft is envisioned as a means of aiding new and existing space enterprises in the areas of satellite connectivity and downlink capability. Meanwhile, an on-orbit servicing (OOS) infrastructure, located primarily in geosynchronous orbit, would provide services such as repair, rescue, refueling, and upgrading of customer spacecraft, in order to alleviate the identified space industry trends. Physics and cost modeling, as well as tradespace exploration, are used to identify optimal LEO infrastructure designs, while system dynamics modeling is used to assess the trends likely to occur in the overall space industry as OOS is incorporated into space architectures. The primary conclusion from the analysis of LEO infrastructure designs is that, when designing for global connectivity, there is an optimal design point between a small constellation of larger spacecraft and a very large constellation of small spacecraft, but this will also depend on the intended mission of the infrastructure and the number of customers expecting to be serviced. Then, for an OOS infrastructure, it is determined that relatively low costs and heavy incorporation of servicing capabilities into customer architectures are needed to ensure long-term sustainability of such a project. Finally, the policy implications for both infrastructure concepts are discussed, with a heavy focus on options for the funding and development regimes employed to implement the infrastructures, as well as the major political and legal implications expected to accompany these projects.

Thesis Supervisor: Dr. Daniel E. Hastings
Title: Cecil and Ida Green Education Professor of Engineering Systems and Aeronautics and Astronautics

ACKNOWLEDGEMENTS

Assisting the author of this thesis was Dr. Daniel Hastings, whose deep knowledge and experience in space systems was crucial to the formulation, guidance, and ultimate delivery of this work. Also massively helpful in the completion of this thesis was Capt. Paul La Tour, USAF, who worked closely with the author for over a year on all of the modeling and methods employed to generate the final results. Dr. Sumanth Kaushik of MIT Lincoln Laboratory provided a wealth of knowledge and guidance throughout, as did Dr. Donna Rhodes, Dr. Adam Ross, and all of the researchers at MIT's System Engineering Advancement Research Initiative (SEAri).

Finally, the author would like to thank his parents, grandparents, brother, and sister for all of their loving support and encouragement. They have never been quite sure what the author has been up to since leaving home, and he hopes it stays that way.

Contents

Figures and Tables 9

Introduction 13

Research Questions and Hypotheses 16

Background and Literature Review 19

Defining Infrastructure 19

Global Positioning System (GPS) 23

U.S. Interstate Highway System 25

U.S. National Airspace System (NAS) 27

On-Orbit Servicing (OOS) 29

DARPA OOS Programs 30

NASA OOS Programs 32

OOS Case Study: Hubble Space Telescope 33

Future Potential for OOS 35

Orbital Communications and Data Relay Infrastructure 38

TDRSS and EDRS 38

LEO Infrastructure Concepts 40

Space Industry Trends 41

Space Policy 45

1967 Outer Space Treaty 45

Space Policy of the United States 47

Space Industry Sectors 48

Methodology 51

Assessing Designs for a LEO Communications and Data Relay Infrastructure 51

Analysis Framework and Utility Function 52

Physics and Cost Modeling 56
Tradespace Exploration 62
Validation of Physics and Cost Models 66
Assessing On-Orbit Servicing: System Dynamics Approach 69
System Dynamics Methodological Background 70
Incorporating Behavioral Economics Theories with System Dynamics 73
Level of Analysis and Modeling Boundary 76
Building the System Dynamics Model 79
Major System Dynamics Model Structures 80
Using System Dynamics To Assess On-Orbit Servicing Effects 94
Results 97
Case Study 1: LEO Communications and Data Relay Infrastructure with Global Coverage 98
Analysis Framework and Parameters 98
Performance and Cost Estimates 100
Discussion 107
Case Study 2: Optimal LEO Communications and Data Relay Infrastructure Designs without Global Coverage 110
Analysis Framework and Parameters 110
Performance and Cost Estimates 115
Discussion 120
Case Study 3: Assessment of On-Orbit Servicing Effects 123
System Dynamics Parameters and Analysis Scenarios 123
Modeling Results and Trend Analysis 126
Discussion 133

Policy Recommendations 137

LEO Communications and Data Relay Infrastructure Policy Recommendations 137

Funding and Development Policies 137

Political and Legal Considerations 140

On-Orbit Servicing Infrastructure Policy Recommendations 145

Funding and Development Policies 145

Political and Legal Considerations 149

Conclusions and Future Work 154

LEO Communications and Data Relay Infrastructure 156

On-Orbit Servicing Infrastructure 159

Vision for Orbital Infrastructure 162

Bibliography 165

Figures and Tables

Figure 1. Design lifetimes of active geosynchronous spacecraft over the last two decades (UCS 2014)... 42

Figure 2. Launch masses of active geosynchronous spacecraft over the last two decades (UCS 2014). ... 42

Figure 3. Cost to initial operating capability (IOC) as design lifetime is increased (Saleh 2008). 43

Figure 4. Cost per operational day of spacecraft as design lifetime is increased (Saleh 2008). 43

Figure 5. Increases in average power of geostationary spacecraft (Jones and Spence 2011). 44

Figure 6. Additional power required of Landsat-8 in order to communicate with the infrastructure over a range of distances and data rates..... 58

Figure 7. Percentage increases in data returned per day by Landsat-8 with the infrastructure over a range of distances and data rates. 58

Figure 8. The locations of actual NASA ground stations were used to determine the infrastructure designs' downlinking capabilities (Mai, 2014). 61

Figure 9. Initial tradespace of infrastructure designs. 63

Figure 10. Graphical representations of a small selection of infrastructure configurations, taken from the tradespace of over 10,000 possible configurations. 64
Figure 11. 36-satellite infrastructure configuration, with six satellites equally spaced in six planes at a 63.5° inclination. 68
Figure 12. System dynamics ontological and epistemological framework (Sterman, 1994) 70
Figure 13. The System Dynamics research cycle (Pruyt, 2006) 71
Figure 14. This simple diagram illustrates the components of a system dynamics model. 73
Figure 15. Information stock structure, showing how the anchor values are represented as stocks, the adjustments as flows, and the interconnected nature of the anchors, adjustments, and space industry variables. 75
Figure 16. The satellite constellation production and replenishment stock and flow structure 81
Figure 17. System dynamics structure for modeling the accumulation of industry experience and the strength of learning curve effects 83
Figure 18. The satellite architecture design structure, which tracks the key design points for space vehicles (Design Life, Mass, and Cost), and adjusts each design point in order to keep cost per operational month low while factoring in any possible exogenous desires for increased performance. 85
Figure 19. System dynamics structure for modeling the risk- and cost-based space vehicle success rate. The negative value associated with SV failures attempts to drive up the success rate, while the costs associated with risk aversion attempt to drive down success rate, leading to a prospect value-based equilibrium. 88
Figure 20. This Value Function reflects the way in which losses (spacecraft failures), multiplied by the steeply decreasing value function below the point (0,0), are felt more strongly than gains (spacecraft successes), which are multiplied by the more moderately increasing value function above (0,0). 89
Figure 21. The Risk to Cost Function reflects the increasing multiplier on cost when success rate is increased beyond a baseline risk of 10% (as well as the cost savings which can be accessed below this baseline risk level). These estimates for cost multipliers were determined solely to aid in the modeling of the space industry and are not considered applicable to all spacecraft types. 90
Figure 22. The increase in average geosynchronous communications spacecraft power over the past three decades (adapted from Jones & Spence, 2011). 91
Figure 23. Normalized Technology Advancements is adapted from the growth in spacecraft power shown in Figure 22, and this acts as a divisor to decrease overall spacecraft mass and cost requirements... 92
Figure 24. Full system dynamics model used for analysis 93

Figure 25. Comparison of the model results for SV launch mass with geosynchronous satellite data provided by the UCS. 94
Figure 26. Comparison of the model results for SV design life with geosynchronous satellite data provided by the UCS. 95
Figure 27. Tradespace of LEO infrastructure designs which satisfy the requirement for global coverage. 102
Figure 28. Closer look at the Pareto front of optimal designs, with the design numbers taken from Table 3. 104
Figure 29. Global infrastructure design 217, which is composed of four orbital planes with eight infrastructure craft per plane (32 total spacecraft). 105
Figure 30. Global infrastructure design 148, which is composed of four orbital planes with 32 infrastructure craft per plane (128 total spacecraft). 105
Figure 31. Global infrastructure designs 12 and 48, which are composed of 10 orbital planes with 32 infrastructure craft per plane (320 total spacecraft). 106
Figure 32. Global infrastructure designs 147, which is composed of 4 orbital planes with 24 infrastructure craft per plane (96 total spacecraft). 106
Figure 33. The five reference infrastructure configurations used for initial testing and validation of the tradespace. 111
Figure 34. Percent contact time of various infrastructure configurations to a reference customer spacecraft. 112
Figure 35. Maximum outage time of various infrastructure configurations to a reference customer spacecraft. 113
Figure 36. Average outage time of various infrastructure configurations to a reference customer spacecraft. 114
Figure 37. Full tradespace of infrastructure designs. Blue points represent designs which satisfy minimum performance requirements for all three cases (Landsat-8, SkyBox, and Planet Labs). Green points represent those which satisfy SkyBox and Planet Labs, and red points satisfy just Planet Labs 117
Figure 38. A closer look at the "knee" of the Pareto front of infrastructure designs. It is in this region that the optimal designs are found, based on both total utility and overall cost. 118
Figure 39. Infrastructure configuration of a single orbital plane containing 16 total satellites (both pictures display the same infrastructure, from different vantage points). 119
Figure 40. Infrastructure configuration of two orbital planes containing 16 satellites each, for a total constellation of 32 satellites. 119

Figure 41. Results of the increasing spacecraft mass scenario. The blue line represents a No Cost OOS infrastructure, the red line represents a Low Cost infrastructure, the green line represents Moderate Cost, and the gray line represents High Cost. The black line displays the results of the Steady State results of the model without OOS implementation..... 127
Figure 42. Results of the constant spacecraft mass scenario. The blue line represents a No Cost OOS infrastructure, the red line represents a Low Cost infrastructure, the green line represents Moderate Cost, and the gray line represents High Cost. The black line displays the results of the Steady State results of the model without OOS implementation..... 129
Figure 43. Results of the decreasing spacecraft mass scenario. The blue line represents a No Cost OOS infrastructure, the red line represents a Low Cost infrastructure, the green line represents Moderate Cost, and the gray line represents High Cost. The black line displays the results of the Steady State results of the model without OOS implementation..... 131
Figure 44. This constellation configuration was identified as one of the optimal solutions for providing constant coverage to LEO customer spacecraft. 157
Figure 45. For a LEO infrastructure not requiring global coverage, it was determined that one orbital plane of 16 infrastructure spacecraft could still provide substantial benefits to LEO customers at signficantly lower costs. 158
Figure 46. These graphs display the most optimal case for OOS implementation; one in which spacecraft mass is reduced as OOS is incorporated into the reference architecture, with the blue and red lines representing free and low cost servicing, respectively. 160

Table 1. Comparison between actual and calculated specifications for Landsat-8..... 67
Table 2. Major system dynamics structures used for the model in this analysis 80
Table 3. Major solutions of interest in the global coverage tradespace; all but Design 12 exist on the Pareto front, indicating greatest performance to cost. 103
Table 4. A selection of the most optimal infrastructure designs, with their orbital configurations and performance characteristics listed..... 116

Introduction

Since the very first man-made satellites were launched into orbit, the space economy has exploded in scope and activity. International commerce now depends on the communication, navigation, and imagery capabilities provided by space systems, and governments and organizations across the globe utilize space for a huge variety of missions and value-generating ventures. New missions and uses for space systems are conceived each year, as space has become a permanent resource utilized by an expanding and diversifying collection of actors and stakeholders. Indeed, in its publication "The Space Economy at a Glance 2014," the Organization for Economic Cooperation and Development identified the following trends in space activity:

1. Globalization of the space sector is accelerating
2. The "democratization" of space is gaining ground
3. Many of the socio-economic impacts from space investments are becoming more visible (OECD, 2014).

Taken altogether, these trends show that the capabilities that the space economy offers will only become more widespread, more accessible, and in greater demand.

With space activity increasing and more customers and stakeholders coming into play, there have also been calls for greater availability of services directly benefiting operational spacecraft. For example, NASA has recently conducted a large-scale study of the benefits and considerations brought about by the implementation of wide-scale on-orbit servicing (OOS) operations (NASA Goddard Space Flight Center, 2010), while the Defense Advanced Research Projects Agency (DARPA) is now soliciting industry for potential designs and technological solutions needed for the development of an OOS infrastructure (DARPA, 2014). At the same time, research into the value of orbital services has begun to look beyond simple cost-benefit analyses, instead focusing on the system attributes, like flexibility and adaptability, which would likely be enhanced through servicing (Cohen, Richardson, Martinelli, & Betser, 2011), (Long, Richards, & Hastings, 2007), (Saleh, Lamassoure, Hastings, & Newman, 2003). Finally, there is a burgeoning commercial space economy in low-Earth orbit (LEO), driven by the increasing

miniaturization of technologies and the promise of lower launch costs (Morring Jr., 2015), (Meyer, 2014). Many of these ventures seek profits through the delivery of massive amounts of timely space-based data to customers on the ground, and thus would be likely to benefit greatly from a widely-available data relay infrastructure.

Services such as those described above are often classified as orbital infrastructure; that is, they represent systems which aid spacecraft in generating value and contribute to overall greater performance than if the spacecraft operated alone. An analogous example of such infrastructure found back on Earth would be national networks of roads, highways, bridges, tunnels, etc., which do not necessarily generate profits on their own but are still crucially depended upon by other enterprises as a means of transporting goods and services. Without such infrastructure, many businesses and organizations would be far more constrained in their ability to expand and service new customers. Similarly, new forms of orbital infrastructure would be highly beneficial to current and future space enterprises, allowing them to increase the performance of their systems and expand operations to a larger, more diverse clientele. For example, rather than launch spacecraft with all the components and fuel needed for 15+ years of operations, firms and agencies could instead launch satellites with the expectation that they will be refueled and upgraded with new components on a regular basis, allowing for a lower launch mass (which should decrease total mission costs) as well as decreasing the likelihood that spacecraft technologies will become obsolete before end of mission (which should increase overall system performance). Furthermore, new ventures in LEO could downlink far more data daily, as well as enjoy near-constant access to their orbital assets, through the utilization of a communications and data relay infrastructure specifically designed as a data backbone for a variety of operations and customer types.

This thesis considers two major concepts for orbital infrastructure: a low-Earth orbit communications and data relay infrastructure and an on-orbit servicing (OOS) infrastructure. Communications and data relay is a crucial component to any space mission, and it is believed that many emerging ventures and customers in low-Earth orbit would benefit significantly from the timely and efficient downlinking of data that could be achieved through subscription to a publicly-available data relay infrastructure. Similarly, OOS has been proposed as a means of

refueling, repairing, and upgrading customer spacecraft, with such a service causing significant shifts in the design and risk management paradigms that drive the space industry. While there are certainly other concepts for orbital infrastructures, these two projects are singled out for analysis in this thesis due to both their technical feasibility and their seemingly obvious benefits to current space architectures. In other words, it was determined that for these two infrastructure concepts especially, much has been done to describe and demonstrate their value, and it is now largely a matter of finding the most optimal, technically-grounded policies to turn these concepts into reality.

In order to achieve this goal of determining the best policies to build and operate these infrastructures, two methodologies were employed. First, cost and physics-based modeling is used to assess the performance of a LEO communications and data relay infrastructure, and the analysis focuses heavily on variables such as: orbital configuration of the infrastructure; number of spacecraft required; power, mass, and data rate design points of infrastructure spacecraft; and overall costs of the infrastructure project. The goal of the analysis was to find the most optimal LEO infrastructure designs, balancing performance and cost as well as assessing the infrastructure's utility for a variety of customer types, and then use these designs to inform the discussion of policy options. Second, system dynamics modeling is used as a means of evaluating the long-term shifts in the overall space industry brought about by the implementation of OOS. By evaluating the trends in program cost, satellite design life, and industry experience simulated in the system dynamics model when risk posturing is shifted through OOS, one can gain insights into the policy regimes most likely to induce long-term viability and sustainability in the infrastructure.

Finally, this thesis concludes with policy proposals for the successful implementation of orbital infrastructure projects, as well as a discussion of the political, legal, and economic issues underlying any such large-scale space systems. It is hoped that the work conducted within the completion of this thesis will aid policymakers and space systems engineers with the development and launch of successful orbital infrastructure projects, to ensure that space remains an active and growing venue for a huge variety of future endeavors.

Research Questions and Hypotheses

Several "big picture" research questions were needed to provide an overall framework for the research conducted for this thesis. These questions were informed significantly by the current state of the space industry and the proposed benefits that orbital infrastructure would contribute, should they be developed. While these questions are necessarily directed at specific infrastructure concepts, the insights derived from the analysis should also be capable of informing the design of many similar projects. For instance, the analysis of the size and cost of a constellation needed for adequate infrastructure performance, as well as the many policy implications surrounding such a project, are intended to provide at least a means for comparison of similar space-based project designs and policy options. Similarly, a system dynamics evaluation of OOS, while not necessarily providing hard estimates of OOS prices and capabilities, should cause policymakers to consider long-term issues surrounding infrastructure projects, as well as the ways in which new technologies and paradigms can affect system variables which they may not have considered previously.

Along these lines, research questions for the case of a LEO communications and data relay infrastructure are as follows:

- What are the optimum architectural design points for a globally-visible (to a LEO spacecraft), radio-frequency enabled communications and data relay infrastructure, which enables the infrastructure to render maximum data relay services for customers on orbit?
- What are the optimal architectural design points for a low-Earth orbit communications and data relay infrastructure which is NOT globally-visible, but which is still designed to benefit the largest number of customers possible?
- What are performance and cost estimates for both infrastructure cases?
- What are the available policy options for both financing and ensuring the long-term success of such an infrastructure?

These research questions were intentionally designed to be somewhat general, in order to allow the tradespace exploration of infrastructure designs to inform as much as possible the final conclusions. The tradespace exploration should grant a degree of clarity between design trade-offs; for instance, does it make more economic sense to have an infrastructure composed of a

few large, powerful spacecraft or many smaller, less powerful spacecraft? The final cost estimates will also do much to inform the best policy choices for the funding and development of the infrastructure. For example, if the final infrastructure cost is relatively low (compared to other similar space systems), then a fully private approach may be best, but if the infrastructure is predicted to be much higher in cost, then some amount of government funding may be recommended. In addition to cost, though, expected return on investment in the LEO data relay infrastructure must also be considered; i.e., private industry would be unlikely to take on a project which does not have a high likelihood of profitability (or if profitability is not expected until far into the future). In such a case the public sector may have to take charge of developing and providing this service, and its costs would likely have to be justified as necessary to supporting science or defense missions (in addition to aiding private ventures).

Next, for the case of on-orbit servicing (OOS), the following research questions were developed:

- Assuming an OOS infrastructure is implemented and incorporated into the overall architecture of a reference spacecraft constellation, what cost structures and policies provide the greatest likelihood of long-term sustainability and beneficial impacts?
- What are the "break-points" in cost and overall architecture risk, after which OOS will no longer remain economically viable?
- What are the major policy considerations that will accompany the development and operation of an OOS infrastructure?

For OOS, it is clear that the research emphasis is less focused on technical design of the infrastructure (OOS having already been demonstrated to be technically feasible by DARPA and NASA, which is discussed further in the Background below). Rather, this thesis is interested in the overall space industry response to the implementation of OOS. More specifically, if it is believed that massive, powerful, expensive spacecraft are bad for the overall industry, and that this trend in spacecraft design will continue largely unabated without a significant paradigm shift, will OOS play a demonstrable role in initiating and sustaining that shift? And what are the general cost and risk levels required for OOS to become a permanent fixture in space architectures?

Finally, as with any new technology, there are several additional policy factors that must be considered for both orbital infrastructure projects, such as liability for damages, issues of security and transparency, the determination of operational codes of conduct, and provisions for international coordination and participation. These issues are tied to the analysis results and discussed in depth.

Background and Literature Review

In order to properly frame the issues addressed in this thesis, the technical and policy backgrounds applicable to orbital infrastructure, as well as several aspects of the space industry in general, must first be discussed. In addition, a review of the pertinent literature on infrastructures in general and the completed work on orbital infrastructure projects is also provided, with an emphasis placed on the lessons which can be learned from past infrastructure projects (both in space and on Earth). With this background knowledge in place, the analysis completed within this thesis can be fully understood, especially with regards to the potential benefits and challenges of the two orbital infrastructure concepts in focus.

Defining Infrastructure

Infrastructure is absolutely integral to the basic function and growth of modern economies. Virtually no one would disagree that developed nations leading the world in economic activity would be in their current position if not for significant investments in infrastructure, from railroads and highways to electrical power grids and high-speed data networks. Infrastructure as a defined aspect of economic and sociotechnical study is a relatively new term, though. The Oxford English Dictionary first listed the word "infrastructure" in 1927, stating that it comprised the "tunnels, bridges, culverts and 'infrastructure' work generally" of French railroads (Batt, 1984). It is a combination of the Latin prefix *infra-*, meaning "under," and the common word *structure*, such that it literally means the underlying foundation or support structure for some activity or enterprise. The term infrastructure also gained prominence in NATO's early strategies for the Cold War, referring to all the necessary logistical and support structures required for effective military operations, from roads and bridges to radio communications and stockpiles of war supplies (Batt, 1984).

Today, the term infrastructure has become so widely used and interpreted that it has become vital to make its definition clear and unambiguous. This is especially important when examining concepts in orbital infrastructure, as they represent relatively uncharted territory in the provision of service-based infrastructures. In a survey conducted in 2011, Beeferman and Wain found the

following to be the most widely accepted definitions for infrastructures among U.S. public sector pension funds:

- Provide[s] services and support that are basic to the functioning of a community, organization, or society and crucial to its economic productivity
- The basic physical and organizational capacities and resources needed for the operation of a society or enterprise or are necessary for an economy to function
- Are capital intensive/have high fixed costs and long economic lives and have strong links to economic development, and a tradition of public sector involvement
- Are essential to driving sustainable economic development and growth, lifting levels of productivity and boosting employment and critical to encouraging business innovation and improving global competitiveness of enterprises (Beeferman & Wain, 2012)

Thus, it is clear that infrastructure is most closely associated with systems of underlying structures or services, which support a variety of activities but are not themselves the producers of goods or economic activity. Furthermore, infrastructure is viewed as crucial to economic productivity and basic operations for many enterprises, and the presence of infrastructure is considered essential to driving economic growth and innovation; i.e., infrastructure can act as a catalyst to greater economic productivity. However, the benefits of infrastructure are also constrained by the large investments and high costs typically associated with such projects.

With the above definition in mind, the following definition for "orbital infrastructure" is proposed and adhered to throughout the remainder of this thesis:

An orbital infrastructure is an architecture that provides on-orbit services, support, or resources to participating space systems or space-based enterprises.

The following collateral classifications are also proposed as expected characteristics of an orbital infrastructure, and with the above definition they establish a taxonomy of infrastructures by which to categorize future architectures.

- The value proposition of the infrastructure is demonstrated solely on orbit (i.e., the infrastructure provides services or support to on-orbit assets, rather than to ground-based assets utilizing satellite services).
- Participating systems or enterprises designed with the expectation of infrastructure utilization would exhibit significantly reduced performance in the absence of the infrastructure.
- The infrastructure is reliable and provides crucial resources or services.
- Use of the infrastructure is available to all customers who adhere to predetermined utilization policies.

Primary examples of space architectures which would fit the above taxonomy, and which are examined in far greater depth in this thesis, are on-orbit spacecraft servicing systems and resource-sharing LEO communications and data relay constellations. Currently, there are only limited operational examples of these infrastructures; for instance, on-orbit servicing has been demonstrated on a small scale several times, but has yet to become a more widely available resource for spacecraft (see background section on on-orbit servicing). Similarly, resource-sharing communications and data relay infrastructures, composed of many small spacecraft and/or hosted payloads, have been described in concept as efficiently allocating in-orbit resources like processing power, link capacity, and data storage amongst disparate and heterogeneous systems. Such a system would potentially allow customer spacecraft to more easily link, store, and downlink data, but this has yet to be validated operationally in a publicly available setting (see background on resource-sharing infrastructures). Thus, both of these types of infrastructures are relatively well-understood systems that now require the investment and development necessary to become fully operational.

An orbital infrastructure for space debris removal and mitigation is also an example of an infrastructure falling within the taxonomy (but not examined as deeply in this thesis). Concepts for a future space debris removal infrastructures represent a somewhat special case, due to the fact that a large aspect of space debris removal—debris detection and classification—has already been largely established and is operational today. In fact, space situational awareness infrastructures are now widespread and have brought together a variety of public and private entities in an effort to detect and avoid orbital collisions (Weeden, 2014). Since these detection

systems are almost entirely ground-based, this also draws out a key point about orbital infrastructures; namely, that the space systems comprising an infrastructure need not be themselves on orbit, but rather must simply provide services and demonstrate their value to orbital assets. In the case of space situational awareness infrastructures, while they can not yet actively remove or mitigate space debris, they do typically provide sufficient warning of increased risk of debris collisions to spacecraft operators. The next step for these infrastructures would of course be the development and fielding of space systems that actually do remove or mitigate the threat posed by space debris. Finally, it should be noted that space situational awareness infrastructures could also prove beneficial to the development of other orbital infrastructures, as the ability to track and monitor space objects would be very helpful to servicing infrastructures and (perhaps to a lesser extent) communications and data relay infrastructures.

Furthermore, the vast amount of coordinated space and ground assets required for human space operations is an incredible example of space-based infrastructure, and the International Space Station (ISS) is the most visible result of this system of systems (International Space Station: Research and Technology, 2015). Indeed, both infrastructures considered in this thesis owe much of the credit for their development thus far to human exploration of space. For example, communications and data relay spacecraft have allowed for continual monitoring and direction of assets like the ISS and Space Shuttle. In addition, many of the technologies needed for OOS were first pioneered within the on-orbit assembly of human space habitats, as well as astronaut-based servicing of assets like the Hubble Space Telescope (described in greater detail below). Thus, the contributions of human space exploration to any future space-based enterprises should not be overlooked. However, within this thesis, human space operations and their accompanying infrastructures are considered less pertinent to the orbital infrastructures in question here, as they are designed specifically to serve within the role of human space exploration, rather than the provision of services to a much larger range of customers and space-based assets.

There are already several operational space systems that perform some of the expected functions of orbital infrastructure. For instance, orbital communications and data relay services have been a fixture of space operations for decades, with large systems like NASA's Space

Network and its Tracking and Data Relay Satellites linking many spacecraft and ground stations (Tracking and Data Relay Satellite Project, 2012). Such systems have developed and validated much of the technology that could one day be used by a resource-sharing data relay constellation. Also, as has been noted previously, small scale tests of satellite servicing have already been successfully completed, and the technologies and precise maneuvers necessary to these operations have been proven many times in both manned and unmanned human exploration missions (Benedict, 2013). If one considers the determination of a satellite's location to be a service, then an example of a successful orbital infrastructure that is now well-established and highly relied upon is the Global Positioning System, and this is examined in greater depth below.

Global Positioning System (GPS)

The Global Positioning System (GPS) is a technological marvel that can provide many lessons for the design of future orbital infrastructures. Requiring at least 24 operational satellites (with often 30 or more in use since reaching full operational capability), GPS provides highly accurate, nonstop, worldwide location information to any user with a compatible receiver (Hegarty & Chatre, 2008). The GPS constellation is situated in medium-Earth orbit (MEO) in six equally-spaced orbital planes at an altitude of approximately 20,200 kilometers, which ensures that at least four satellites are visible to users anywhere on the Earth's surface (GPS.gov 2014). Due to many of the satellites lasting much longer than expected, the GPS constellation configuration was recently adjusted to an "Expanded 24" configuration, allowing three additional satellites to enter the operational configuration and create "the most optimal geometry in its 42 year history, maximizing GPS coverage for all users worldwide" (50th Space Wing Public Affairs, 2011).

The origins of GPS can be traced back to an early Cold War desire for a globally available navigation system, which would allow military users to accurately determine their location as well as perform more advanced functions like weapons targeting. Indeed, the concept of a satellite navigation system was first validated experimentally when two U.S. physicists tracked *Sputnik* and realized that, just as they could pinpoint the first manmade satellite's position using Doppler shift effects, so too could a user on the ground determine his own location using signals

received from satellites passing overhead (Guier & Weiffenbach, 1997). The plans for a U.S. navigational system quickly followed, and in 1964 the first successful satellite navigation program, TRANSIT, entered operational service, using six satellites to provide U.S. Navy ships with an accurate location fix once an hour (Danchik, 1998). Meanwhile, the U.S. Air Force and Army began pursuing their own satellite navigation systems, and in 1973 it was decided that the technologies developed and lessons learned from all services would be combined into one large joint system, called the Navstar-GPS program. The first 11 prototype satellites for this system were launched between 1978 and 1985, directly followed by the launching of operational satellites, and over 57 GPS have now been successfully placed in orbit (Hegarty & Chatre, 2008).

While GPS was originally developed as an exclusively military system, President Reagan ordered that it be made available for international civil use upon reaching operational capability, in response to the downing of Korean Airlines Flight 007 after it had accidentally strayed into Soviet Union territory in 1983 (Pace, et al., 1995). Civil use then became popular almost immediately, with many nations and international organizations adopting the use of GPS even before fully global coverage was achieved with the launching of the 24th satellite in 1994. In addition, the highest quality GPS signals were originally reserved only for military use, while civil signals were intentionally degraded in accuracy, but this practice was terminated by President Clinton in 2000 (and in 2007 the U.S. announced that future GPS satellites would no longer be manufactured with the capability to selectively degrade civil signals) (Hegarty & Chatre, 2008). While operated and maintained by the U.S. Air Force, GPS is now considered a U.S. national asset, with significant assurances provided in both Presidential policy directives and Congressional legislation that international, unrestricted GPS access has become a permanent fixture of the system (GPS.gov 2013).

The development and evolution of GPS provides several pertinent lessons to the design and implementation of any orbital infrastructure project. The positioning, navigation, and timing (PN&T) services provided by GPS, while largely confined to Earth-bound applications, have also been shown to provide useful location information for other spacecraft, even those which orbit far higher than the GPS constellation (Winternitz, Bamford, & Heckler, 2009). Thus, GPS can effectively be considered an orbital PN&T infrastrastructure, even if it may not have

originally been conceived and designed as such. Furthermore, the way in which GPS was originally developed as a military navigation system, but has since gained far greater prevalence in terrestrial civil applications, reflects one path by which new orbital infrastructure projects could be funded and developed. Indeed, few could have predicted just how integrated GPS has become in international commerce and how essential the system now is to a huge variety of civil enterprises. As GPS enabled an abundance of new technologies and ventures in many economic sectors on Earth, so too could new forms of orbital infrastructure create opportunities for economic activity and enhanced services in space.

Finally, the fact that GPS navigation services are provided completely free of charge should not be overlooked as a key component of its policy regime. While it is difficult to tell now how the proliferation of GPS technologies would have differed if they were accompanied with participation fees, it is likely that the use of GPS would not be as ubiquitous—and would have much less influence on operations outside the US military—were it not free of charge to civilians. In contrast to this is the European Union's Galileo satellite navigation system (still in development), which is slated to offer both free navigation services to all users, as well as a fee-based, higher-accuracy signal (Galileo Services, 2015). While it is not yet known exactly how much it will cost to access this more accurate signal, it is likely that this will have a significant effect on Galileo's uptake within its target customer base of European consumers, especially if the free Galileo signal proves to be much less accurate than its GPS counterpart.

U.S. Interstate Highway System

One of the most obvious examples of a successful national infrastructure program is the Dwight D. Eisenhower National System of Interstate and Defense Highways (more commonly known as the U.S. Interstate Highway System). This nationwide network of highly standardized roadways was promoted, authorized, and enacted largely by its namesake, President Eisenhower, throughout his presidency in the 1950's. Before the passing of Eisenhower's legislation, interstate highway construction had largely occurred in short spurts, with the individual states each building their own roadways. The federal government had passed bills allowing it to match funding for roadway construction and enforce some basic highway standards, but there was a

lack of nationwide planning and coordination. President Eisenhower had seen first-hand as a young Army officer in 1919 how difficult it was to motor his unit across the country, and he was also impressed by Germany's system of autobahns, or network of standardized rural highways which allowed its military to quickly move troops and vehicles across the country. Thus, it was through Eisenhower's leadership that Congress passed the Federal Aid Highway Act of 1956, which established design standards and designated over 41,000 miles of roadways as part of the National System of Interstate and Defense Highways. This act also set the federal government's share of the project costs at 90%, to be funded through the Highway Trust Fund (which received revenue from federal gas and other motor-vehicle user taxes). Indeed, this trust fund was integral to Eisenhower's vision of a "pay-as-you-go" development paradigm, in which the program would be largely self-financed and would not contribute to the federal budget deficit (History of the Interstate Highway System and Interstate FAQ's, 2015).

States began constructing interstate highways almost immediately after the legislation was passed, spurred on by the influx of cash from the federal government. Indeed, a key aspect of the Federal Aid Highway Act was that it left the states in charge of construction and maintenance of their roadways; the federal government simply provided funding and oversight. By placing ownership of interstate highway on the states, the federal government ensured that the responsibility for maintenance and traffic enforcement would remain on a local level, while it still retained the right to direct overall interstate highway policy and the direction of funding. As for the source of this funding, it is estimated that 70% of the construction and maintenance costs over time for interstate highways are paid through user fees, primarily fuel and other motor-vehicle taxes (Weingroff, 2011). The remaining costs of the highways are covered through general state and federal funding, bond issues, and certain designated property and other taxes. Also, highways that were constructed before the passing of Eisenhower's bill in 1956 are often still operated as toll roads (and are generally ineligible for federal funding for maintenance). Furthermore, as some American suburbs have expanded and some highways have come into greater use and demand, many states are returning to the use of toll roads as a means of ensuring adequate funding for new infrastructure (History of the Interstate Highway System and Interstate FAQ's, 2015).

The development, construction, and maintenance of the U.S. Interstate Highway System, like the GPS constellation, offer many lessons for the planning of future orbital infrastructure projects. Similar to GPS, the highway system originally had military applications in mind, in that it would allow troops and equipment to cross the country much faster than if there was not a nationwide network of standardized roadways. However, civilian use has far outstripped any military needs for the highway system, and economic activity has benefited enormously from this infrastructure project, to the point where the interstates are a vital and indispensable component of national commerce. Furthermore, the way in which the initial construction of the interstates was almost entirely funded by the federal government, but ownership of the roadways was still passed on to the states, also offers a potential concept for the development of orbital infrastructures. The interstate system has long been regarded as a national asset, with federal funding and oversight to reflect this, similar to how an orbital infrastructure could one day be regarded as a national (or even international) asset which is indispensable to future space operations. Finally, the use of toll roads on new, heavily-traveled, high-demand routes provides an additional funding concept for future orbital infrastructure projects, where demanded services could be paid for on a per-use basis.

U.S. National Airspace System (NAS)

Another example of an extensive, federally administered infrastructure is the U.S. National Airspace System (NAS). As aviation systems have increased in complexity and economic importance over the past century, it is crucial that there be sufficient infrastructure underlying aviation operations. In the earliest days of aviation, pilots largely relied on their own facilities and flying experience to conduct aerial operations; naturally, this led to an extremely dangerous, poorly-defined flying environment. It was the aviation industry leaders themselves who called for federal action in the regulation and standardization of aviation practices, and this led to the Air Commerce Act of 1926, which charged the Secretary of Commerce with "fostering air commerce, issuing and enforcing air traffic rules, licensing pilots, certifying aircraft, establishing airways, and operating and maintaining aids to air navigation" (Federal Aviation Administration, 2015). Then, in 1934, the first air traffic control (ATC) centers were established as a means of tracking and coordinating flights over large distances. While initially funded and operated by

private airlines, these early ATC centers were soon transferred to federal authorities, and while most aircraft at the time lacked radio communications, the centers coordinated in the tracking of aircraft arrivals and departures as an early means of airspace management (Federal Aviation Administration, 2015).

The foundations of the modern aviation infrastructure lie in the Federal Aviation Act of 1958, which established the Federal Aviation Agency (now the Federal Aviation Administration, or FAA) and charged it with the direction and control of the navigable airspaces within the U.S. In line with these duties, the FAA created the NAS as a means of establishing a safe and efficient airspace environment for civil, commercial, and military aviation. The history of the NAS is marked with continual evolution and improvement, as new technologies and challenges have caused shifts in the aviation landscape, and throughout this history the airspace infrastructure has only grown in scope and importance. The NAS is now composed of a vast network of air navigation facilities, ATC facilities, and airports, which provide the necessary services for thousands of domestic and international flights throughout the U.S. each day. Strict rules are also in place to govern air traffic, and these rules are clearly updated and promulgated on a regular basis (Federal Aviation Administration, 2015). Indeed, one of the greatest signals of the success of the NAS is the way in which it operates so quietly and efficiently; it is likely that few among the millions of airline passengers each year realize the incredible amount of manpower and planning that goes into the management of U.S. airspaces.

As with the GPS constellation and Interstate Highway System, the NAS offers several lessons useful to future infrastructure projects. First, and most importantly, the FAA has established very clear, unambiguous regulations with regards to the participation of pilots and airlines in the NAS, such that virtually every aspect of aviation operations has a corresponding rule. While such a massive amount of regulations might at first seem onerous and perhaps even detrimental to private enterprise, the FAA largely mitigates this by regularly issuing new compilations of aviation rules and by providing notices to operators of any new or updated rules (Federal Aviation Administration, 2014). The FAA also maintains strict licensing and pilot currency requirements, with a major component of this process being the demonstration of advanced knowledge of FAA regulations. Clear, unambiguous rules for national airspaces also allow for the seamless integration of multiple types of operations, with each actor—commercial,

private, or military—knowing exactly how their conduct within a particular airspace is controlled and regulated. In addition, use of the NAS infrastructure comes at very low cost to users. ATC facilities and navigational aids are free and open for use by all pilots, and while most airport facilities are operated privately on a for-profit basis, the majority of basic ground service fees represent a very small portion of the overall cost of operating an aircraft (Federal Aviation Administration, 2006).

In its regulation of the NAS, the FAA has also maintained relatively close ties with the private aviation industry and other crucial stakeholders, demonstrated by the way that some components of the NAS are operated publicly while others are provided by private partners, and in this manner the infrastructure can be reshaped and reorganized in ways which best suit those who will actually use it. Furthermore, while some may criticize the FAA for not updating regulations fast enough, it should be noted that over its history this organization has demonstrated a marked willingness to update and revise regulations as technologies and expectations for airspace use evolve. This is an important consideration; few regulatory statutes are likely to remain relevant forever, and thus for any large-scale infrastructure it is crucial that policymakers regularly enact reasonable revisions to rules and regulations, to address changes in the operational environment.

On-Orbit Servicing (OOS)

The ability to repair, refuel, and reposition spacecraft in orbit has often been proposed as a means of radically restructuring the conduct and planning of space missions. As the predominant spacecraft design paradigm stands today, once a satellite has been placed into its predetermined orbit, there is absolutely no way for engineers to in any way modify or fix that satellite, beyond changes in software or self-repositioning (which is likely impossible, or at least very undesirable given the large amount of precious on-board fuel which is required for significant orbital maneuvers). Naturally, this paradigm places significant constraints on satellite design parameters and mission flexibility, forcing engineers to construct satellites that can only be estimated to remain suitable in performance throughout those satellites' lifetimes. Also, should an anomaly strike, mechanical or otherwise, before the end of a satellite's operational

lifetime, then options for repair and recovery of prior levels of service are very limited. While engineers have certainly devised ingenious methods for salvaging value from damaged or malfunctioning spacecraft in the past (for instance, by shifting the mission's concept of operations to accommodate the spacecraft as is), few would dismiss a means of fixing the spacecraft and continuing its mission as originally planned, if such a service was available.

DARPA OOS Programs

To this end, several preliminary tests and studies have been conducted to examine the feasibility of on-orbit servicing. One of the most prominent projects in this arena was the Defense Advanced Research Projects Agency's (DARPA) Orbital Express program, which validated the "technical feasibility of robotic, autonomous on-orbit refueling and reconfiguration of satellites" (Walker, 2007). During this mission, two spacecraft were launched: a prototype serviceable satellite called NextSat and a prototype servicing satellite called ASTRO. Launched into a circular, 492-km orbit in 2007, the two spacecraft performed or attempted the following operations over a three month period (all adapted from Walker 2007):

- Transfer of hydrazine monopropellant between ASTRO and NextSat.
- Use of ASTRO's robotic arm to detach and reattach a flight computer from a bay onboard ASTRO.
- Use of the arm to detach a battery from its bay on ASTRO and transfer it to a similar bay onboard NextSat.
- Use of ASTRO's arm to couple to a NextSat grappling fixture and "berth" the spacecraft, allowing a separate capture mechanism to perform final docking.
- Multiple rendezvous and capture scenarios with a number of different sensors, at ranges up to seven kilometers (4.4 miles).
    - "Soft" (low momentum transfer) docking with a special, three-fingered capture mechanism.
    - Use the robotic arm to grapple and pull NextSat into the capture system envelope.

- Use of an infrared camera, multiple optical cameras, a laser rangefinder, and NASA's Advanced Video Guidance System, which uses retroreflector targets on NextSat and a laser on ASTRO to estimate a client spacecraft attitude and facilitate docking (Walker, 2007).

The DARPA Orbital Express mission "met all of its mission success criteria" (Orbital Express On-Orbit Mission Updates, 2007). In so doing, it provided resounding evidence of the technical feasibility of on-orbit servicing, provided that the spacecraft being receiving fuel and upgraded components is compatible with the servicing spacecraft.

Another DARPA program which is currently working to validate OOS capabilities is the Phoenix program, which seeks to "develop and demonstrate technologies that make it possible to inspect and robotically service cooperative space systems in GEO and to validate new satellite assembly architectures" (Phoenix Goals, 2015). The Phoenix program has three primary components:

- Developing advanced GEO space robotics, which can perform functions such as assembly, repair, asset life extension, refueling, etc. and which would eventually become part of a future robotic assembly platform operating continuously at GEO
- Developing and testing the "satlets" concept, in which a spacecraft is composed of small individual modules that are aggregated on orbit. This would allow new technologies and capabilities to be added to the spacecraft over time, with each new module being easily attached to prior modules and all the components sharing common power and data resources.
- Testing the Payload Orbital Delivery (POD) system, in which a standardized mechanism allows new components (possibly satlets or other needed parts) to be carried up to GEO onboard commercial craft and then taken out and used easily by a servicing spacecraft (Phoenix Goals, 2015).

All of the components of the Phoenix program are aimed at creating a servicing infrastructure at GEO that can quickly and easily provide existing and future satellites with needed repairs and upgrades. While some spacecraft may require specific design alterations to become capable of taking advantage of satlets and other capabilities offered, the Phoenix program is still aimed at

validating an OOS concept which would not require inordinate changes to current spacecraft design paradigms.

NASA OOS Programs

NASA has also conducted a great deal of satellite servicing tests and operations, with its Satellite Servicing Capabilities Office (SSCO) spearheading many of the efforts. Currently, one of its primary programs is the Robotic Refueling Mission (RRM), which involves testing of satellite servicing components on the exterior of the International Space Station (ISS). While the main thrust of the mission is described in the name, one thing that makes this program unique from other servicing demonstrations is that it actively seeks solutions for servicing satellites which were not originally designed for such operations (Escobedo Jr., 2014). Thus, RRM is working towards solutions that could potentially be applied not simply to satellites purpose-built for servicing, but to a large portion of currently operational spacecraft. In order to perform this testing, the ISS's Canadian-built robotic arm, Dextre, uses four unique tools to perform various tasks on a custom-built RRM module mounted on the ISS exterior and containing common satellite components. This RRM module is roughly the size of a washing machine and contains about 0.4 gallons (1.7 liters) of ethanol to simulate fluid transfer on orbit (Escobedo Jr., 2014). Dextre remains under the control of engineers on the ground during servicing operations, and according to Escobedo, there are six primary RRM tasks:

- Launch Lock Removal and Vision - On September 6-7, 2011, mission controllers used the Dextre robot to successfully release the "launch locks" on the four RRM servicing tools. These locks kept the RRM tools secure within the RRM module during the shuttle Atlantis' flight to the International Space Station. Once this task was complete, the RRM team then used Dextre's cameras to image the RRM hardware in both sunlight and darkness, providing data that NASA's Satellite Servicing Capabilities Office used to develop machine vision algorithms that work in spite of harsh on-orbit lighting. All subsequent RRM tasks were performed using RRM Tools.

- Gas Fittings Removal - Marking the first use of RRM Tools on orbit, Dextre used the RRM tools to remove the fittings that many spacecraft have for the filling of special coolant gases.
- Refueling - After Dextre opened up a fuel valve that was similar to those commonly used on satellites today, it tested the transfer of liquid ethanol through a sophisticated robotic fueling hose.
- Thermal Blanket Manipulation - Dextre practiced slicing off thermal blanket tape and folding back the thermal blanket to reveal the contents underneath.
- Screw (Fastener) Removal - Dextre robotically unscrewed satellite bolts (fasteners).
- Electrical Cap Removal - Dextre removed the caps that would typically cover a satellite's electrical receptacle (Escobedo Jr., 2014).

Engineers working on the RRM also hope to use the data gathered from all of these tasks to improve simulators back on Earth, such that future servicing missions can be practiced in the most accurately simulated environment possible before moving on to the real tasks on orbit. The RRM continues today, with new components being delivered over time and more demonstrations of on-orbit tasks scheduled, including tests involving the inspection of satellite components and experiments involving solar cells and paint materials (Robotic Refueling Mission, 2015).

OOS Case Study: Hubble Space Telescope

While DARPA and NASA have proven experimentally that OOS is technically feasible, an actual operational example of the benefits of OOS is provided by the Hubble Space Telescope, which underwent five separate servicing missions to fix and improve many aspects of this well-known and long-lived spacecraft. The first of these missions was absolutely critical, in that it fixed Hubble's incorrectly-shaped optics; without this fix, completed in December 1993, the telescope would have severely diminished performance and would likely not have been capable of returning the incredible images for which it has become so well-known. The second mission, completed in 1997, focused less on repairs of the Hubble and more on upgrades, including the addition of more advanced instruments which utilized technologies which were not available when the telescope was originally constructed. Repairs were again the focus of the third mission

(actually conducted in two parts, 3A and 3B), as four of Hubble's six gyroscopes had failed due to age and use. Hubble also needed replacement components for its solar array and infrared telescope, along with the installation of new scientific instruments. Finally, the last servicing mission to Hubble was a busy one, being the culmination of several efforts to make lasting, final upgardes. Completed in 2009, the last mission successfully added several new instruments, conducted an impromptu on-site repair of two instruments, and replaced many of Hubble's basic components with upgraded versions (Team Hubble: Servicing Missions, 2015). Taken altogether, these missions have left NASA with an orbital telescope that is far more capable and enduring than the one launched into orbit over 20 years ago.

Treated less as a spacecraft and more as a floating national observatory, the Hubble Space Telescope was designed to have easily accessible and upgradeable components. It also required servicing visits by astronauts operating out of the Space Shuttle, who came ready with specially-designed tools and abundant experience gained from extensive practice runs of each servicing mission back on Earth. Thus, the Hubble does represent a special case amongst servicing missions, in which cost was largely not an obstacle and significant NASA resources were at engineers' disposal. Indeed, the high cost of the Hubble servicing missions has most likely acted as one of the strongest deterrents against designing other orbital observatories for regular servicing (at least by human operators). The last Hubble mission cost over $900 million, and was so expensive and controversial that it affected several of NASA's downstream projects (Parks & Berger, 2006). Furthermore, NASA's 2010 OOS study identified the excessive costs of Hubble servicing as one of the major contributors to the "myth" that servicing is necessarily a very costly, fundamentally prohibitive means of increasing space systems' performance (NASA Goddard Space Flight Center, 2010)

Still, experiences with Hubble offer many lessons for future OOS operations. The fact that the entire Hubble science mission was rescued from the brink of failure by the first servicing mission speaks volumes for one of the most common rationales for developing OOS capabilities; that is, it reinforces the usefulness of OOS for fixing incredibly expensive spacecraft which otherwise would be completely lost or would experience significantly degraded performance due to some initial anomalous failure. Furthermore, the Hubble has received significant upgrades in technology over its lifetime, and it has even incorporated scientific instruments which had yet to

be conceived when the telescope was originally constructed. The Hubble's lifetime has also been significantly extended due to servicing missions, with the last mission providing it with the new batteries and other components needed to allow it to operate well into the current decade (Team Hubble: Servicing Missions, 2015). Without a doubt, the Hubble Space Telescope would be a far less capable spacecraft, as well as one which would likely have been decommissioned long ago, if it had not been designed with OOS in mind.

Future Potential for OOS

Clearly, a great deal of technical research and testing has been completed, such that few would argue against the technical feasibility of OOS. However, the overall space industry has yet to definitively embrace design paradigms which incorporate OOS; indeed, the case of the Hubble Space Telescope begs the question of why, if OOS has been demonstrated to be so beneficial, more orbital observatories and other missions have not pursued similar capabilities. A massive study of satellite servicing's history, capabilities, and future prospects was conducted in 2010 by NASA's Goddard Space Flight Center, and within this study the authors identified three **"myths"** which have hampered OOS's development and uptake by the overall industry:

- **Myth #1: "There is nothing to service."** As the study found, and as is described in further detail below, a significant number of servicing opportunities have been identified each year. These OOS opportunities range well beyond the rescue of failed spacecraft, to include refueling and inspection of the commodity satellites vital to space communications infrastructures.

- **Myth #2: "Servicing is too costly."** This is generally quoted in reference to the Hubble servicing missions, but the study authors made clear that Hubble is an extreme case of servicing, involving manned spaceflight and highly sophisticated operations, which is not representative of the vast majority of potential servicing activities. This is especially true when considering far simpler, far less expensive refueling or repositioning activities involving unmanned telerobotic operations.

- **Myth #3: "Satellites cannot be serviced unless they were designed to be serviced."** Again using Hubble as an example, many of that observatory's greatest servicing successes were conducted on components that were never actually meant to be serviced. Agencies like DARPA and NASA have also conducted extensive technological demonstrations of servicing on non-compliant components (and this was described in greater detail above) (NASA Goddard Space Flight Center, 2010).

As evidenced by this report, it has largely been industry risk aversion and resistance to changing entrenched architectural paradigms that have hampered the wider proliferation of OOS activities beyond missions like Hubble. Indeed, spacecraft operators are no longer justified in citing costs or technological shortcomings in their aversion to OOS implementation, especially when these operations take place in the relatively accessible regions of LEO and geosynchronous orbit.

Furthermore, in describing the rationale for the need of OOS capabilities, Benedict describes how potential users have been hesitant to embrace this technology. Servicing missions including refueling, towing, rescue of satellites which have suffered launch shortfalls or orbital anomalies, and even the removal of orbital debris would of course be welcome, but the space industry is conservative and will not embrace on-orbit servicing until organizations like NASA and DARPA prove it is feasible. Once this has been accomplished (and many would say it has), demand for missions is expected to greatly increase. Indeed, supporters of a shift to OOS have argued that there are several cases where it may be less expensive and more efficient to service a satellite than to replace it. For example, if a commercial communications satellite at GEO requires an investment of $200-$400 million, most firms would pay a sum below or even equal to that initial investment in order to save a satellite from an otherwise mission-ending anomaly. The case for refueling or servicing well-functioning spacecraft is more difficult to sell, though, as firms often prefer replacement of the spacecraft with a more technologically capable one, rather than service of their current asset (Benedict, 2013).

In addition, Richards developed a methodology to assess serviceability across the four activities of rendezvous, acquire, access, and service, and he estimates that there are approximately 25 servicing opportunities for GEO spacecraft each year. Similar to Benedict, he

also finds that the main challenges to implementing more OOS are not technical but social and political—that is, the technology must be proven safe and reliable before it receives wider implementation, leading to a "chicken-and-egg" problem of maturing the technology when few wish to undertake the required funding and development. Designers have also become comfortable, if grudgingly so, with current design paradigms which are lacking a servicing option, and a very strong business case for servicing may have to first be presented, beyond NASA and DARPA technical demonstrations, before the wider industry will begin a paradigm shift towards OOS (Richards, 2006).

Next, according to Long et. al., on-orbit servicing of satellites appears to be technically feasible but may not yet be economically feasible, and a new value proposition for satellite architectures must therefore be developed. This new value proposition would be oriented towards rapid response to technological or market change and design of satellites with less redundancy; in other words, satellites must be designed with OOS explicitly in mind. Long et. al. find that 130 servicing tickets of various urgency could be expected to be generated over a 5 year period by the set of 335 GEO satellites, but that the economic feasibility of on-orbit servicing remains questionable. The current generation of spacecraft is too reliable to ensure market share for OOS, and a paradigm shift to GEO spacecraft with shorter lifetimes may be required. New value propositions must be incorporated into the design and operation of satellites, and OOS must be able to provide large performance gains over the long design lifetimes of satellites. Also, if customers begin to more heavily value flexibility and responsiveness in their space operations, then OOS could capitalize on the shifting paradigm (Long, Richards, & Hastings, 2007).

Orbital Communications and Data Relay Infrastructure

Communications and data relay between spacecraft is a crucially vital component of nearly all modern space operations, especially those which require the coordination of many orbital assets. Even more crucial is the ability to quickly and efficiently downlink spacecraft data to the users who need it on the ground; indeed, even the most advanced spacecraft instruments are next to useless if the data they collect is not transmitted when and where it is actually needed. To this end, an orbital communications and data relay infrastructure, one that allows customer spacecraft to link in and transmit data from spacecraft-to-spacecraft and from spacecraft-to-ground, has been proposed as a highly beneficial enhancement to new and existing space architectures. With such an infrastructure in place, it is believed that not only would customers be able to downlink and process more spacecraft data (making orbital assets even more valuable to customers worldwide), but entirely new ventures could be explored, ones which could never have succeeded without the data relay service provided by a global, widely-available resource-sharing infrastructure.

TDRSS and EDRS

Many of the earliest space missions clearly demonstrated a need for greater connectivity between spacecraft and the controllers on the ground. During these initial stages of space utilization, NASA depended on a large collection of international ground stations, which in addition to being unable to provide truly global coverage for most spacecraft could also be expensive to man and maintain, as well as remaining subject to the whims of the host nation. Thus, the Tracking and Data Relay Satellite System (TDRSS) project was established in 1973, with the first launches of TDRSS spacecraft occurring in the 1980's. NASA procures the spacecraft directly from manufacturers, and this constellation of geosynchronous satellites now require just three ground stations, two in the mainland US and one in Guam, to provide global, round-the-clock coverage to spacecraft operating in low-Earth orbit (Tracking and Data Relay Satellite Project, 2012). TDRSS has been operating continuously for over three decades and is relied upon heavily for human space operations, as well as to return significantly more data from science payloads than those spacecraft could return on their own (and often in near-real time, which is only possible for low-Earth orbit assets on the odd chance that the relevant data is being

collected during or just before a scheduled ground station pass). TDRSS, as well as other NASA communications assets, can also be used by other entities, public and private, to relay their own data for a relatively small fee and as permitted by NASA scheduling (NASA Goddard Space Flight Center, 2012). While TDRSS is an excellent system, providing outstanding service to NASA science missions and other NASA partners, its main drawback is that its services are in very high demand and as such it is unlikely to be able to aid many smaller commercial customers (especially newer entities which do not necessarily have the priority of more established players).

Also lending evidence to the ever-growing demand for timely, reliable communications and data relay capabilities is the ESA's European Data Relay System (EDRS). Its overall mission remains much the same as TDRSS, in that it seeks to greatly increase the amount of data which can be downlinked from non-geosynchronous space assets, as well as allow for significantly improved transmission of time-critical data to users on the ground. This planned constellation of geosynchronous satellites also hopes to surpass TDRSS's capabilities through the use of optical communications from geosynchronous orbit to dedicated ground stations, and its first payload is slated to be launched on board a host spacecraft in 2015 (European Space Agency, 2014).

While TDRSS and EDRS both operate in geosynchronous orbit, the infrastructure proposed and assessed in this thesis would be placed in low-Earth orbit. While this does not provide the very large, hemispheric visibility to the Earth seen by GEO assets, it does offer other benefits. For instance, transmission latency issues are much less significant in low-Earth orbit, simply because the distances between customer spacecraft to the infrastructure and from the infrastructure to the ground would be much less than the links to and from geosynchronous spacecraft (approximately 1,000 km in low-Earth orbit vs. over 35,000 km in geosynchronous). This large difference in transmission distance also means that a low-Earth orbit infrastructure could theoretically be composed of much smaller, less powerful, and simpler spacecraft than those that must be launched to GEO. Such a difference in spacecraft size and complexity could also potentially lead to decreased infrastructure costs, even if many more spacecraft are required for global coverage in LEO than in GEO. Furthermore, the fact the LEO infrastructure would have to be composed of many more spacecraft also lends a significant degree of resiliency to the overall infrastructure; i.e., a LEO infrastructure could theoretically experience failures in some of its spacecraft without significant degradation of the overall system's performance, while a GEO

constellation relying on a small number of large spacecraft would be far more damaged in the event of just one spacecraft failure. In summary, while there are definite benefits to placing a communications and data relay infrastructure in both GEO and LEO, it is believed that the case for a LEO infrastructure is one which has not been pursued as actively by space agencies and spacecraft operators, but which now is deserving of real consideration.

LEO Infrastructure Concepts

Along these lines of assessing emerging concepts in LEO infrastructures, Golkar et. al. have performed extensive research into the hosted, disaggregated space infrastructure concept within Golkar's Federated Satellite Systems (FSS) framework. The FSS framework integrates the study of technical feasibility, market assessment, and business case proposition validation of FSS terminals mounted as payloads on hosting spacecraft. FSS allows spacecraft to share in-orbit resources in an opportunistic fashion, with examples of in-orbit resources including processing power, link capacity, and data storage. Commercial implementations of the FSS concept are designed to create in-orbit markets of space resource commodities, and to enhance sustainability, cost-effectiveness, robustness, and reliability of participating missions. Also, one key test bed that Golkar identifies for the implementation of an orbital infrastructure is the International Space Station, as it has all the necessary power and hardware needed to test new infrastructure technologies (Golkar 2013). Optical hosted payloads are identified as one of Golkar's preferred solutions, as they provide simpler, lighter, and more effective data relay and processing systems when compared to RF. The ISS as a hosted payload hub could create a market of millions to hundreds of millions in Euros for optical communications and data processing, and potential markets of interest would be those LEO spacecraft in sun-synchronous orbits conducting Earth observation and telecommunications missions (which are highly pressed to return large amounts of low-latency data to customers on the ground). The FSS concept would also make use of the unused telecommunications capacity available in participating spacecraft at any given time.

The communications and data relay infrastructure considered in this thesis is slightly different from Golkar's concept, however, in that this analysis considers only free-flying, dedicated infrastructure spacecraft, rather than hosted payloads. The work of Golkar et. al. does

provide significant insight into how the resources of the infrastructure can be efficiently allocated amongst customers, and it would be expected that any well-designed communications infrastructure would follow the same principles for shared, assured, efficient connections between the infrastructure and its customers.

Finally, it is important to note that the time has never been better for the implementation of a LEO communications and data relay infrastructure. Investors and venture capitalists are increasingly bullish on technologies and enterprises based in LEO, with many of these new ventures, within such missions as space-based internet and optical imagery, being driven by the increasing miniaturization of technologies and the promise of lower launch costs. As Frank Morring Jr. writes in a recent article covering the booming LEO economy, there is an "insatiable demand for data communications and real-time information about our planet," and a data relay infrastructure could act as a major enhancement to many upcoming systems (Morring Jr., 2015). Indeed, if an infrastructure is implemented and found to offer significant value to customer spacecraft, it is likely that entirely new systems could then be designed specifically with the infrastructure in mind, much as terrestrial businesses can rely on the highways, airways, and sea lanes to move goods back on Earth.

Space Industry Trends

As with almost all modern industries, spacecraft technology has progressed significantly in the last two decades. One would expect that the cost and mass of the typical spacecraft would have decreased or at least remained mostly constant, even as the power and performance of each craft is improved. However, this has largely not been the case. Over the past two decades, the major trends in geosynchronous spacecraft manufacturing have been towards larger, more expensive, longer-lasting spacecraft. This is shown in Figure 1and Figure 2, which depict the design lifetimes and launch masses of U.S. civil, commercial, and military spacecraft launched to geosynchronous orbit over the last two decades (Union of Concerned Scientists, 2014). Similarly, it has been shown by Saleh that the cost of spacecraft also increases with design life, and that cost-per-operational-day estimates exhibit diminishing returns after a design life of approximately 8 years (shown in Figure 3 and Figure 4) (Saleh J. , 2008).

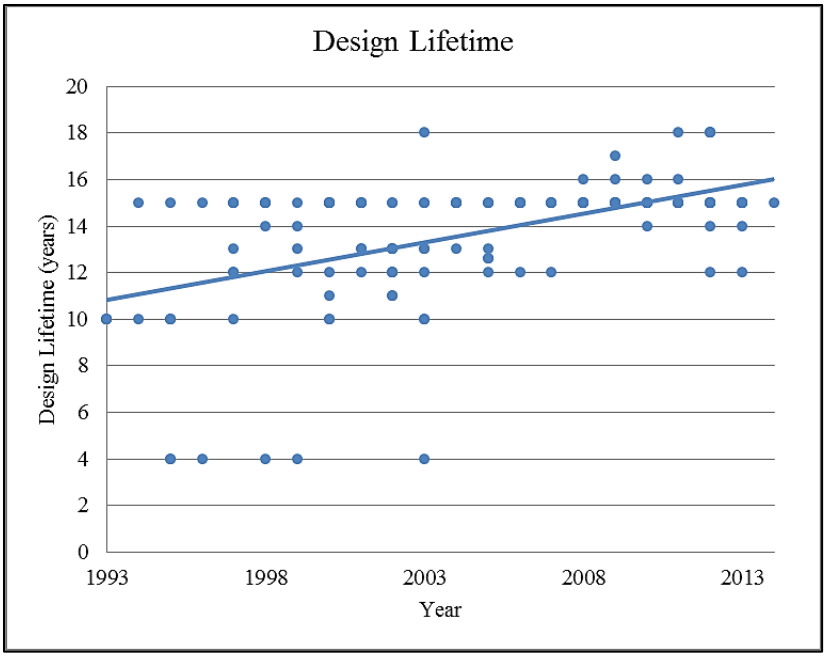


Figure 1. Design lifetimes of active geosynchronous spacecraft over the last two decades (UCS 2014).

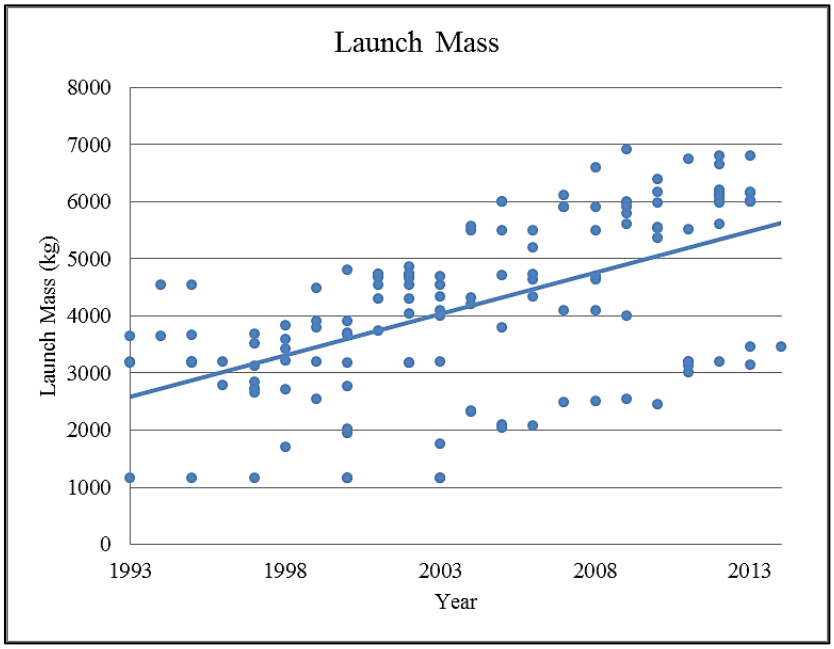


Figure 2. Launch masses of active geosynchronous spacecraft over the last two decades (UCS 2014).

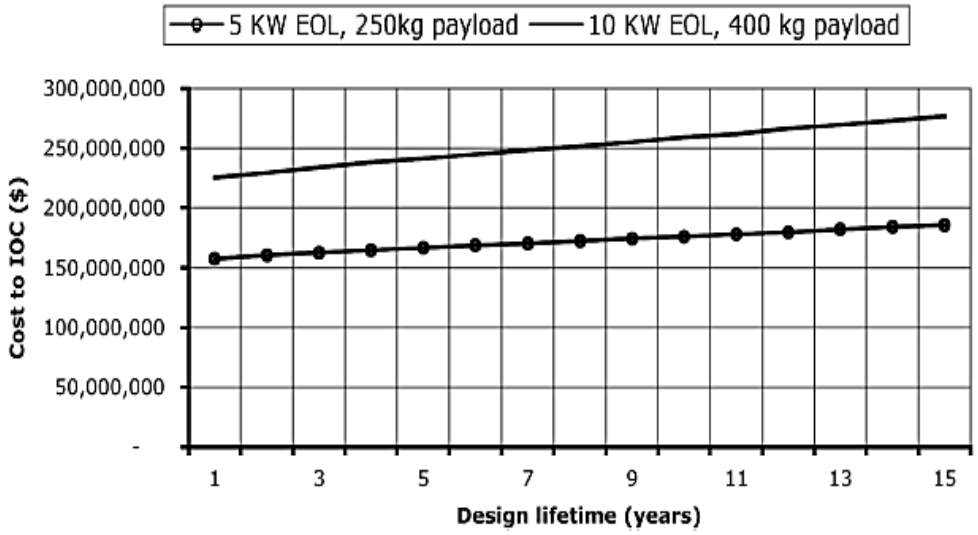


Figure 3. Cost to initial operating capability (IOC) as design lifetime is increased (Saleh 2008).

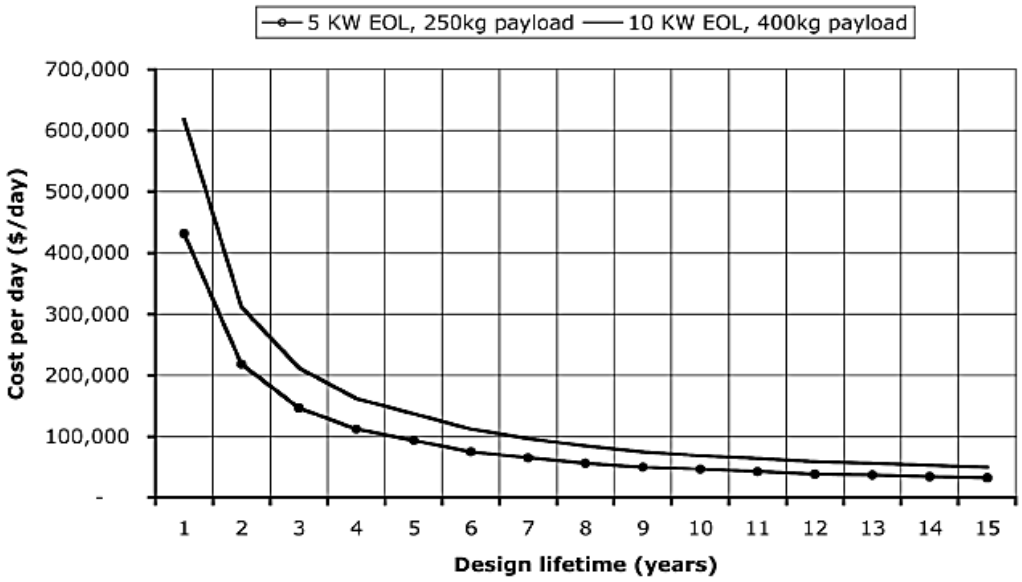


Figure 4. Cost per operational day of spacecraft as design lifetime is increased (Saleh 2008).

This is not to say that the performance of spacecraft has only held steady. In fact, the average power of spacecraft has been shown to fit an exponential growth curve as technology has progressed over the last several decades (shown in Figure 5). Using power and mass as

proxies for spacecraft performance, it is evident that spacecraft architecture designers have consistently opted for higher-performing, longer-lasting, increasingly massive space vehicles.

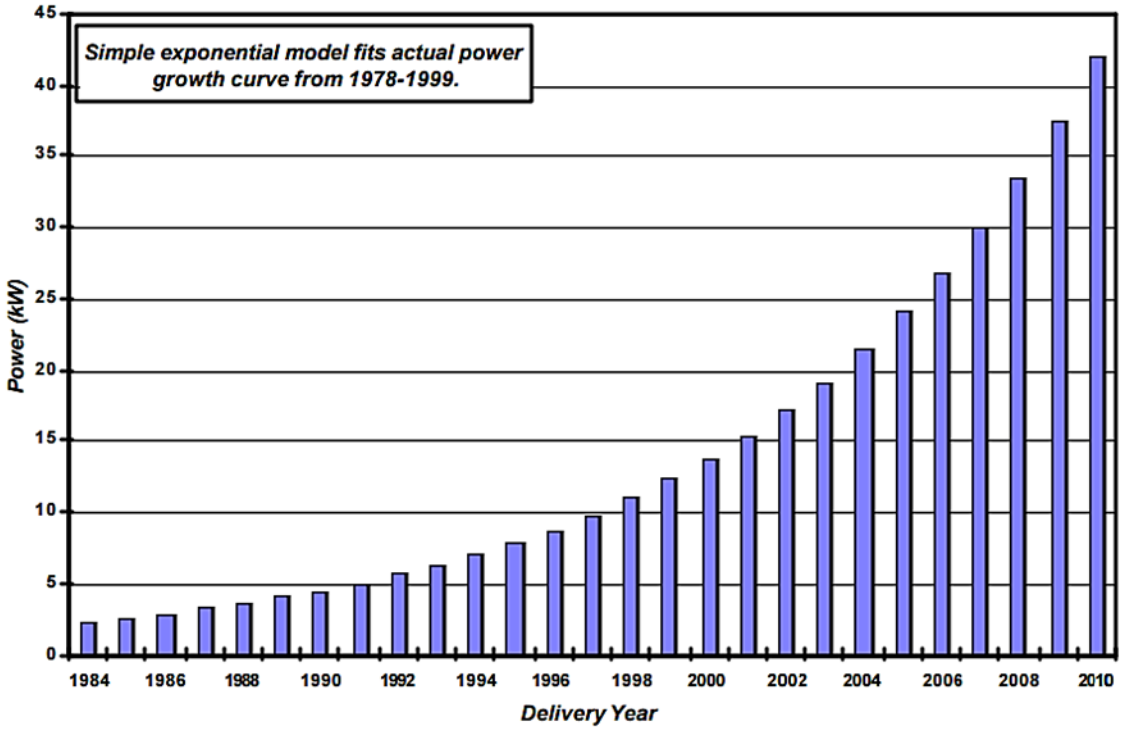


Figure 5. Increases in average power of geostationary spacecraft (Jones and Spence 2011).

Current spacecraft perform exquisitely and reliably, but cracks in the overall space industry are beginning to show. Launch costs have remained stubbornly high, even as the technology behind launch vehicles has matured, and the development schedules required to build, test, and deliver new spacecraft typically extend a decade or more. Meanwhile, cost and schedule overruns on new spacecraft are the norm, and the space industry as a whole has exhibited a high degree of consolidation and monopolization as firms chase larger and lengthier contracts (Wertz J. , 2011). Proponents of on-orbit servicing often argue that spacecraft design paradigms incorporating OOS would help to alleviate some of these industry stressors, allowing for new spacecraft to halt the steady increase in mass and design life in favor of more flexible designs—ones which would plan for periodic refueling and/or technology upgrades (Long, Richards, & Hastings, 2007). Indeed, in the methodology presented below, it is assumed that an effective

OOS program is one that can significantly alter long-term spacecraft design trends while also allowing for a more flexible, lower-cost industrial risk posture.

Space Policy

The field of space policy has experienced significant shifts over the past few decades, as space activity has moved from an almost exclusively government- and military-driven focus to today's increasingly commercial and democratized environment. In addition, ever since the first satellites were launched, space policy has necessarily been a global affair, as orbital dynamics dictate that any non-geosynchronous spacecraft will be constantly overflying other nations' territories. The major international treaty governing space activities is discussed below, in addition to the current space policy of the United States. The three major sectors of U.S. space activities—commercial, civil, and defense—are also discussed.

1967 Outer Space Treaty

From the very earliest days of space utilization, the United Nations has regarded space as a highly international arena. Perhaps the most important, foundational UN treaty—one which has guided the direction and development of nearly all nation's space activities for decades—is the 1967 Outer Space Treaty, also known as the "Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, including the Moon and Other Celestial Bodies." The major provisions of this treaty, which are relevant to this thesis, are discussed below:

- Article 1: Outer space is regarded as the "province of all mankind," and the exploration and use of outer space should be directed towards the interests of all countries. All areas of outer space, including the Moon and other bodies, are free to exploration and use without discrimination of any kind.
- Article II: No nation can claim sovereignty or territorial control of any part of outer space.

- Article III: All nations carrying out space activities must do so in the "interest of maintaining international peace and security and promoting international cooperation and understanding."
- Article IV: States should not place in orbit or on any celestial body any weapons of mass destruction, nor should nations test weapons or establish military fortifications in outer space. However, military personnel and facilities may be used for peaceful purposes.
- Article VI: States bear international responsibility for all national activities in outer space, regardless of whether they are carried out by government or non-governmental entities. In addition, each State must monitor and regulate the space activities of its non-governmental entities.
- Article VII: States which claim ownership of a launched object, in addition to the States from which the launch is procured (if those are not the same), are held internationally liable for any damage caused to another State or entity by any object launched.
- Article VIII: States retain jurisdiction and control over any and all objects they launch, or procure the launch of, into outer space, for as long as they remain in outer space (UN Office for Outer Space Affairs, 1967).

From this treaty it is clear that space is regarded as a common resource, open and accessible to all. Space activities are to always be conducted in a peaceful, non-intrusive manner, with the utmost care taken to ensure no harm is done to other nations' space assets. Weaponization of space is highly discouraged (with weapons of mass destruction being banned outright), and it is expected that nations will cooperate and aid one another in the conduct of space activities. Also very important is the way in which all nations are held responsible for the objects they launch into outer space—not just while those objects are actively engaged in space operations, but for the entire duration of their time in orbit. From this one can see the way in which an orbital infrastructure would have to be managed; i.e., even if an infrastructure provides services to many nations and entities, its ultimate ownership and liability for damages would still be held by the nation (or national entity) which launched it. Also, within the spirit of the Outer Space Treaty, a nation which develops and launches an orbital infrastructure project would also likely need to do so in as transparent a manner as possible, in order to make clear the peaceful purposes of the project.

Space Policy of the United States

The United States has been a leader in space activity since the very first spacecraft were launched into orbit, and as such maintains a robust, expansive space policy. The latest U.S. National Space Policy, released in 2010, lays out several guiding principles for the conduct of U.S. space operations, and these are discussed in detail below:

- All nations should act responsibly in their use of space, with a specific focus on openness, transparency, and sustainability within space operations. Nations should also work to prevent accidents and mistrust in the course of space activities, as well as help others enjoy the benefits of space utilization.
- Commercial activity in space is a crucial component of a healthy space industry, and as such the U.S. will encourage and support further growth of the U.S. commercial space sector. It is the goal of the U.S. to have a competitive, robust, and innovative commercial space sector that also helps to advance U.S. global technological leadership.
- The right of all nations to explore and utilize outer space for peaceful purposes is a vital component of international law. Furthermore, the U.S. regards "peaceful purposes" as including international and homeland security activities.
- The U.S. supports international law forbidding claims of national sovereignty and upholding the free access of all to space utilization. In addition, the U.S. regards any "purposeful interference" with spacecraft or supporting infrastructure as an infringement of nations' rights.
- The U.S. will work to assure space is open and accessible to all. Consistent with the right to self-defense, the U.S. will also seek to deter others from interference and attack, as well as defend its own and allied partners' space systems. Should deterrence fail, the U.S. reserves the right to actively defeat incoming threats (National Space Policy of the United States of America, 2010).

The main point from the National Space Policy relevant to this thesis is that the U.S. seeks to heavily leverage the capabilities of commercial space players to achieve its objectives in space, rather than rely solely on government agencies to deliver space-based capabilities. Indeed, a robust U.S. commercial space sector is considered to be a national asset, one which can be called upon to deliver innovative, cost-efficient solutions to satisfy new and existing space mission

requirements. Thus, it is assumed that a commercial or partially-commercial solution for the implementation of an orbital infrastructure project would be heavily preferred, although almost certainly some amount of governmental oversight would still be required.

In addition, the commitment of the U.S. to an open, transparent, accessible space environment should grant a degree of confidence to international partners hoping to utilize the services provided by an orbital infrastructure. For instance, one of the political hindrances to on-orbit servicing is the fact that a spacecraft which can repair or manipulate the components of a satellite would also likely be capable of intentionally damaging or disabling that satellite. However, the U.S., by virtue of its national space policy and its past record of being a global leader in open space operations, could be trusted enough by international partners to implement such a project and extend its services to a wide range of spacecraft. Thus, the open, transparent nature of U.S. space policy could actually have a strong supportive effect on the long-term development of an internationally-accessible orbital infrastructure.

Space Industry Sectors

While the three major sectors of the U.S. space industry—commercial, civil, and defense—are all closely linked and share many of the same technologies and architectural paradigms, they also all have relatively unique requirements and capabilities, stemming from their sometimes-disparate missions and stakeholder expectations. For commercial entities, the main driver of satellite and space architecture design is profit maximization and reliable service to the customer. The primary commercial space mission is telecommunications and global data connectivity, with many of these enterprises having their origins in civil or military communications architectures which were subsequently spun off into private firms. Indeed, many commercial space entities are also strongly linked to governments and public agencies, providing launch services, communications, some on-orbit infrastructure, etc., and thus they may be strongly reliant on government contracts or subsidies. Furthermore, with the U.S. national space policy emphasizing an increased reliance on the commercial space sector, it is likely that new players in the space industry will continue to look to government and/or military as an initial and/or primary customer for new enterprises. As commercial space matures and technologies advance,

though, many private firms are beginning to look beyond the provision of services to governments or large agencies, and instead are exploring new opportunities in missions like global internet services and terrestrial imagery to provide services to a much wider range of consumers (Figliola, Behrens, & Morgan, 2006).

Civil space systems typically lack the explicit need to turn a profit on their space activities, but they are still held accountable by policymakers and taxpayers' representatives in government to provide useful returns on investment. Therefore, civil space architectures are often engaged in missions such as scientific exploration, Earth science, and data relay between orbital assets; i.e., those activities which may not be profitable enough to be provided by commercial entities, but which are still valued for their scientific findings, usefulness to other missions, or simply greater understanding of our planet and the overall space environment. Civil systems, such as those launched and operated by agencies like NASA and NOAA, are also tasked with providing crucial services to both large organizations and individuals, such as space-based weather systems and terrestrial imaging. Furthermore, civil space architectures may be used to test new technologies and operational concepts for overall viability before transitioning successful projects or capabilities over to commercial users (Figliola, Behrens, & Morgan, 2006).

Finally, the defense space sector shares several similarities with civil systems, but with the important distinction that defense systems are focused on aiding in military operations. Defense space systems typically have an additional desire for increased survivability and resiliency against a much larger spectrum of threats, beyond those faced by all spacecraft in the harsh space environment, and this can often lead to greater overall requirements from defense architectures than from commercial or civil. It is primarily along these lines that a new thrust for increased resiliency and disaggregation of space architectures has emerged, especially from the U.S. Air Force Space Command, as a means of increasing overall survivability, responsiveness, and flexiblity while also decreasing costs (U.S. Air Force Space Command, 2013). In addition, defense systems oftentimes do have useful civil applications (GPS is a primary example of this); however, it is usually only after a defense system has satisfied its military requirements that its potential civil applications can be explored. Furthermore, missions in the defense sector are typically viewed as vital to national security, thereby increasing overall architecture needs and costs but also allowing for higher final price tags than may be acceptable in the civil or

commercial sectors (Figliola, Behrens, & Morgan, 2006). It is worth noting, however, that many of the most recent defense architectures have been characterized by increasing costs and lengthening development cycles, and because of this there is now greater focus on responsiveness and increased cooperation with other entities—civil and commercial—in an effort to drive down costs and development lags; for instance, through the hosting of less vital missions on board commercial satellite busses (Office of the Under Secretary of Defense for Acquisition, May 2003). Finally, much like civil systems, defense architectures sometimes act as testbeds for emerging space technologies, which can then be spun off for use by civil or commercial entities.

Methodology

The methodologies chosen for the analysis in this thesis follow directly from the research questions posed above. For a LEO communications and data relay infrastructure, optimal solutions were sought for constellation and spacecraft design, which will render the greatest benefits to a customer spacecraft. Insights as to the best policy regimes for infrastructure implementation can then follow directly from this estimation of architecture size, key design points, and costs. To this end, a robust physics and cost model was developed for the evaluation of a tradespace of LEO data relay infrastructure designs, and the specifics of these models are described in much greater detail below.

Then, to inform the discussion of OOS policy options, a system dynamics model was used to track the long-term trends in spacecraft design and industry health after the implementation of servicing operations. Unlike the LEO data relay infrastructure case, the research questions posed did not call for the design of a specific system, but instead required an evaluation of overall system trends in cost, design, and industry experience. Since OOS is described as a means of significantly shifting architectural paradigms (DARPA, 2014), system dynamics modeling can act as a useful methodology for the investigation of which policies will most likely effect these shifts over a long time horizon. Again, the specifics of this methodology are discussed in much greater detail below.

Assessing Designs for a LEO Communications and Data Relay Infrastructure

Modeling the positions and velocities of spacecraft as they orbit the Earth has long been a vital component to the design and evaluation of new and existing missions. Such models are crucial to understanding orbital parameters such as altitude, distances between communications links, ground coverage, and revisit times, in addition to scheduling satellite passes over ground stations. The equations governing orbital mechanics are complex and time-intensive for a human to solve, but when implemented in computer code they can be executed over and over again many times a second, allowing a user to continuously track the path and performance of a satellite or constellation. Indeed, many agencies and firms involved in space operations maintain

their own in-house orbital modeling suites, while commercially available software packages like Satellite Toolkit (STK) allow any user access to a high-fidelity, highly validated orbital propagator.

Then, the orbital parameters of a satellite or constellation, along with the performance requirements of the space mission, directly inform satellite design choices such as transmitter power, battery sizing, fuel requirements, and data storage needs. Again, many sources are available to aid space mission architects with the design and performance specifications of satellites, with one of the most well-known and reputable sources being Microcosm's *Space Mission Engineering: The New SMAD* (Wertz, Everett, & Puschell, 2011). This text is used extensively to inform the analysis conducted in this thesis, as described in further detail below. Once orbit and design parameters have been determined, overall costs can then be estimated using statistics-based cost models, which use the costs of past space missions as baseline estimates for future missions.

All of these mission parameters—orbitology, satellite designs, and costs—come together to form one cohesive architectural design, and while some space missions require highly specific designs, many can be accomplished through a range of architecture configurations. Thus, in addition to simply selecting the performance requirements of a satellite or constellation, the space mission designer must also evaluate the costs and trade-offs between various designs. For instance, one large, very powerful optical imagery spacecraft could likely downlink a huge amount of high-resolution data with each pass over a ground station, but a constellation of smaller, less powerful satellites with lower resolution may actually be preferred if the designer values lower latency and greater ground coverage. Similarly, the design of a communications and data relay infrastructure encompasses a huge variety of trade-offs to achieve its mission of connecting customer spacecraft and downlinking data, and the process for determining the most optimal designs is described in detail in the sections below.

Analysis Framework and Utility Function

In order to assess various designs for a communications and data relay infrastructure, a widely applicable reference customer constellation must first be determined. Of course, a

potential infrastructure would aid a variety of customers in downlinking their spacecraft data to users on the ground, but it was determined that satellite imagery is one mission in particular which would benefit greatly from the ability to very quickly downlink freshly collected data. Because optical imagery constellations were singled out for analysis, new commercial firms such as Skybox and Planet Labs were considered to be primary customers for utilizing the resources provided by the infrastructure. With this in mind, a utility function was developed which captured the value of various optical imagery constellation designs, with and without the downlinking capabilities of a publicly-available infrastructure. This value function included both a spacecraft's imaging capabilities as well as the constellation's ability to downlink collected data quickly and efficiently.

The first five value metrics identified for analyzing orbital imagery constellation designs were average resolution, peak resolution, time in view over a target, mean gap time between passes over a target, and percentage of global coverage. These metrics are specifically related to the design and orbitology of the imaging spacecraft, and must be considered in the design of any such system. The next four metrics are related to the constellation's ability to return imagery data to the ground: percentage daily contact time with ground stations, the daily percentage of collected data which can actually be transmitted to the ground, the average outage time between downlink passes, and the maximum outage time between downlink passes. These four metrics are those which can be augmented by infrastructure, and the value of the many possible infrastructure configurations must be measured in conjunction with that of the optical imagery constellation designs (i.e., the infrastructure provides an improvement in downlink performance and interconnectivity for customer spacecraft, but this improvement must be measured relative to stand-alone performance). With this in mind, the value metrics for an infrastructure's utility are a combination of both the infrastructure's performance and its performance in relation to a customer constellation, and these metrics are listed below:

1. Percentage daily contact time with the receiving ground stations;
    - A maximum percentage of 100% is assigned a value of 1, while 0% is assigned a value of 0, with other values assigned linearly in between.

2. Normalized average outage time between passes of the infrastructure over the ground stations (or LEO-to-ground latency);
    - An average outage of 0 minutes is assigned a value of 1, while an average of 450 minutes or greater is assigned a value of 0.
3. Normalized maximum outage time between passes (another measure of LEO-to-ground latency);
    - A maximum outage of 0 minutes is assigned a value of 1, while a maximum outage of 900 minutes or greater is assigned a value of 0.
4. Normalized total amount of time per day that the infrastructure is actually transmitting data to the ground;
    - A maximum of 24 hours of contact a day is assigned a value of 1, while a minimum of 60 minutes or less is assigned a value of 0.
5. Normalized maximum outage time between the infrastructure and a participating spacecraft;
    - A maximum outage of 0 minutes is assigned a value of 1, while a maximum outage of 90 minutes or greater is assigned a value of 0.
6. Normalized average outage time between the infrastructure and a participating spacecraft (or LEO-to-LEO communications latency);
    - An average outage of 0 minutes is assigned a value of 1, while an average outage of 45 minutes or greater is assigned a value of 0.
7. Percentage of "priority" data which can be transmitted each day; that is, how much data can be sent from a participating spacecraft, through the infrastructure, to the ground, in real-time as the data is being collected;
    - A maximum percentage of 100% is assigned a value of 1, while 0% is assigned a value of 0.
8. Percentage of data received from a participating spacecraft, which is then actually downlinked to a ground station;
    - A maximum percentage of 100% is assigned a value of 1, while 0% is assigned a value of 0.

9. Total combined daily percentage of contact time of both the infrastructure and a participating spacecraft with receiving ground stations.

- A maximum percentage of 100% is assigned a value of 1, while 0% is assigned a value of 0.

The utility function was designed to be as straightforward as possible, with each of the nine metrics receiving an equal weighting of one-ninth. The means of determining the "score" for each metric, or the assigned value between 0 and 1 determined for each metric, was also designed to be straightforward, with each score extrapolated linearly between predetermined maximum and minimum performance values (listed beneath each metric above). The utility function can also be written as shown below:

$$
\begin{aligned}
Design\ Utility\ = \frac{1}{9}\ [\ &(\%\ contact\ time\ with\ ground\ stations) \\
&+\ (Normalized\ average\ outage\ time\ with\ ground\ stations) \\
&+\ (Normalized\ maximum\ outage\ time\ with\ ground\ stations) \\
&+\ (Normalized\ total\ contact\ time\ per\ day\ between\ infrastructure\ and\ ground) \\
&+\ (Normalized\ maximum\ outage\ time\ between\ infrastructure\ and\ customer\ spacecraft) \\
&+\ (Normalized\ average\ outage\ time\ between\ infrastructure\ and\ customer\ spacecraft) \\
&+\ (\%\ of\ "priority"\ data\ transmitted\ each\ day) \\
&+\ (\%\ of\ data\ downlinked\ from\ customer\ to\ ground\ per\ day) \\
&+\ (\%\ of\ combined\ daily\ contact\ time\ with\ ground\ stations)\ ]
\end{aligned}
$$

Equation 1. Utility function used to compare utility of infrastructure designs. The "score" for each metric is a value between 0 and 1, extrapolated linearly from maximum and minimum values for each metric, and each metric is assigned an equal weighting in the final calculation of design utility.

The ultimate goal of the analysis was to produce optimal infrastructure designs that maximize the connections between customer spacecraft and the infrastructure, thereby increasing connectivity between customer spacecraft and ground stations. Furthermore, the inclusion of a "priority" data metric was essential for evaluating the infrastructure's overall value to

participating systems. It is believed that customers of an infrastructure would not only wish to downlink more overall data than is possible with just their own spacecraft; they would also desire the capability to immediately downlink data being collected in real-time, often from an imaging spacecraft that is many minutes or hours away from its next ground station pass. In order to do this, the imaging spacecraft must be in contact with an infrastructure spacecraft, and then at least one infrastructure craft must be in contact with a ground station, such that data can be sent immediately from the imaging spacecraft, through the infrastructure, and to the customer on the ground. This "priority" data metric represents a major improvement in system capability, one that can only be achieved through infrastructure.

Physics and Cost Modeling

With the analysis framework and utility function thus conceptualized, the next step for analyzing an infrastructure's performance and characteristics was through the construction of a physics-based orbital propagator, executed through MATLAB. This model incorporated the rotation of the Earth, the division between night and day, and oblateness. The script executed one-minute time steps, solving for a satellite's or a constellation's position and velocity over the Earth after each step. The model was assumed to cover a simple 24-hour period, on the Julian date represented by 1 January 2014. This was tested and found to be a large enough time scale to analyze the performance of low Earth orbit spacecraft (which orbit many times a day), in addition to easing the computational requirements for analysis. The chance that a particular orbit is a "worst-case" orbit for the date chosen also becomes increasingly insignificant as larger LEO constellations are analyzed; in other words, the larger number of spacecraft circling the Earth with respect to the unchanging location of the ground stations negates the irregularity or unsuitability of any particular orbit. It should also be noted that while orbits were all assumed to be "circular," this does not equate to perfect circles (due to the inclusion of Earth oblateness and perturbation). This propagator's performance was validated against several sample cases of Satellite Toolkit (STK) simulations, and was shown to achieve the necessary fidelity and resolution in its results to provide for effective analysis (Systems Toolkit (STK) version 10, 2015).

A key module added to the orbital propagator, especially for the analysis conducted for this thesis, was one that examined radio frequency link closure. It allows the modeler to specify a data rate and a maximum allowable communications range between spacecraft in the infrastructure or between spacecraft and the ground (in Mbps), and then returned the required power to achieve link closure (assuming a directional antenna with an average level of encoding and line losses). All calculations involving data rate and power requirements assumed line-of-sight communications only, as well as 50% average cloud cover when communicating with ground stations. Furthermore, a major parameter for this module was the fact that transceivers were sized such that they could transmit all data collected at 100% duty cycle; that is, the transceivers are sized as if they will run continuously. Even though this is not likely to be the case in real operations, it does ensure that spacecraft power and mass estimates are as conservative as possible. This also represents the over-engineering that would likely go into a real infrastructure project, one that is expected to service a large number of customers daily. The power and size of the transmitters required were then fed into an overall spacecraft sizing module, which provided final figures on total spacecraft mass and power requirements. This then led directly to launch vehicle selection and the determination of total on-board fuel requirements. Mass, power, and link budget closure estimates were based on several resources provided in Space Mission Engineering: The New SMAD (Wertz, Everett, & Puschell, 2011). Refer to Figure 6 and Figure 7 below for graphical representations of calculated power requirements for communication with the infrastructure and increases in daily downlinked data through use of the infrastructure. These calculations directly influenced total spacecraft mass, as higher power and data rate requirements lead to larger batteries and solar panel areas.

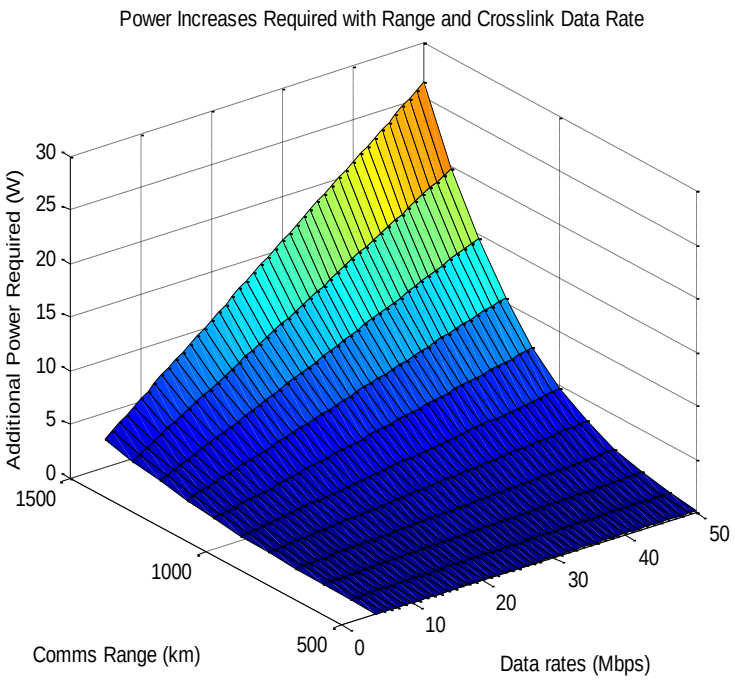


Figure 6. Additional power required of Landsat-8 in order to communicate with the infrastructure over a range of distances and data rates.

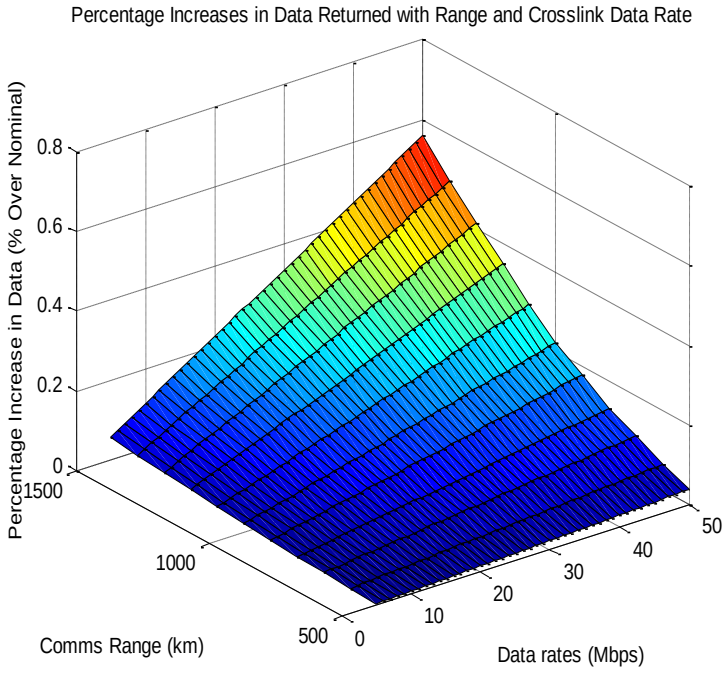


Figure 7. Percentage increases in data returned per day by Landsat-8 with the infrastructure over a range of distances and data rates.

The modeling of the infrastructure spacecraft required a few key parameters and assumptions. First, each infrastructure spacecraft was assumed to be able to link with other infrastructure craft that are positioned both in-front-of and behind itself within the same plane, other infrastructure craft from a different plane that comes within its range (e.g. a craft on an equatorial plane that passes by a descending or ascending infrastructure craft on a polar plane), as well as to two additional customer spacecraft anywhere within range. It must be noted that the inter-infrastructure links were sized differently than those of crosslinks with other spacecraft; that is, it was assumed that inter-infrastructure links were assured but may have had to cover much greater distances, while crosslinks with other spacecraft would have a shorter allowable range but could potentially be of a higher data rate. The different designs required for these two different requirements would almost certainly exclude one transmitter/receiver for all links, and as such this was accounted for within the infrastructure spacecraft sizing module. Additionally, each infrastructure spacecraft would always be capable of communicating with any visible ground station, subject to a minimum elevation angle of 10 degrees (the same constraint was also applied to the customer spacecraft communicating directly with the ground).

Once the designs and orbits for both the optical imagery craft and the infrastructure were determined, the next step was to determine how much all of it would cost (the infrastructure itself plus any modifications to the customer spacecraft required to participate in the infrastructure). Cost modeling for spacecraft was conducted using several resources: Space Mission Engineering, the Aerospace Corporation's Small Satellite Cost Model (SSCM), and the US Air Force Space and Missile Systems Center's Unmanned Space Vehicle Cost Model (USCM) (Wertz, Everett, & Puschell, 2011). The cost model developed was a bottoms-up cost analysis of the entire space architecture, including development costs, programmatic costs, IA&T costs, reproduction costs, launch costs, and ground operational costs. There were also over 50 cost and performance metrics included in the final determination of each individual spacecraft's cost, which were tied directly to the spacecraft sizing and design carried out in the physics model. Wherever possible, conservative estimates were used, and these were compared to actual space missions. The performance of the cost model is discussed further in the validation example. The cost of the required launch vehicles was also added to overall program costs, with the cost modeling of launch vehicles assuming best estimates for several of the most commonly

used launchers. Also, the selection of launcher for each infrastructure configuration was based on the total combined infrastructure mass, with aggregated launch constrained to a maximum of six satellites per launcher (this constraint is a conservative estimate, based on current technology and launch practices; for instance, the IridiumNEXT constellation is slated to launch 10 satellites at a time on SpaceX's Falcon 9 rockets (de Selding, 2014)).

In addition, there are several key assumptions underlying the modeling and analysis conducted in this paper. First, a standard communications scheme was assumed to be established between all participating spacecraft and the infrastructure, which does not conflict with other missions or space enterprises. Second, for each infrastructure configuration, it was required that there be enough satellites with adequately-sized transmitters such that global crosslinks are always maintained within the infrastructure itself; in other words, each plane of an infrastructure constellation must be a permanently closed and connected network. This could be achieved either through more powerful transmitters or by adding more infrastructure craft to each orbital plane (thereby decreasing crosslink distance), and these trade-offs were explored in the tradespace shown in the next section. It was also assumed that the infrastructure system would self-correct itself in the event of losing a satellite in the infrastructure constellation; for example, one way this could be done would be by having the remaining spacecraft maneuver and realign such that they return to equidistant spacing about their orbital plane. Achieving self-correcting capability requires the infrastructure spacecraft to always have the capability of bridging the communications gap in the event of a failure; in order to accommodate that requirement, the model assumed over-engineering of the power and RF link requirements. Incidentally, this assumption also enhances the infrastructure's ability to communicate with another plane, because of the greater power and link closure capabilities of each spacecraft (which is more in line with the goals of a robust infrastructure anyway).

With reference to downlinks and ground stations, established NASA ground stations were considered the primary means for the infrastructure to downlink data. The list of ground stations included those at Fairbanks, AK and Wallops, VA, in addition to those which NASA operates in partnership with other nations and agencies, such as the stations in Norway, Sweden, Germany, and at McMurdo Station in Antarctica (U.S. Geological Survey, 2014), (Mai, 2014). These five

ground stations, spread globally, were included in each analysis, except for those cases in which the infrastructure was tested specifically for its ability to replace downed ground stations. It was also assumed that the infrastructure and its participating spacecraft would be linked to the same ground stations, and that all participating ground stations would have the required capacity to service all spacecraft in view. A map of NASA ground stations, which includes the locations used in the analysis, is shown in Figure 8.

Figure 8. The locations of actual NASA ground stations were used to determine the infrastructure designs' downlinking capabilities **(**Mai**,** 2014**)**.

Also assumed during physics modeling was the systems' inherent ability to mitigate the significant Doppler shift effects that occur between any systems traveling at orbital speeds. While not a trivial problem (e.g., up to +/- 100 kHz in S-band), a review of literature on the topic suggests that a synchronization penalty preceding links between spacecraft and the infrastructure was sufficient to account for any Doppler effects (Vilar & Austin, 1991). That is, when a spacecraft first comes within communications range of the infrastructure, a time penalty of two minutes was forced before the model can actually begin to count the time in contact. The

forced penalty also holds true when the spacecraft is visible to more than one infrastructure payload. For example, before switching between payloads, the two minute penalty was still accounted for while making the switch (i.e. time in contact was not counted if the new payload being switched to had not already been visible for at least two minutes). The fact that this analysis only considers point-to-point communications between spacecraft, and not the more complex problem of multiple payloads communicating with multiple other crafts simultaneously, means that the assumption of a solved Doppler shift problem is almost certainly valid. It is also worth noting that this assumption forces much larger payloads overall, further reinforcing the conservative nature of spacecraft size and cost estimates.

Tradespace Exploration

With the orbital propagator and the associated modules developed, all of this data could then be entered into a tradespace exploration module. An extensive tradespace of over 10,000 infrastructure configurations was assembled, varying infrastructure design parameters such as altitude, inclination, number of planes, number of satellites per plane, and data rate for both crosslinks and inter-infrastructure links. Tradespace exploration is an increasingly vital tool for systems engineers, as it allows a huge variety of system designs to be evaluated and compared, based upon a common utility function (Ross & Hastings, 2005). Often, design tradespaces number in the thousands of designs, and while there are inevitably many inadequate system designs, a "Pareto front" of optimal designs is typically identified. This front contains the designs with the highest performance-to-cost ratios, with the most optimal points centered around the "knee" of the front (or the point at which there begin to be diminishing increases in performance with continued increases in cost).

The designs in this tradespace were analyzed using the utility function for the combined customer spacecraft-infrastructure system described previously, with each of the nine metrics being an equally-weighted component of an overall final performance score. Cost was also calculated for each design, such that a final tradespace of designs, plotted with utility vs. cost, was ultimately delivered and analyzed to find the most Pareto-optimal infrastructure designs (i.e. those designs which maximize utility while also minimizing cost). The full tradespace of infrastructure designs alone (without any association to a participating imagery constellation) is

shown in Figure 9. This tradespace is explored much more extensively in Case Study #2 below, with a much heavier focus on the performance increases experienced by varying customer types. In addition, graphical representations of a small selection of infrastructure designs are shown in Figure 10. These are provided in order to show the reader just how extensive the tradespace of designs actually is, and how widely different were the various designs considered. Also, refer back to Equation 1 above for the utility function derived to compute the utility of the designs displayed in the tradespace below.

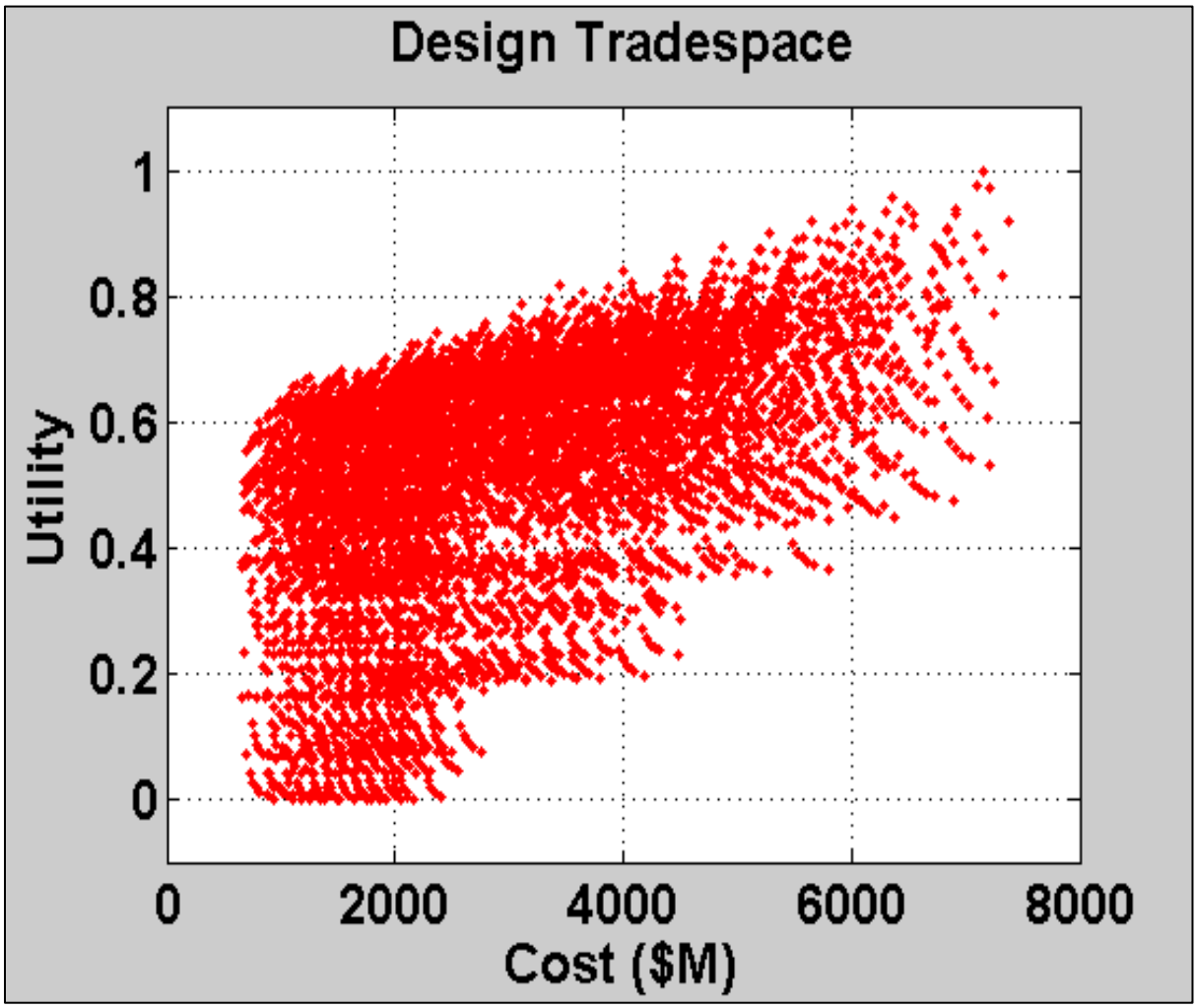


Figure 9. Initial tradespace of infrastructure designs.

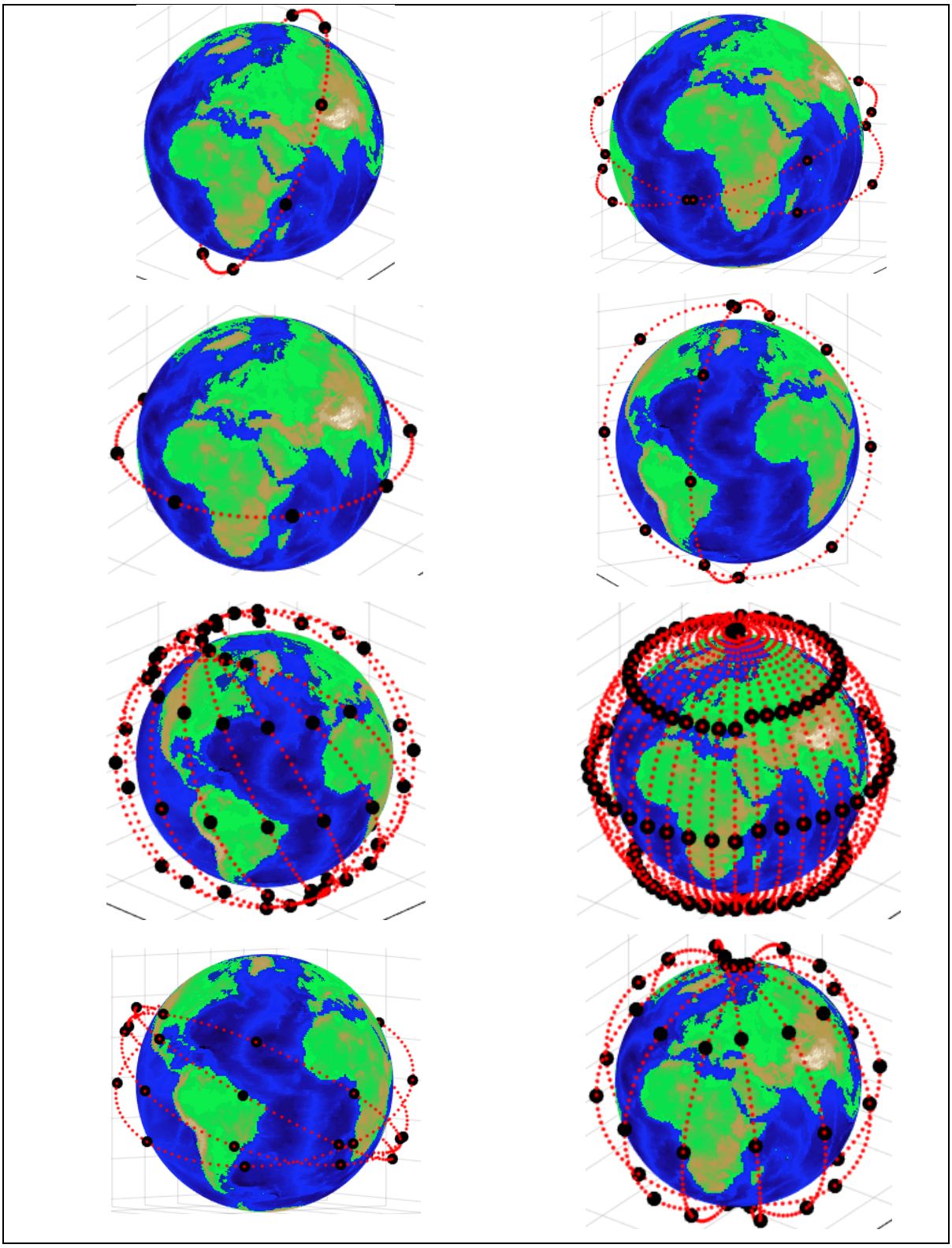


Figure 10. Graphical representations of a small selection of infrastructure configurations, taken from the tradespace of over 10,000 possible configurations.

The infrastructure tradespace shown in Figure 9 considers only each design's ability to transmit data to the ground, using a singular reference spacecraft in Landsat-8's orbit. In other words, the tradespace reflects maximum performance of the designs as a simple downlink architecture, from one spacecraft through the infrastructure to the ground, without the additional requirement of servicing any other missions or performing earth observation functions. As a first result (computed for the validation section discussed further below), it was found that under the utility and cost functions developed for this thesis, one of the most optimal infrastructure configurations was a design of 16 satellites divided equally into 2 planes in a polar orbit, with relatively small crosslinks of 10 Mbps. Based on the original utility function, which sought to maximize global coverage and ground contact time while minimizing outage times, this is indeed a very efficient infrastructure. However, it has not taken any specific customer requirements into consideration and as such may not actually be of use to a wider range of customer spacecraft. To accommodate this, a mask was later added to the tradespace analysis module, which forced out any designs that did not meet minimum customer requirements. This mask was based on the percentage increase of data downlinked over the customer's baseline, such that if a design does not meet a certain minimum percent increase (e.g. 20% over the baseline downlinked data total), then that design's utility is automatically reduced to zero. Designs remaining after the application of the minimum performance mask, which exhibited decreases in maximum outage and average outage, as well as increases in the overall amount of data that is downlinked each day, were also given higher utility scores.

The tradespace of infrastructure configurations shown in Figure 9 is informative, and it served as a useful tool in the overall validation of this thesis' methodology, but ultimately an infrastructure must be evaluated according to the value it can offer to actual space missions. Thus, to complete the analysis, a tradespace of infrastructure utility was developed for three major cases: a highly disaggregated constellation of very small imagery spacecraft (such as Planet Labs' Flock 1), a moderately disaggregated constellation of small-to-medium sized imagery spacecraft (such as SkyBox's SkySats), and a large monolithic imagery spacecraft (such as NASA's Landsat-8) [(Niles, 2014), (Skybox Imaging, 2015), (U.S. Geological Survey, Landsat-8 History, 2014)]. While the utility function works differently for each of these cases, and thus the results cannot be compared directly, it remains highly useful as a tool for analyzing

infrastructure designs and finding the "best fit" infrastructure for the largest variety of customers. The output of this tradespace analysis is presented in Case Study 2 of the Results section.

Validation of Physics and Cost Models

As a means of validating the physics and cost models described above, an example case was constructed using NASA's Landsat-8 as a template. Landsat-8, formerly known as the Landsat Data Continuity Mission, is an optical and thermal imagery satellite which continues the Landsat program's mission of providing scientists with global observations of land use and land cover change. Landsat-8 was launched on February 11, 2013 into a sun-synchronous, 98.2° inclination orbit at an altitude of 705 km. Its main payload, the Operational Land Imager, collects at least 400 images a day, with a panchromatic resolution of 15 m and a multispectral resolution of 30 m (U.S. Geological Survey, 2014). Landsat-8's orbital characteristics and technical requirements (such as resolution and data downlinked per day) were used as inputs to the orbit propagator, optical payload module, RF link module, and cost model. In addition, the locations of the actual Landsat-8 ground stations were used to determine daily access times and overhead pass information (refer to Figure 8 for a map of the ground station locations) [(U.S. Geological Survey, 2014) (Mai, 2014)]. A comparison between Landsat-8's actual specifications and those returned by the models are shown in Table 1.

LandSat-8	Actual	No Infrastructure	Maximum Infrastructure Case
Altitude (km)	705	705	705
Inclination (deg)	98.2	98.2	98.2
Number of Sats	1	1	1
Number of Planes	1	1	1
Crosslink Data (Mbps)	0	0	50
Crosslink Power (W)	0	0	14.7
Total SV Power (W)	1006 (max)	697.5	712.2
Solar Area (m^2)	3.6	3.75	5.02
Downlink Speed (Mbps)	300	278	278 Landsat (+50 Infrastructure)
Total Data (Mbits)	3200000	3200300	3200300
Dry Mass (kg)	1512	1383	1459
Wet Mass (kg)	2623	2459	2594
Fuel Mass (kg)	386	373	394
Bus Mass (kg)	421	854	905
Payload Mass (kg)	450	529	554
Bus Cost ($M)	194	119	123
Payload Cost ($M)	348	267	292
Total Program Cost ($)	930	703	743

Table 1. Comparison between actual and calculated specifications for Landsat-8.

While not exact, it is evident from Table 1that the model did return suitable estimates of the specifications for a Landsat-8 type spacecraft. This model is meant to return required parameters for an optical imagery spacecraft, and while conservative formulations were used in every instance possible during model development, it is still unlikely that the models would capture the significant factors of safety that a high-priority NASA mission like Landsat would certainly entail. For instance, the actual Landsat-8 spacecraft has higher mass and power than which was sized in the model, which was due to the increased redundancy and over-engineering of the actual spacecraft. Also, it is shown that the estimate for bus mass was well off from the data

provided for Landsat-8. It was determined that this was due to the fact the data for Landsat-8 reflected the stripped-down, absolute minimum value for its spacecraft bus, while the model returned a value which included much more of the spacecraft's essential components (such as batteries, solar panels, piping, and similar hardware). The fact that the model's final values for dry and wet spacecraft mass were relatively similar to the real world data also provided assurance that its results are reliable. Furthermore, the cost model consistently underestimated the actual costs of the Landsat mission (though not enough to indicate significant flaws), and this is likely due to the cost overruns endemic amongst modern space missions.

To further test the model's capabilities, data was generated for a hypothetical case of a Landsat-8 type spacecraft being supported by an infrastructure. This validation-case infrastructure was sized as a 36 satellite constellation, with six satellites equally spaced in six planes at a 63.5 degree inclination (shown graphically in Figure 11 below). The increase in power requirement and data downlink rate were calculated over a range of allowable communications distances and infrastructure data rates (500-1500 km and 5-50 Mbps, respectively). The rightmost column of Table 1 shows the calculated data related to the "maximum infrastructure" case, in which Landsat-8 is designed to communicate with the infrastructure at 1500 km and 50 Mbps.

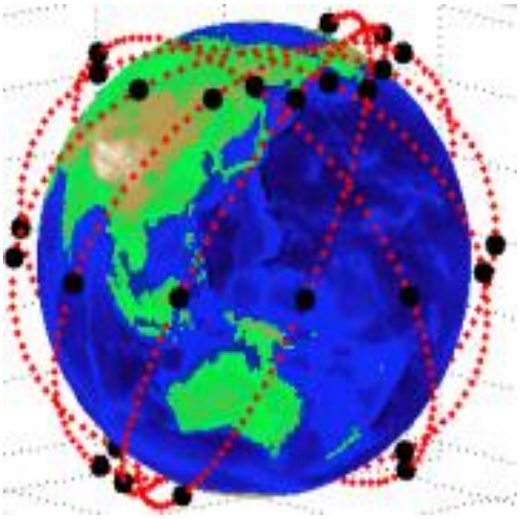


Figure 11. 36-satellite infrastructure configuration, with six satellites equally spaced in six planes at a 63.5 degrees inclination.

Thus, with all parameters in place and the model validated as shown above, it could then begin the process of evaluating designs for a LEO data relay infrastructure. The final analysis was broken down into two case studies: the first requiring total connectivity to a LEO customer (i.e., no outages in infrastructure availability due to it being out of range), and the second case not requiring this strict capability but still designed for optimal performance. The parameters for these cases are described in greater detail in the introductions to each case study below.

Before continuing to the case study results, however, the system dynamics methodology used to evaluate the case of OOS must first be covered in depth.

Assessing On-Orbit Servicing: System Dynamics Approach

The next section describes the system dynamics methodology, with an emphasis on explaining why this modeling methodology is useful for the evaluation of long-term space industry trends. System dynamics differs significantly from the physics and cost modeling used for the case of LEO communications and data relay infrastructure. In that case, the research questions called for the determination of optimal constellation and spacecraft designs, in order to find the configurations that would yield the greatest performance at the lowest overall cost. In this next case, however, it is assumed from the outset that industry has already determined the best technical designs for an OOS infrastructure and its components (as there is significantly less variability in the design of a servicing spacecraft than of an entire constellation of disaggregated communications spacecraft). Therefore, the research questions call for an evaluation of the long-term implications of OOS implementation, especially with respect to the ways in which added costs and shifts in risk posturing will affect the direction of basic architectural and industrial trends.

Along these lines, system dynamics was identified as a useful and pertinent methodology, due to its ability to capture the interactions of several interconnected system variables and simulate the overall behavior of the very complex system found within a basic space architecture. Through analysis of the trends produced by the system dynamics model, insights

can then be gained as to the viability of various policy regimes for the implementation of an OOS infrastructure.

System Dynamics Methodological Background

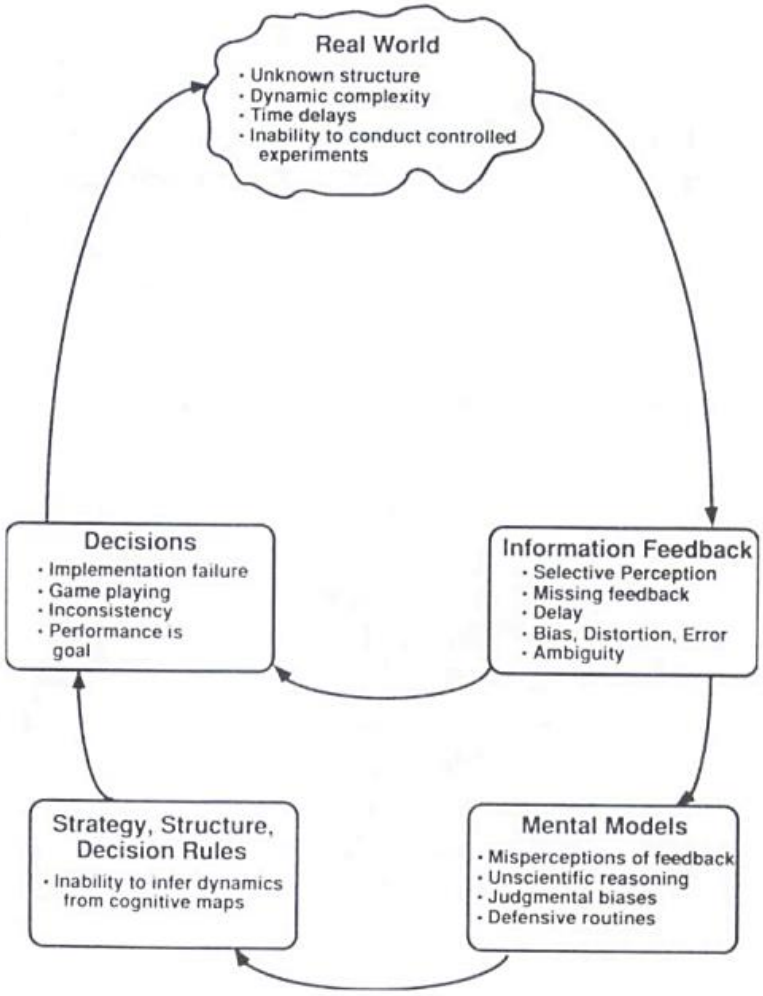


Figure 12. System dynamics ontological and epistemological framework (Sterman, 1994)

The modeling methodology of system dynamics has gained significant credibility over the past few decades, and Sterman provides a seminal description of the underlying reasoning and justifications behind system dynamics modeling in his 1994 article, *Learning in and about*

complex systems. He describes the system dynamics worldview as one in which decision-making is a constant iterative process, influenced by information feedback and the decision-makers' own mental models and beliefs about the system. Sterman also describes the way in which humans are typically unable to fully visualize the causal relationships amongst the many disparate variables in a system, and as such they may not recognize the ways in which certain decisions bring about unintended consequences (especially when there are significant delays in the realization of those consequences). In addition, Pruyt outlines a system dynamics research cycle, using the causal loop diagram shown in Figure 13, which is in fact itself one of the main tools for system dynamics modeling. This diagram shows the overall system dynamics method, including the way in which the researcher incorporates external data and his/her own mental model of the world and then simulates in an iterative approach to converge on a better approximation of the system under analysis. Thus, system dynamics modeling acts as a means of capturing the relationships amongst system variables as well as simulating the effects of policy changes and management decisions throughout the entire structure (Sterman, 1994).

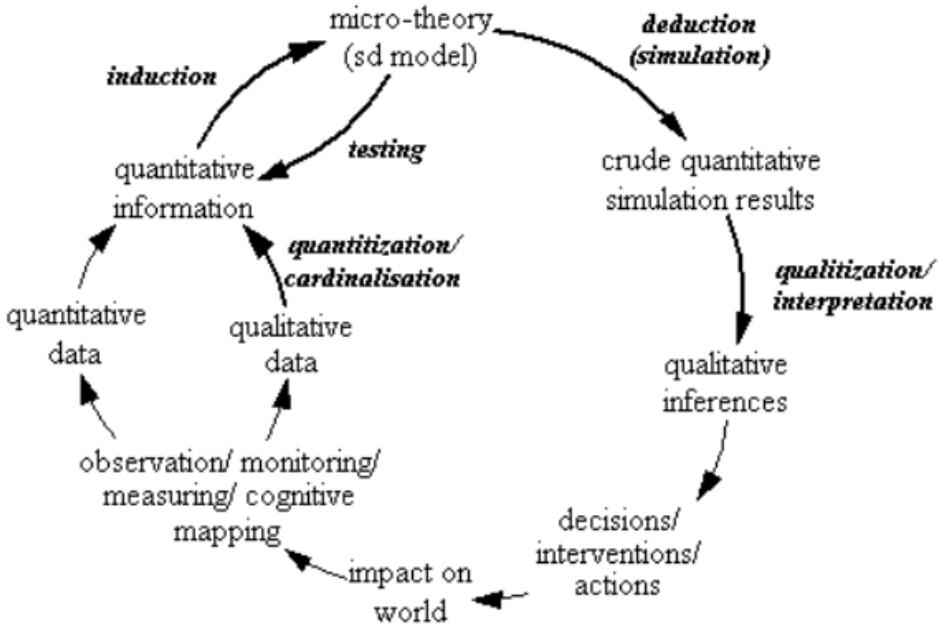


Figure 13. The System Dynamics research cycle (Pruyt, 2006)

Next, Wolstenholme in the 1980's made justifications for why system dynamics should be classified as a *methodology* in its own right. He reasons that system dynamics is "capable of

assisting with practical problem definition, analysis and change in a wide range of systems and with a potential to provide a more significant contribution to current general system practice" (Wolstenholme, 1983). The processes involved with system dynamics not only aid in providing a qualitative description and evaluation of a complex system (through the visualization of causal feedback loops affecting quantifiable stocks, inflows, and outflows), but they also allow modelers to arrive at quantified descriptions and solutions to system problems. Using system dynamics to identify and resolve complex issues has also been shown to grant modelers and clients a better overall understanding of complicated systems, affording them insights which had not been arrived at using other methodologies.

Additionally, the Systems Dynamics Society outlines system dynamics as a *methodology* as well. The society focuses on the word *model*, because that is the final output or goal of the research, but they are clearly describing a mixed-method in development of the *model* which serves a research purpose. "…System dynamics modelers not only use traditional econometric methods to estimate model parameters using quantitative data, but also routinely augment those methods with qualitative research methods including use of archival documents, interviews, and ethnographic methods and direct observation of decision making and organizational processes. Model testing involves quantitative assessment of the ability of the model to reproduce the behavior of the system of interest, and a wide range of additional tests including structure assessment, dimensional consistency, extreme condition, behavior reproduction, surprise behavior, sensitivity analysis, and system improvement tests" (System Dynamics for Academia, 2014).

The actual process of system dynamics modeling involves the use of interconnected loops of stocks, flows, and variables, which all together represent a larger complex system. Stocks can represent both materials and information; that is, they represent a tangible accumulation of some quantity or concept within a system. Thus, in a space industry context, stocks can represent actual satellites in orbit or under production, as well as represent the level of experience in the industry or the design choices (like mass, design life, and cost) of new generations of spacecraft. Flows, meanwhile, represent changes in the stocks, and these changes are in turn driven by the myriad variables that influence the long-term development and behavior of a system. Links are

drawn between all of these stocks, flows, and variables, such that several interconnected feedback loops are formed. It is these loops which ultimately drive system behavior, as one component of the system affects another, which affects another, which then eventually comes back to inform the output of the first component. Mathematically, the stocks and flows represent a system of coupled, non-linear differential, which are calculated smoothly and efficiently through computer simulations. By correctly identifying the dominant links and feedbacks within a complex system, system dynamics seeks to understand overall system behavior and then inform policy decisions which can direct the system towards optimal outcomes (Sterman, 2000). A graphical representation of all of the concepts discussed above is provided in Figure 14 below. The symbols used in this figure are also used throughout the remainder of this thesis as graphical representations of a theoretical space architecture and the underlying space industry.

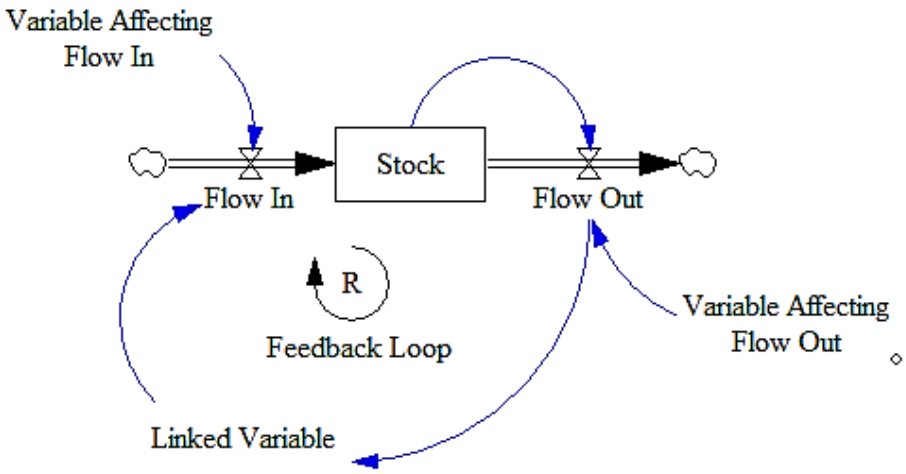


Figure 14. This simple diagram illustrates the components of a system dynamics model.

Incorporating Behavioral Economics Theories with System Dynamics

A short discussion is needed to explain the relationship between emerging theories in behavioral economics and their place within the system dynamics methodology. Concepts in behavioral economics are a natural fit with system dynamics, as one of the core tenets of the

methodology is to properly capture the mental models that decision-makers consult when faced with choosing policies. One key heuristic that has already been widely used by system dynamicists is anchoring and adjustment (Tversky & Kahneman, 1974). This describes the way in which humans typically base their decisions on some "anchor" value, and then they adjust their estimates as new information is obtained (with adjustments typically being inadequate, due to the much larger influence of the original anchor value). Sterman showed how this heuristic causes wild oscillations in a stock maintenance game, as human players attempt to minimize a fictional firm's cost by maintaining an appropriate stock of inventory throughout small shifts in demand. Widely known as the Beer Distribution Game, players almost always become anchored to certain levels of inventory and then fail to adjust adequately as conditions change (Sterman, 1989). It is postulated that the same heuristic is influential in space systems designs. For instance, space systems decision-makers likely become anchored to a certain level of design lifetime and cost per operational day, and then they become unwilling to change to a new level as conditions change. This would also provide partial explanation for why spacecraft have slowly become more massive, more powerful, and longer-lasting, as designers become anchored to the design paradigm of increasing mass, power, and design life with each generation and then fail to adjust far beyond such design preferences (even if significant cost savings and utility gains could potentially be gained through other design paradigms).

System dynamics allows this heuristic to be implemented as an adjustable information stock. The anchor value is a stock that begins at a reference point; for example, in 1993 the average design lifetime was approximately 10 years, the average mass was approximately 2600 kg, and the average power was about 6 kilowatts (refer to Figure 1, Figure 2, and Figure 5 above). As the simulation progresses this anchor value can be adjusted, such that it changes over time according to the rules implemented by the driving variables of the structure. This reflects the way in which an anchor value can be changed slightly by one decision-maker, with the new value then acting as the anchor for the next decision-maker. Figure 15 below provides a basic graphical representation of this phenomenon, showing the way that design points like mass, design lifetime, and power are adjusted over time according to other space industry variables, with these variables then also being influenced by those design points.

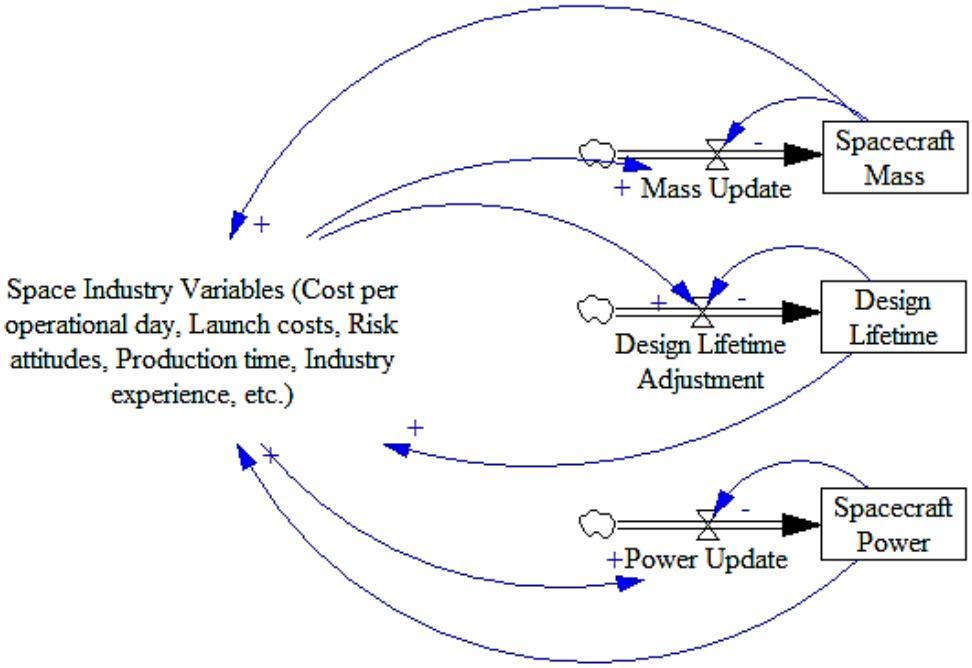


Figure 15. Information stock structure, showing how the anchor values are represented as stocks, the adjustments as flows, and the interconnected nature of the anchors, adjustments, and space industry variables.

Risk also dominates the space system decision-making process. Spaceflight is an inherently risky enterprise; Wertz estimates that about 10% of launches fail to insert spacecraft into orbit, and once the craft is in orbit there is a near-infinite number of failure modes, both known and unknown that can prematurely end a space mission (Wertz J. , 2011). Space is also a field where successes largely go unrecognized, while failures often take the form of highly visible explosions or losses of critical services. And of course, a typical space mission requires hundreds of millions (if not billions) of dollars, making the stakes that much higher and giving decision-makers ample justification for avoiding any and all risks. Still, it is likely that intense risk aversion is directly linked to increasing costs and development schedules, to the point where decision-makers are overweighting the probability of mission failure and thus missing opportunities for cost and schedule reductions. This follows from one of the core tenets of Prospect Theory; i.e., that losses are felt more strongly than gains, and therefore a small chance of failure will be overweighted by the decision-maker (Tversky & Kahneman, 1979).

Spaceflight is also a field where engineers and managers are under extremely high pressure to succeed, and as Curley et. al. make clear, the certainty of evaluation of one's decisions by peers and superiors has been strongly linked to a desire to avoid risks and ambiguity (Curley, Yates, & Abrams, 1986). Furthermore, it has been shown that people are willing to spend significant amounts of money to avoid ambiguous processes (like those associated with new, unproven space systems) in favor of other processes which are known but entail a normatively equal level of risk (Trautmann, Vieider, & Wakker, 2008). The biases introduced by all of these phenomena can be included in a system dynamics model, by forcing extra costs and delays proportional to the perceived level of risk and then comparing the inefficiencies this causes to the case of a perfectly rational decision-maker. These risk features also allow for the testing of a key rationale for the implementation of OOS: that it will reduce the level of risk aversion widely felt in the space industry and allow for decreased spacecraft costs and mass.

Level of Analysis and Modeling Boundary

As described in the sections above, system dynamics as a methodology seeks to understand complex systems by breaking down their components into interconnected loops of important stocks, flows, and variables. Within the context of on-orbit servicing, this methodology can aid space mission architects and policy-makers in understanding the long-term impacts of OOS implementation, in addition to shedding light on the policies most likely to ensure the success and viability of this technology. Next, the framework and level of analysis for analyzing the space industry with system dynamics are covered, and followed by an in-depth discussion of the components of the model created for analysis within this thesis.

To examine the potential impacts of on-orbit servicing on the wider space industry over a long timeframe, an appropriate a level of analysis must be identified from the following hierarchy of the overall space industry:

- Satellites
- Satellite Constellations
- Unites States Space Industry

- Unites States Industry

The level of analysis for this research is *satellite constellations*. These constellations are effectively the units of analysis that form the overall U.S. Space Industry. While the U.S. Space Industry includes many diverse actors and stakeholders, from civil agencies to the military to commercial entities, this analysis seeks to examine the shift in risk posturing and satellite design choices that are projected to result from the widespread use of OOS. This is because current trends in spacecraft design are not viewed as an issue affecting a single firm or entity, but rather as an industry-wide phenomenon which has ramifications for all stakeholders.

This decision is directly in line with the system dynamics methodology, as it defines the boundary for both the research and the modeling effort. Setting the level of analysis also leads to the definition of what is exogenous, or outside the control of the model's feedback structures, and what is endogenous, or should be included as internal variables in the structure. Having set the level of analysis as satellite constellations within the US Space Industry, the model developed for this analysis focuses primarily on their design and replenishment, and tracks some of the effects that OOS may have on individual constellations and the overall industry. This also implies that activities in other technical sectors are exogenous to the model structure.

The system dynamics model for this analysis focuses on the design choices affecting the development and upkeep of a theoretical geostationary communications constellation, as a proxy for other activities within the greater U.S. Space Industry. This modeling choice is justified due to the fact that this particular area of the space industry (geostationary satellite communications) is firmly established, comprising over a third of all unmanned space activities and with a great deal of publicly available data (The Tauri Group, 2014). Additionally, the established nature of this sector of the space industry directly lends itself to becoming a first and primary customer for spacecraft servicing; indeed, NASA has identified the geosynchronous as one of the most likely sectors to benefit from the introduction of OOS (NASA Goddard Space Flight Center, 2010). Because the analysis uses satellite communications as a proxy, the hierarchy of levels of analysis changes slightly to become the following:

- Individual Communications Satellite
- Communications Constellation
- Space Communications Activities
- U.S. Space Activities
- Space Industry

Additionally, the following list contains the key variables identified for analysis of the space industry:

- Satellite Design Life
- Production Time
- Cost per Satellite
- Cost per Operational Month
- Strength of Learning Curves
- Industry Experience
- Satellite Mass
- Number of Launches
- Number of Satellites
- Cost of Launch
- Kilowatts of Power (as a proxy for satellite performance)
- Spacecraft Failure Rate

Many of these variables can be considered as clustered variables; i.e., they represent a larger idea that is actually composed of several variables that may not be directly quantifiable. For example, the single variable of satellite power, measured in kilowatts, is put forward as an appropriate measure of satellite "performance" (an otherwise tricky variable to quantify, given the highly specialized nature of most satellites). Satellite power determines the number of customers that can be serviced as well as the data-rate associated with that servicing; as such, power can be used as a single proxy variable to represent the performance that communications satellites are most interested in. The system dynamics methodology functions primarily on the

core tenet that it is through this clustering or combining of variables that the actual system behavior can be modeled and analyzed, and so the fact that some variables seem only tangentially related should not detract significantly from overall model validity. Indeed, the usefulness and ability of the variables to grant insight to the issues facing the space industry with and without the implementation of servicing will almost certainly be limited by the incredibly complex nature of such a large, multi-faceted system.

Finally, risk plays a key role in this model, and it is implemented using a "Success Rate" information stock, which tracks the overall rate of successful spacecraft launches and can be adjusted over time according to cost and risk preferences. This structure is covered in greater depth in the next section. The main variables governing the movement of success rate, though, are two separate functions: a risk-based value function and a risk-to-cost function. The risk-based value function reflects the prospect theory concept of losses being felt more strongly than gains (in this case, spacecraft failures being felt more strongly than successes) (Tversky & Kahneman, 1979). This value function can be adjusted to vary the relative strength of losses over successes, but is typically set such that the structure will always push the success rate towards 100%. The success rate is balanced on the other side by the risk-to-cost function, which increases the cost of each spacecraft when the success rate is adjusted above a baseline risk level (to reflect the more costly engineering required to decrease risk). The success rate reaches equilibrium when the value function and risk-to-cost function reach an equivalent level of adjustment, reflecting the assumption that decision-makers will decrease risk only to the extent allowed by cost considerations (i.e. when prospect theory utility is maximized). Finally, the impact of the value function is tied to the amount of time required to produce replacement spacecraft, reflecting the fact that decision-makers will be far more risk averse when they know that any failure will require years to replace.

Building the System Dynamics Model

The design of satellite constellations and the influence of human bias over time were modeled and simulated within a deductive model, with the major structures of this model

outlined in Table 2. Each structure, its implementation, and its purpose are further examined in the sections below.

Major System Dynamics Model Structures	
Production Pipeline	Implements a mechanism to trigger replenishment of the constellation and ensures capacity requirements are met; also implements appropriate time delays between requesting and fielding satellites. Key Variables: Number of satellites being produced, Production Time, Satellites on orbit, and failure rate.
Industry Learning Curves and Experience	Keeps track of long term activity in the US Space Industry and translates this into reductions in cost or production time over generations. Key Variables: "Experience" and "Learning Curves."
Space Vehicle and Architecture Design	Updates constellation design points as the simulation progresses, according to activity in other structures as well as exogenous design demands. Key Variables: Satellite Cost, Satellite Launch Cost, Satellite Mass and Cost Per Operational Month
Prospect Theory-based Success Rate	Implement a prospect theory-based optimization between cost and failure rate, based on cost and value functions. Key Variables: Success Rate, Value Function, Risk to Cost Function.
Exogenous Technology	Translate real world data on geosynchronous communications satellites performance and technology changes over the last thirty years into a model-usable technology advancement curve, as well as provides design updates to each new generation of spacecraft. Key Variables: Kilowatts of Power (proxy for performance) and Mass Reductions from Technology Advancements.

Table 2. Major system dynamics structures used for the model in this analysis.

Production Pipeline Structure

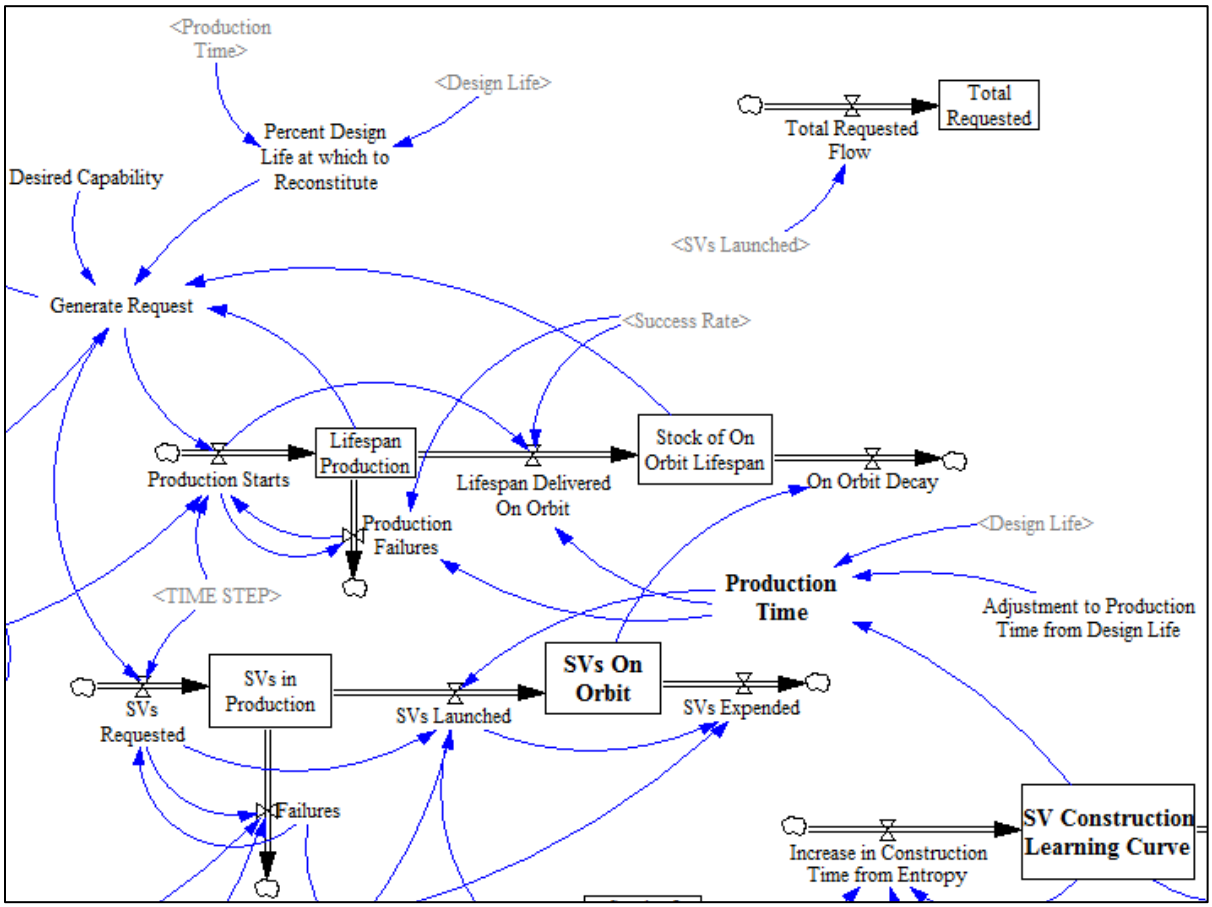


Figure 16. The satellite constellation production and replenishment stock and flow structure

The first requirement of the system dynamics model is that it be capable of modeling a basic satellite architecture. For this analysis, the architecture was composed of 12 satellites on orbit, requiring regular replenishment as satellites naturally die at the end of their design lifetime. It was also decided that replacement satellites must be purchased in the mode of campaign acquisition, where the entire block of 12 satellites is purchased at the same time, rather than only as they are needed (however, replacements for failures were allowed to be purchased immediately after the failure occurred as add-ons to the previous block). This acquisition method most closely resembles that of most government and commercial space enterprises. Of course, in anticipation of failures extra satellites could have been requested as part of the initial purchase; both approaches are seen in real acquisitions. The overall difference between these approaches is that if re-orders occur after failure, then the production pipeline is kept open longer and "experience" is given a small smoothing, whereas if extras are ordered initially then the

model grants a slightly higher initial "experience" level but also degrades further in-between purchases. Neither approach significantly changed results within the test cases modeled.

The basic satellite replenishment loop was implemented using two parallel stock and flow chains, shown in Figure 16. The top stock and flow chain tracks "on-orbit lifespan," which is a product of the number of satellites on orbit and the current generation's design lifetime. Directly below that stock and flow chain is another chain, which tracks the actual number of satellites on orbit at any given time. Two chains were required because the signal to send a request for more satellites is a function of both the number of satellites on orbit and how much longer they would last, as well as the amount of time required to build a new generation ("Production Time"). Thus, the model keeps track of all three of these variables to determine when to request more space vehicles (SV's), which also ensures that at least 12 satellites are always present in the "SVs On Orbit" stock. It should be noted that if the "Percent Design Life at which to Reconstitute" is low enough, then the model will not be able to keep 12 satellites on orbit at all time; it must be greater than or equal to the Production Time divided by the Design Life. This is logical, as acquisitions takes time and satellites and rockets can have production times of years.

Industry Learning Curves and Experience Structure

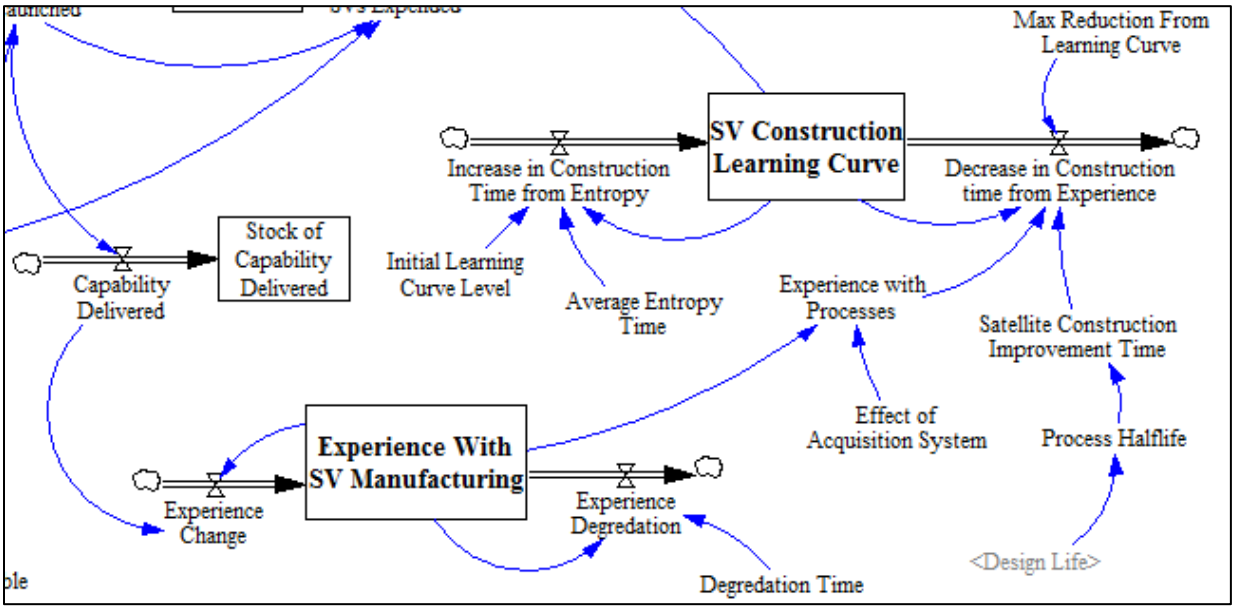


Figure 17. System dynamics structure for modeling the accumulation of industry experience and the strength of learning curve effects

Directly below the satellite replenishment structure is the industry learning curves structure (shown in Figure 17). This starts with the "Capability Delivered" flow, which is exactly the same as "SVs Launched" but allows the modeler to keep a running tally of total satellite capability delivered over time. This then drives an increase in "Experience with SV Manufacturing," which increases when new SVs are launched but also slowly decreases over time due to natural experience decay. The "SV Construction Learning Curve" stock and flow chain is more complex. It begins at an initial level of 1 (representing full price, full production time), and is then decreased by experience gains (down to a minimum level arbitrarily set as .2 or 20%). It must be noted that this learning curve represents the cost of creating new programs and technologies; i.e., it does not represent cost savings within individual generations. The gain in learning curve effects is caused primarily by the "Experience with Process," or the relationship between the process half-life and how much relevant activity is occurring within this time period. The combination of these two structures is tuned to represent gains from activity as well as loss from entropy, or the slow degradation of human capital and the effects of technological obsolesce. The model is tuned to expect two generations of technology advancement per decade; this implies that, after 10 years, knowledge is only worth 25% of its original value with respect

to current activities. The model is also tuned such that the marginal gain in knowledge over time decreases with each individual unit (e.g. the cost and production time decrease between the second and third SV constellation build is substantially more than between the 52nd and the 53rd.) In summary, the structure tracks the number of satellites which are produced within a time period, and the efficiency at which this knowledge is captured by the industry, in order to determine the impact of this activity via gains in the industry learning curve and subsequent decreases in satellite cost and production time.

It should be noted that experience also increases with SV failures. Even though a failed SV must be replaced, and this replacement increases overall costs, experience is still gained by the overall system with each failure. This reflects one way in which very reliable spacecraft can actually *decrease* overall industry experience and learning curve effects, especially when the spacecraft are very long-lived and there is a great deal of time between the launching of new generations. This also lends support for many of the proposed benefits of OOS, which in addition to allowing spacecraft operators to save failed spacecraft by executing the option for servicing, should also theoretically allow for the adoption of a more flexible risk management posture (which should also decrease costs). This shift in risk posturing is explained in much greater detail in the prospect theory-based risk posture structure below, but within the context of industry experience, the overall system benefits not only from less costly failures but also from the experience gained from allowing and fixing a greater number of spacecraft failures.

Space Vehicle and Architecture Design Structure

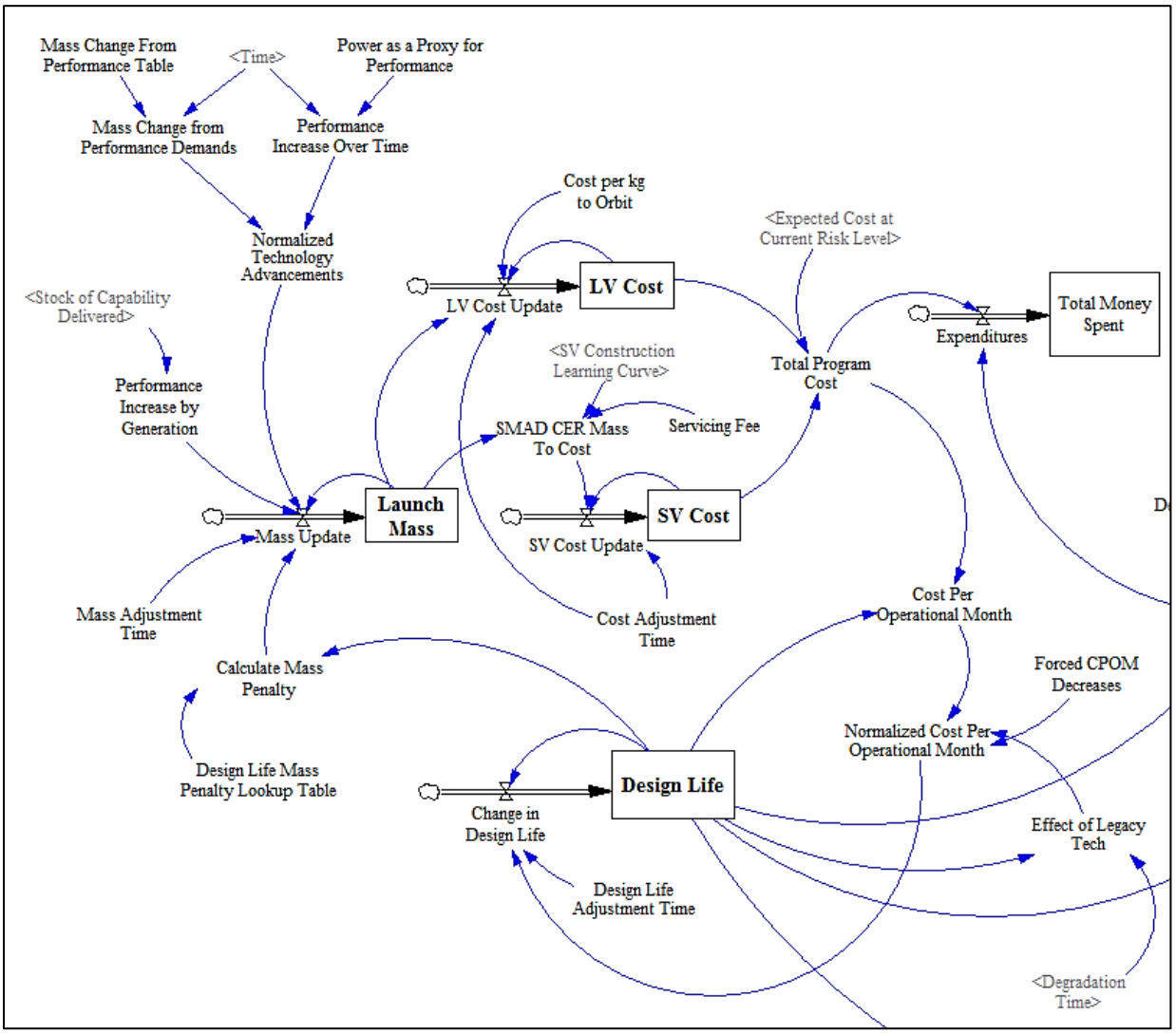


Figure 18. The satellite architecture design structure, which tracks the key design points for space vehicles (Design Life, Mass, and Cost), and adjusts each design point in order to keep cost per operational month low while factoring in any possible exogenous desires for increased performance.

The satellite and architecture design loops were constructed as a means of tracking over time the changes in average SV mass, design life, and costs (with this structure displayed in Figure 18). The stocks of Design Life, Mass, and SV Cost always begin at the same initial level, but can increase or decrease based on the parameters of each simulation. For instance, mass is increased due to an exogenous demand for greater performance (assuming satellite mass is an adequate proxy for overall power and capability), and this increase in mass is linked to the

delivery of each generation of spacecraft. While SV mass is not a perfect proxy for overall performance, since different technologies and designs will inherently lead to different masses, it is considered adequate for the case of large, standardized geosynchronous communications spacecraft. For spacecraft of this type, overall performance can largely be measured by their total power, number of transmitters on board, and total fuel for repositioning, all of which contribute to SV mass (e.g., more power means larger, heavier solar panels and batteries). However, it is also recognized that, over time, technology advancements will lead to miniaturization and decreased component masses, and as such SV mass is also decreased by a "Normalized Technology Advancements" curve, which tracks the improvements in all technologies over time. This model assumes technology allows an approximately 10-fold reduction in satellite mass below the initial level over the course of the 30 year simulation.

Then, any increases or decreases in SV mass lead to proportional increases in space vehicle and launch vehicle costs, which both contribute to Total Program Cost. This program cost is then divided by the Design Life in order to compute Cost per Operational Day (CPOM). This cost is normalized, and then linked back into Design Life, such that an increase in CPOM leads to an increase in Design Life. The structure is built this way—with CPOM directly linked to Design Life—due to the core assumption that satellite architecture designers will call for an increase in Design Life in order to decrease CPOM and amortize increasing program costs over a longer time frame (with this assumption being supported by the findings of Saleh et. al.). Increases in Design Life, though, also result in Mass increases (because of increased redundancy, greater structural integrity, and increased on-board fuel requirements), such that the entire structure becomes a reinforcing feedback loop.

Of course, Design Life can not be increased indefinitely, as longer-lived SV's begin to exhibit the negative effects of technology obsolescence. To model this, a variable described as the "Effect of Legacy Tech" is included, which is linked to the degradation time of technology also being used to effect industrial experience. This effect of legacy tech tracks the increase in Design Life and begins to slow its increase once design life is equal to or greater than the technology degradation time; i.e., once the age of satellite technology on orbit becomes so advanced that it begins to conflict with the experience retention of engineers on the ground, and

it also becomes far less valuable compared to emerging, present-day technologies. This is a somewhat rough estimate of the limits on satellite design life; indeed, scientists and engineers could theoretically design a spacecraft that would remain active indefinitely (as many very old, still functional spacecraft have done inadvertently). However, for the purposes of this model, limits must be placed on design life increases, and linking these limits to the technology degradation time was a logical connection.

This structure also incorporates some of the effects of on-orbit servicing. When OOS is implemented in the model, mass can be frozen at its pre-implementation value, such that it does not increase or decrease once OOS becomes available. This reflects the assumption that SV's will continue to be designed for the same basic mass level; i.e., even if the refueling capabilities of OOS decrease total on-board fuel requirements, SV designers will simply make up for these mass savings by adding more transmitters or other similarly value-adding components (this assumption is supported by the case of all-electric satellites, whose mass savings have simply led satellite operators to order and launch more spacecraft for greater value on orbit) (Clark, 2012). Mass can also be decreased over time as space architectures incorporate more of the benefits of OOS (such as refueling and technology upgrade capabilities). However, it is not expected that these mass savings from OOS will be completely free; indeed, an increase in spacecraft cost can also be implemented with OOS, which is reflected in the "Servicing Fee" variable. This servicing fee can be adjusted to assess the overall model's reaction to a wide range of OOS costs, from a completely free infrastructure to one that would charge spacecraft operators as much to service their spacecraft as it would cost to launch a new SV. This also reflects the range of policy options available for the implementation of OOS, and this is discussed further in the analysis results.

Prospect Theory-Based Risk Posture Structure

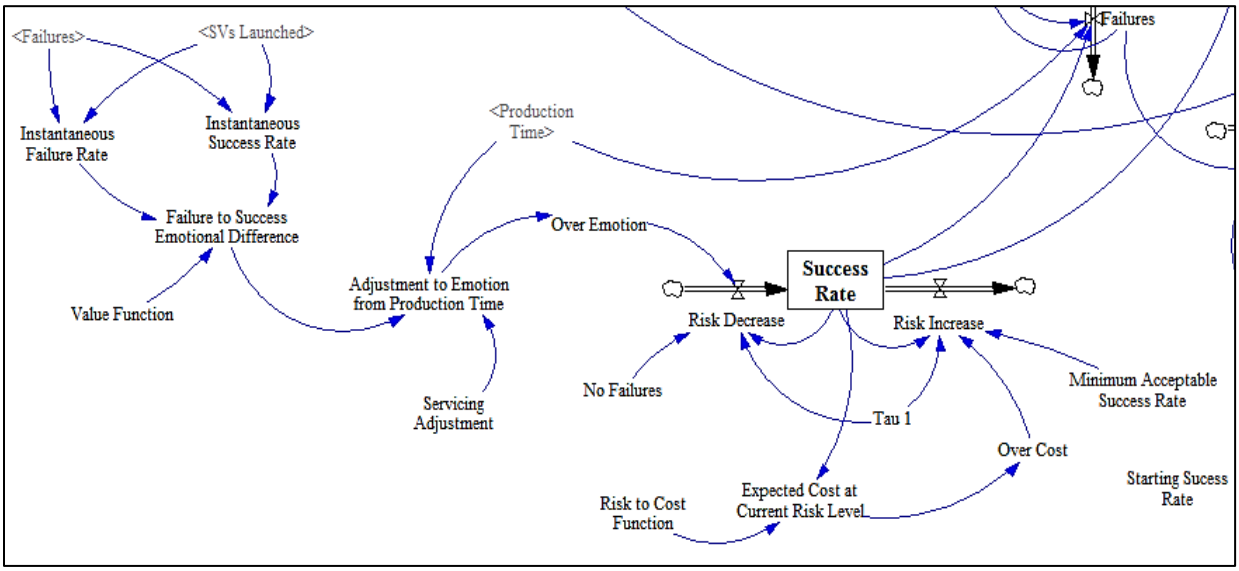


Figure 19. System dynamics structure for modeling the risk- and cost-based space vehicle success rate. The negative value associated with SV failures attempts to drive up the success rate, while the costs associated with risk aversion attempt to drive down success rate, leading to a prospect value-based equilibrium.

The structure for modeling the effects of prospect theory-based risk aversion (shown in Figure 19) hinges primarily on a "Success Rate" information stock. This stock is increased by the effects of the spacecraft failures and successes as determined by the Value Function (Figure 20), which will attempt to push the Success Rate to 1 (or 100% success) as long as there are spacecraft failures (with this effect becoming less and less strong as failures decrease and success approaches 100%). The "Emotion" value is also affected by spacecraft Production Time, such that a longer Production Time will lead to an increasingly strong push to adjust Success Rate upwards. At the same time, there is also a cost variable attempting to push the success rate down, as less reliable spacecraft are also assumed to be less expensive (due to less risk overhead, less shielding and redundancy, less testing, etc.). This cost variable is driven by a Risk-to-Cost Function (Figure 21) which provides some cost savings below a baseline success rate but will also cause substantial cost increases as designers attempt to make spacecraft more reliable. In this model, the baseline success rate was considered to be 90%; that is, the added costs associated with risk aversion will only be implemented once the model begins to push the success rate above 90%. This specific threshold was informed by Wertz, who states that it can

be expected that roughly nine out of ten space missions will experience some form of failure (Wertz, Everett, & Puschell, 2011). The Success Rate then reaches an equilibrium when the negative effects of failure are balanced by the cost considerations represented in the risk-to-cost function. It should also be noted that the Success Rate may never reach an equilibrium if the Production Time is changing, due to the way that an increasing Production Time increases the effects of failures in the Value Function.

This structure is also where the effects of on-orbit servicing are primarily enacted. One of the major proposed benefits of OOS is its effects on spacecraft risk management, such that SV's can be "saved" by OOS in the event of failure (in the form of anomalous mechanical issues, technological obsolescence, or fuel depletion). By having the option of OOS at their disposal, it is postulated that SV designers will experience less risk aversion, and will thus be much more tolerant of a decreased success rate. This is reflected in the "Servicing Adjustment" variable, which decreases the "emotional" effects of failure when OOS is implemented. One of the main benefits of this is that less reliable spacecraft are also modeled as costing less than those with a very high level of reliability. Of course, OOS itself also comes with a cost, and this was implemented in the SV and Architecture Design structure explained in the section above.

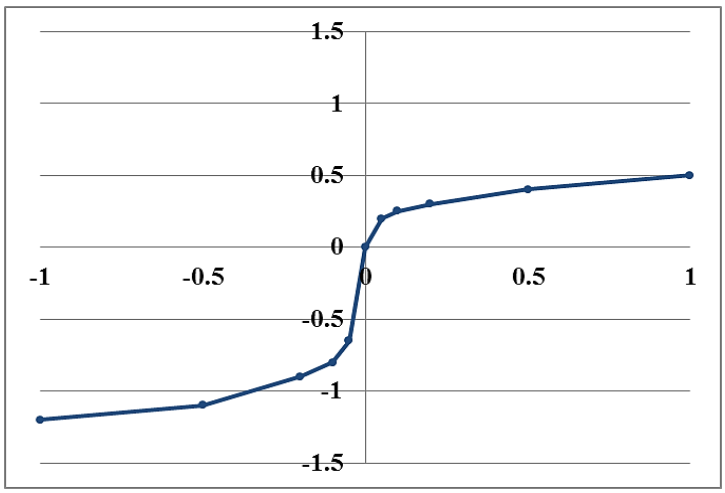


Figure 20. This Value Function reflects the way in which losses (spacecraft failures), multiplied by the steeply decreasing value function below the point (0,0), are felt more strongly than gains (spacecraft successes), which are multiplied by the more moderately increasing value function above (0,0).

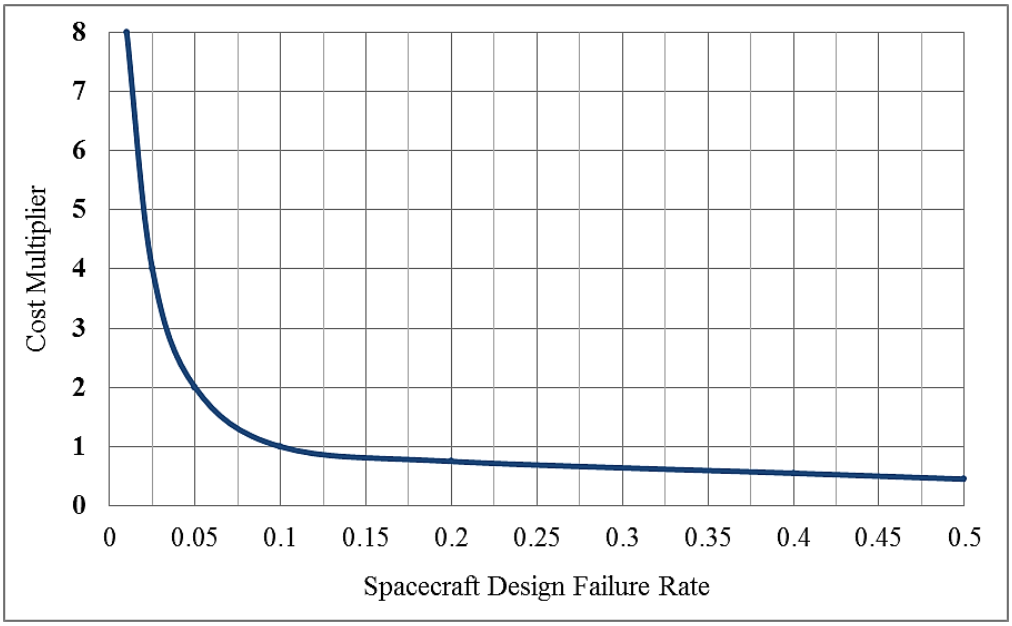


Figure 21. The Risk to Cost Function reflects the increasing multiplier on cost when success rate is increased beyond a baseline risk of 10% (as well as the cost savings which can be accessed below this baseline risk level). These estimates for cost multipliers were determined solely to aid in the modeling of the space industry and are not considered applicable to all spacecraft types.

Exogenous Technology

Because this model is meant to simulate the space industry over a very long time frame (of at least three decades), the overall advancement of technology external to the space industry must also be included. It was assumed that technology advancements which benefit spacecraft production come from both inside and outside the industry; that is, some technology improvements came from space industry-specific research and development, while others came from advancements made in other industrial sectors but which have applications to spacecraft. The overall increase over the past three decades in geosynchronous spacecraft power (displayed in Figure 22) was used as a proxy for technology advancements, due to the fact that overall spacecraft power reflects a variety of technologies, from photovoltaic cells to batteries to more efficient hardware. This growth in spacecraft power was then adapted to a normalized technology advancement curve, shown in Figure 23. These advancements act as a divisor to

spacecraft mass and cost, allowing for decreases in both as the simulation advances (provided there are not overriding exogenous increases to mass or cost).

It is worth noting that the process of developing and implementing OOS could also spur greater technological advancement; i.e., the production of the servicing spacecraft and servicing mechanisms themselves could actually lead to even greater advances in overall space technologies. However, for the purposes of this analysis, the introduction of OOS to the overall space industry did not alter the progression of technology, but instead was modeled solely as providing the services necessary to shift spacecraft risk, cost, and mass.

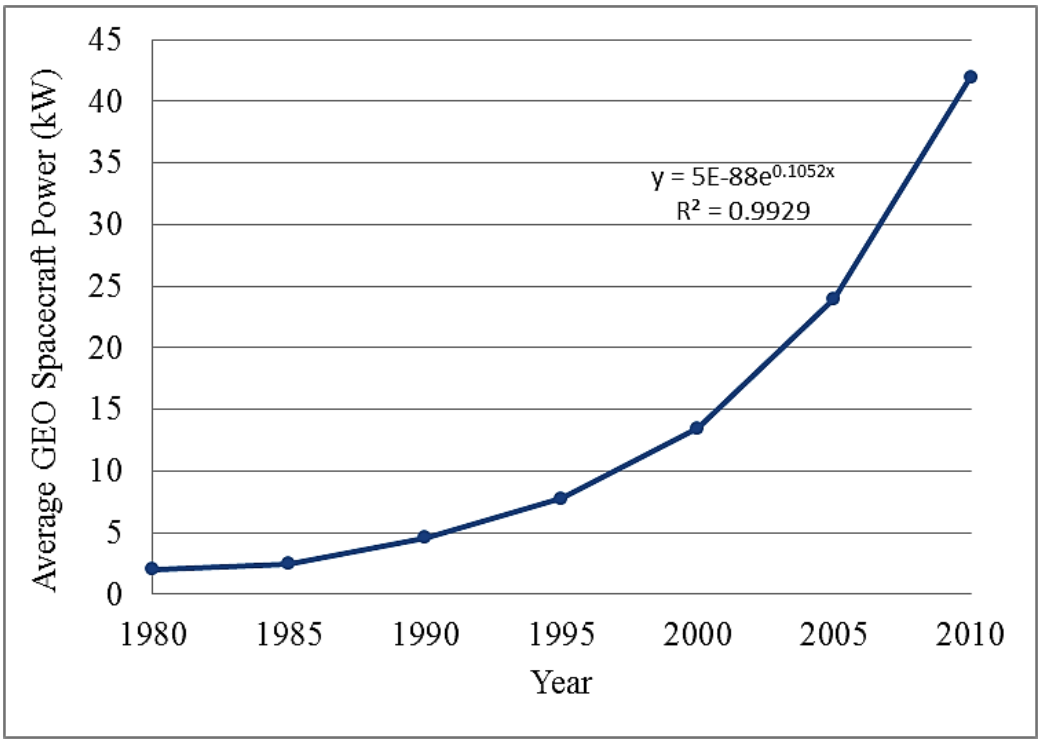


Figure 22. The increase in average geosynchronous communications spacecraft power over the past three decades (adapted from Jones & Spence, 2011).

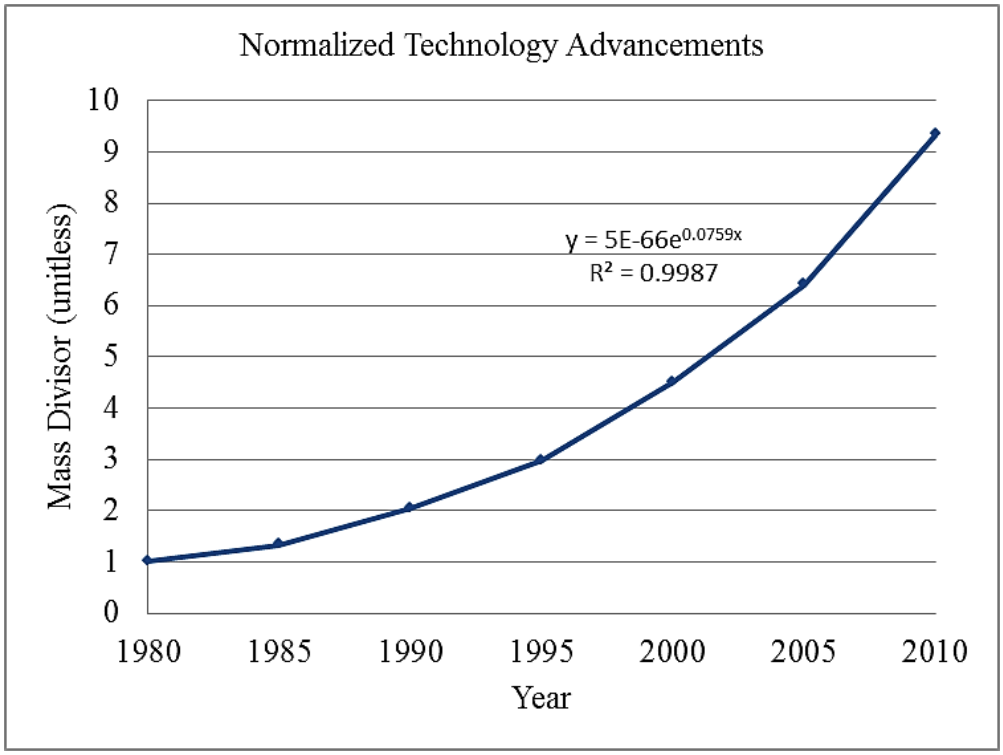


Figure 23. Normalized Technology Advancements is adapted from the growth in spacecraft power shown in Figure 22, and this acts as a divisor to decrease overall spacecraft mass and cost requirements.

Finally, shown in Figure 24 below is the entire system dynamics model used in this analysis. This shows the links between all the variables described in the sections above, in addition to providing graphical proof of the complexity involved in modeling a relatively basic space architecture. In the upper left corner of the figure is the Space Vehicle and Architecture Design structure. To its immediate right is the Production Pipeline structure, which actually tracks the number of spacecraft on orbit and the number needed in production. Below this pipeline structure is the Industry Learning Curves and Experience structure, which tracks experience gains and degradation in the industry and applies a learning curve to SV cost and production time. Finally, in the lower left corner is the Prospect Theory-based Risk Posture Structure, which analyzes the number of spacecraft failures and successes and then determines the designed-for spacecraft success rate, balancing both feeling of risk aversion as well as the increasing costs associated with a higher success rate.

Full Model

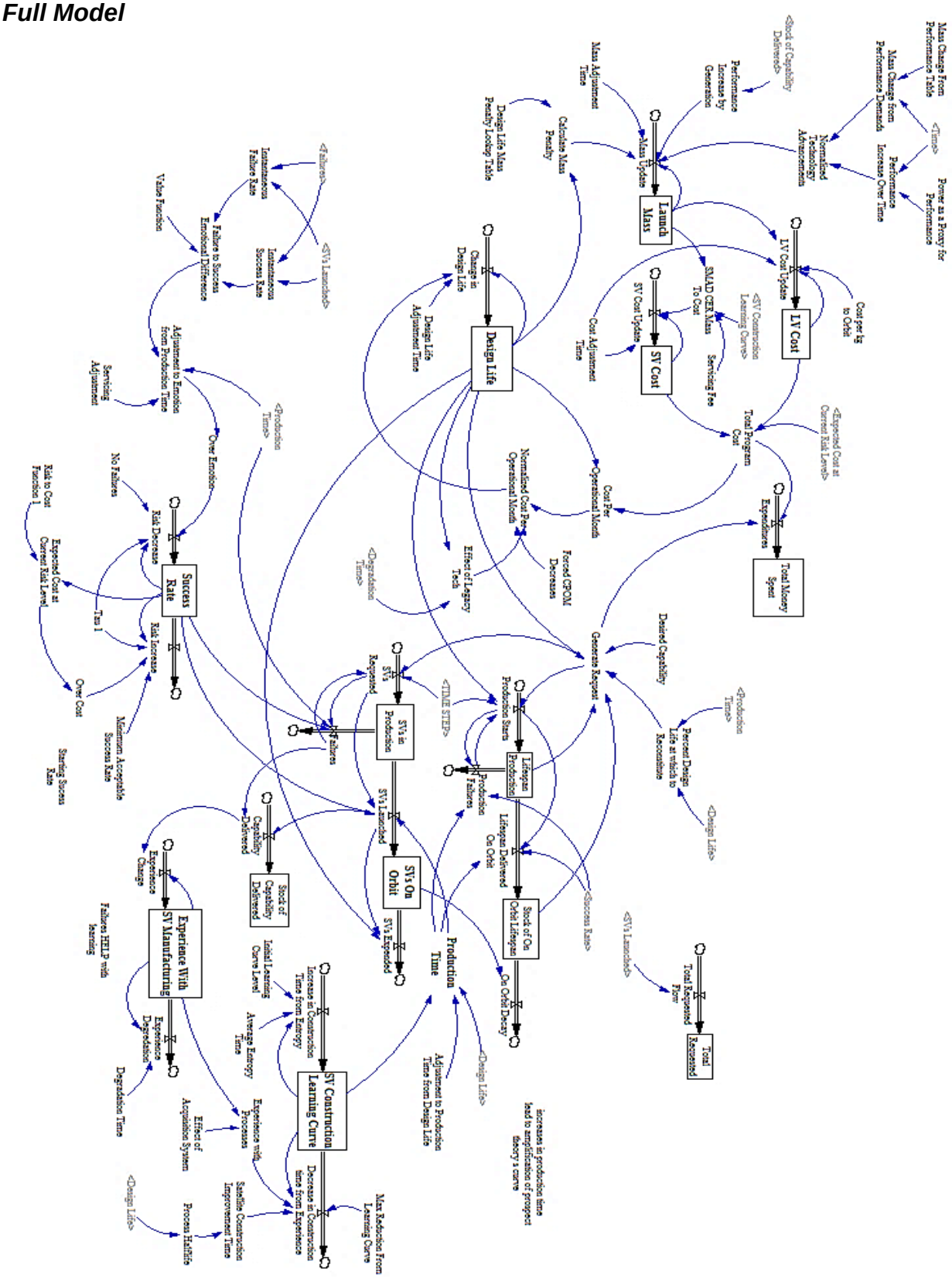


Figure 24. Full system dynamics model used for analysis.

Using System Dynamics To Assess On-Orbit Servicing Effects

To test the wider effects of the introduction of servicing spacecraft to overall space architectures, the system dynamics model can be tuned to simulate a pre-servicing/post-servicing industry evaluation. This was achieved primarily by running the model until the simulated space industry resembled present day paradigms, as shown by the SV mass and design life closely resembling the real-world progression of geosynchronous communications spacecraft. This also served as a means of validating the model's performance, and the results of this test are shown in the figures below. As is shown, the model returned SV mass and design life values very closely resembling that of real world satellites, lending a degree of confidence to the model's ability to predict overall space industry and operational effects into the future.

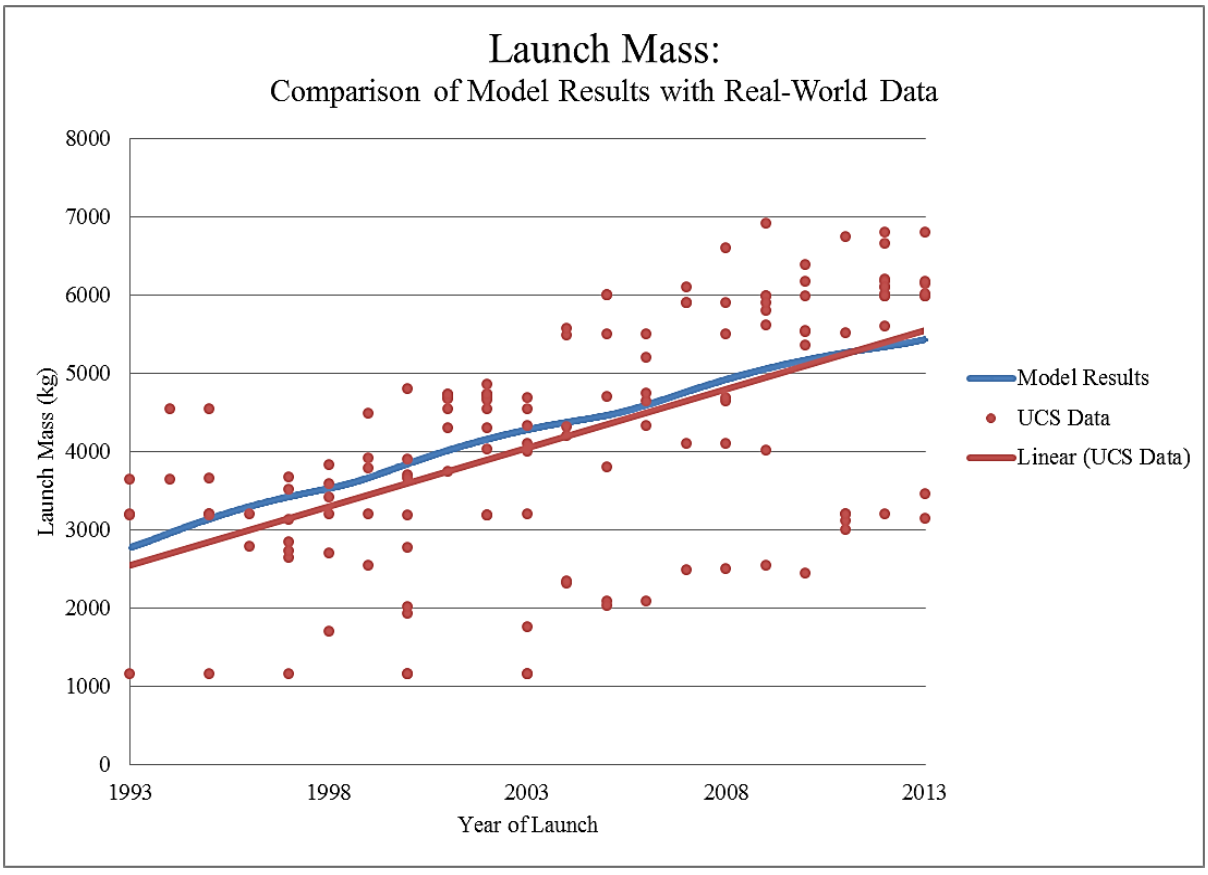


Figure 25. Comparison of the model results for SV launch mass with geosynchronous satellite data provided by the UCS.

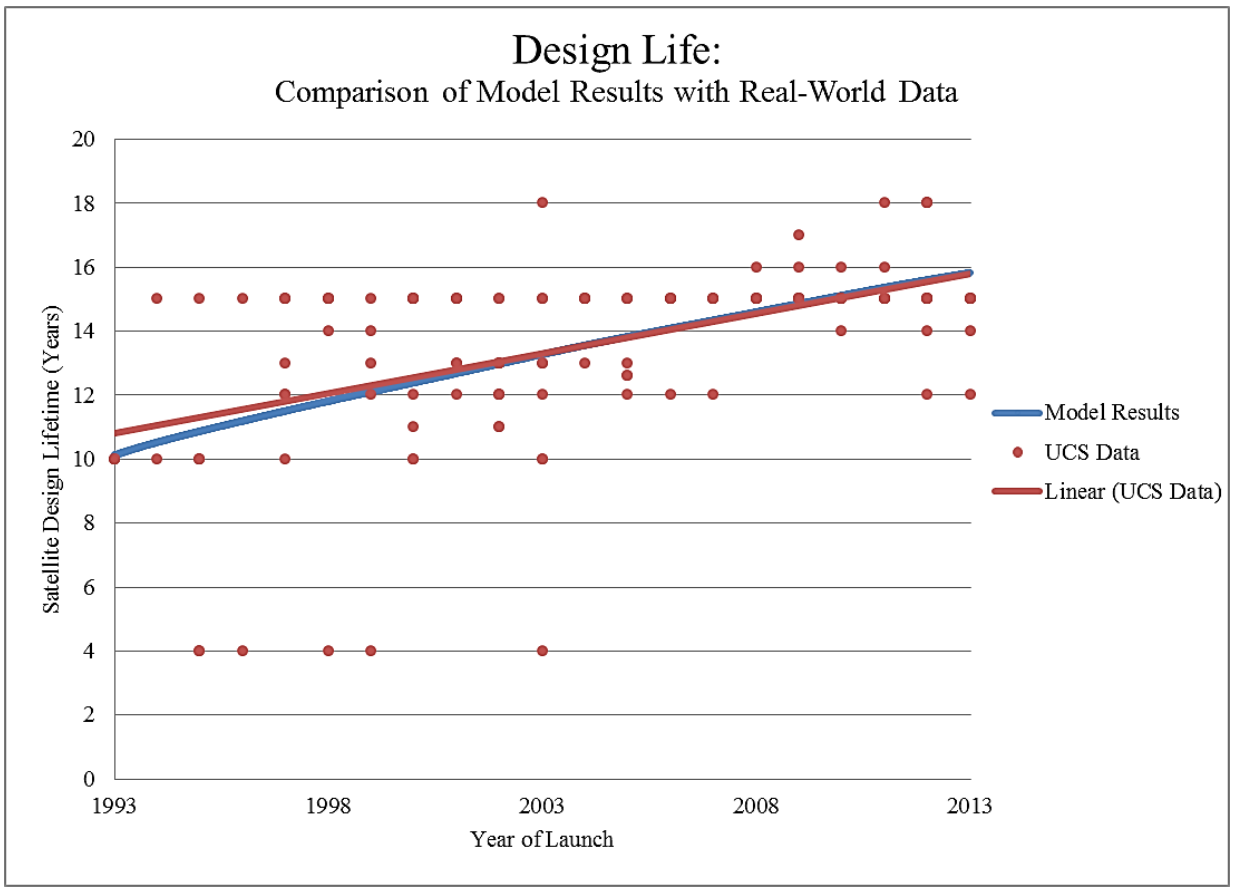


Figure 26. Comparison of the model results for SV design life with geosynchronous satellite data provided by the UCS.

Then, with model parameters resembling those of present day satellite architectures, a significant reduction in the architecture's risk aversion can be forced, as spacecraft operators begin to rely upon the repair and repositioning capabilities of the servicing infrastructure. This shift in risk posturing is forced through manipulation of the prospect theory-based value function, such that designers become willing to accept more risk (i.e., feel less negative emotion from failures) in the interest of cost savings and greater use of the servicing infrastructure. The shift in the risk function is applied in tandem with a series of new pricing structures: free, partial, or full price of a replacement satellite, applied over all spacecraft produced. All of this is meant to alter the overall "attitude" of the model from one of high risk aversion to one which is more accepting of risk, given a relatively "cheap" reduction in risk from reliance on the servicing infrastructure and the cost savings associated with building less reliable spacecraft.

A key focus of this analysis is to identify the "break-points" at which servicing becomes a preferable option to simply increasing mass or design life of satellites. In other words, it is believed that there is some price point at which it is cheaper, and thus preferable, to design for the option of servicing than to not have the option at all (i.e. designers would rather exercise the option to service a spacecraft, and accept a slightly less reliable overall design, than to completely replace it). From this point, the willingness to pay for a servicing infrastructure can be estimated, such that if the fees to use the infrastructure were any higher, then the average customer would be unwilling to pay for servicing and instead would simply launch a new satellite. It should be noted that, while servicing is often cited as a significant hedge against the risk of major mission failure, it can also be used as a means of upgrading or refueling operational satellites, such that their performance is maintained or even improved even as spacecraft age or technological obsolescence threatens long term value. Finally, this willingness to pay for servicing can also be tied back to various policy options for implementing OOS, such that the model can provide insight into which policies are most likely to ensure the long-term success of an OOS project.

The aim of this analysis is thus to examine the wider industrial effects of the introduction of servicing spacecraft. For example, what effects could this have on industry experience and learning curves? What price point and risk level does the industry settle on over time? How are spacecraft design lives, mass, cost, production times, etc. all affected by the introduction of servicing? It is not expected that any one simulation will answer these questions. Rather, a variety of simulations altering cost and risk functions (as well as other parameters) have to be run, and then general trends can be identified from the aggregate. Given the nature of system dynamics to identify trends rather than specific price points, it is also not expected that the model will return actual cost estimates of new spacecraft or OOS technologies. Instead, the developed model can be given general cases, designed to reflect actual policy options, to investigate which policies have the greatest potential for shifting architectural paradigms and ensuring the long-term success of OOS.

Results

With these methodologies in place, analysis of the two infrastructure concepts could then be conducted. For the LEO communications and data relay infrastructure, this thesis explored constellation designs, varying parameters such as number of spacecraft per orbital plane, number of orbital planes, satellite crosslink and infrastructure interlink data rates, and other constellation orbital characteristics. Using the physics and cost models and the tradespace exploration module outlined above, two case studies were conducted for the LEO infrastructure: one which ensured global coverage for LEO customers (Case Study 1), and another which lacked global coverage but still sought to optimize infrastructure performance (Case Study 2). Then, the system dynamics model outlined above was used to examine various infrastructure implementation scenarios for the OOS, varying parameters such as infrastructure cost and customer incorporation of OOS capabilities, in order to determine what policies and actions by the space industry would do the most to promote the long term success and sustainability of an OOS infrastructure. These results are displayed and discussed in Case Study 3.

The results below have varying degrees of generalizability to wider sets of space systems and infrastructure concepts. The first two case studies, which cover global and non-global LEO data relay constellations, are perhaps the most generalizable to other systems, in that the solutions can inform most any space system which seeks a great deal of coverage and performance at the lowest cost, especially when utilizing disaggregated constellations of relatively small spacecraft. Meanwhile, the OOS case study is somewhat less generalizable, in that large, relatively standardized communications spacecraft at geosynchronous orbit were the primary focus during model development and validation. Of course, these types of spacecraft represent a large portion of commercial space activity and would likely be considered primary customers for any OOS infrastructure project (DARPA, 2014). The overall generalizability of the three case studies are discussed in greater depth in each section.

Case Study 1: LEO Communications and Data Relay Infrastructure with Global Coverage

The first case study analyzed possible constellation design points for a LEO communications and data relay infrastructure which had complete global coverage for its LEO customers. Of course, this requirement for global coverage placed significant constraints on the total size of the tradespace of possible designs. The analysis thus focused heavily on the comparison between a relatively small number of powerful spacecraft versus a much larger constellation of less powerful spacecraft, and using the tradespace exploration module to examine which constellations made the most economic sense in terms of their total cost and their ability to downlink customer data. Various crosslink and infrastructure interlink data rates were also explored, as it was hypothesized that the highest data rates may not necessarily be the best option.

Analysis Framework and Parameters

Forcing the tradespace exploration module to return only those designs which ensured global LEO coverage was a relatively simple exercise: any designs which were not calculated to have at least 100% contact time between a customer spacecraft and the infrastructure satellites had their utilities automatically set to zero. Thus, in the results presented in this first case study, only those designs having global coverage to a LEO customer were assigned non-zero utilities. "Global" in this context means that a reference LEO customer spacecraft will not experience any outages of service at any point in its orbit. Because of the way in which the model was developed—assessing LEO-to-LEO communications links between spacecraft, and then the number and duration of spacecraft passes over specified ground stations—it should be noted that it is not specifically constructed to calculated coverage over all terrestrial points. However, by assuring constant coverage to a customer in LEO, it is assumed that, by extension, the model also assures global ground coverage. Again, the model was not constructed to assess ground coverage, and thus any instances of "global" coverage refer to the infrastructure's visibility to a customer spacecraft in LEO.

Of course, other case-specific parameters were also input. The link-up penalty of two minutes was dropped, as it was assumed that the spacecraft would always be in communications with the infrastructure. Furthermore, to ease the computational load, only the medium class of customer was considered (the customer class modeled by the SkyBox constellation (SkySat Imaging Program of SkyBox Imaging, 2014)). This was considered a valid modeling parameter, especially since a global infrastructure would be available to all customers, and would increase the performance of each customer class similarly. Another crucial parameter included in this case was the way in which additional units were discounted in construction and ground costs. An aggressive unit-over-unit cost reduction factor was used in both cases: 85% manufacturing cost improvement for each additional satellite constructed, and 70% improvement in ground station costs for each additional satellite supported. These values were taken from established cost models, and were meant to maximize the way in which unit costs would decrease as more of the same model of infrastructure spacecraft are constructed, in addition to assuming that ground costs would not increase excessively simply because there are more spacecraft in the infrastructure (Fox, Brancato, & Alkire, 2008). It was theorized that without these discount factors, then the largest constellations would be unfairly estimated to be much higher in cost than is realistic given industry scale effects.

Only fully interlinked rings of infrastructure spacecraft in polar orbits were considered, with more spacecraft added in additional orbital planes. This is because it was assumed that at least one high latitude ground station was always available, such that the infrastructure would always be capable of downlinking customer data to this polar station (regardless of the customer spacecraft's relative location to the station or to the infrastructure). The polar station considered within the analysis of each design is the one located in Svalbard, Norway, as it was assumed that this location is one which would have terrestrial data links to customers anywhere else on the globe (refer to Figure 8 for a map of NASA ground stations). This then requires the assumption that ground capabilities are such that data can be sent from this polar station to anywhere it is needed on the globe, but as this thesis is only concerned with the service provided to a LEO spacecraft, this was considered a valid and necessary assumption. Of course, other ground stations would also likely be used to downlink from the infrastructure, but none would have the same constant passes of infrastructure spacecraft like a polar station would. In many ways these parameters for global coverage borrow heavily from the design of the Iridium constellation

(Iridium Global Network: Satellite Constellation, 2015), but rather than ensure global ground coverage, the infrastructure designs in this tradespace were analyzed on their ability to provide constant coverage to a reference LEO customer.

Finally, to simplify the tradespace and reduce the computational load, all constellations were designed to have an altitude of 800 km, inclination of 95 degrees, and satellite design life of five years. These choices were informed and derived from the results of the initial validation (discussed in the methodology above), which found that an altitude of 800 km represented a beneficial compromise between the increased visibility of even higher altitudes and the lower power requirements of infrastructures situated at lower altitudes (and therefore closer to customer spacecraft and ground stations). In addition, a polar orbit of 95 degrees for all infrastructure configurations was deemed necessary for complete coverage, and a relatively short design life of five years followed from the assumption that a large constellation such as this would utilize relatively simple, short-lived, standardized spacecraft. Also, only crosslink ranges of 2000, 2500, and 3000 km were analyzed in the tradespace, as it was assumed that shorter ranges would be unrealistic for a globally-linked infrastructure, while forcing larger ranges would lead to spacecraft which were oversized for this particular mission and constellation type (this assumption was also supported in the initial testing of the model). Finally, the utility function was adjusted to only include total daily downlinked data, as all the other metrics in the function (such as percent contact time, average outage, maximum outage, etc.) would be completely satisfied by the consideration of only designs providing global coverage. Thus, it is assumed that a higher performing global infrastructure design is one that will simply downlink more data per day than another.

Performance and Cost Estimates

The tradespace developed for this analysis differed slightly from what is typically found in tradespace exploration, in that there is no definite "knee" in the Pareto front of optimal designs. This is to be expected, as the nature of a global LEO infrastructure dictates that, all else being equal, increasing the total number of spacecraft will increase the total amount of data which can be downlinked and therefore increase the overall performance of the system. Where this analysis

does provide insight, however, is where the infrastructure designer can expect to see a diminishing rate of increasing performance as number of spacecraft increases, as well as which solutions are likely to be "good enough" for the purposes of most LEO customers.

The results shown in Figure 27 are fully indicative of this. Keeping in mind that all solutions in this set provide round-the-clock coverage for LEO coverage, it is the designs along the indicated Pareto front which are configured in such a way as to provide the maximum downlink performance for the minimum cost. Key designs along this front are displayed in Table **3**, and a closer look at the Pareto front of designs is shown Figure 28. Also included in the list of interesting solutions is Design 12, which is the largest, most disaggregated constellation of lowest-power, lowest-mass spacecraft. This was considered to be a good reference constellation, given the greater interest now being shown in highly disaggregated space systems (U.S. Air Force Space Command, 2013). Graphical representations of a few of the key designs from Table 3 are then shown in Figures 29-32.

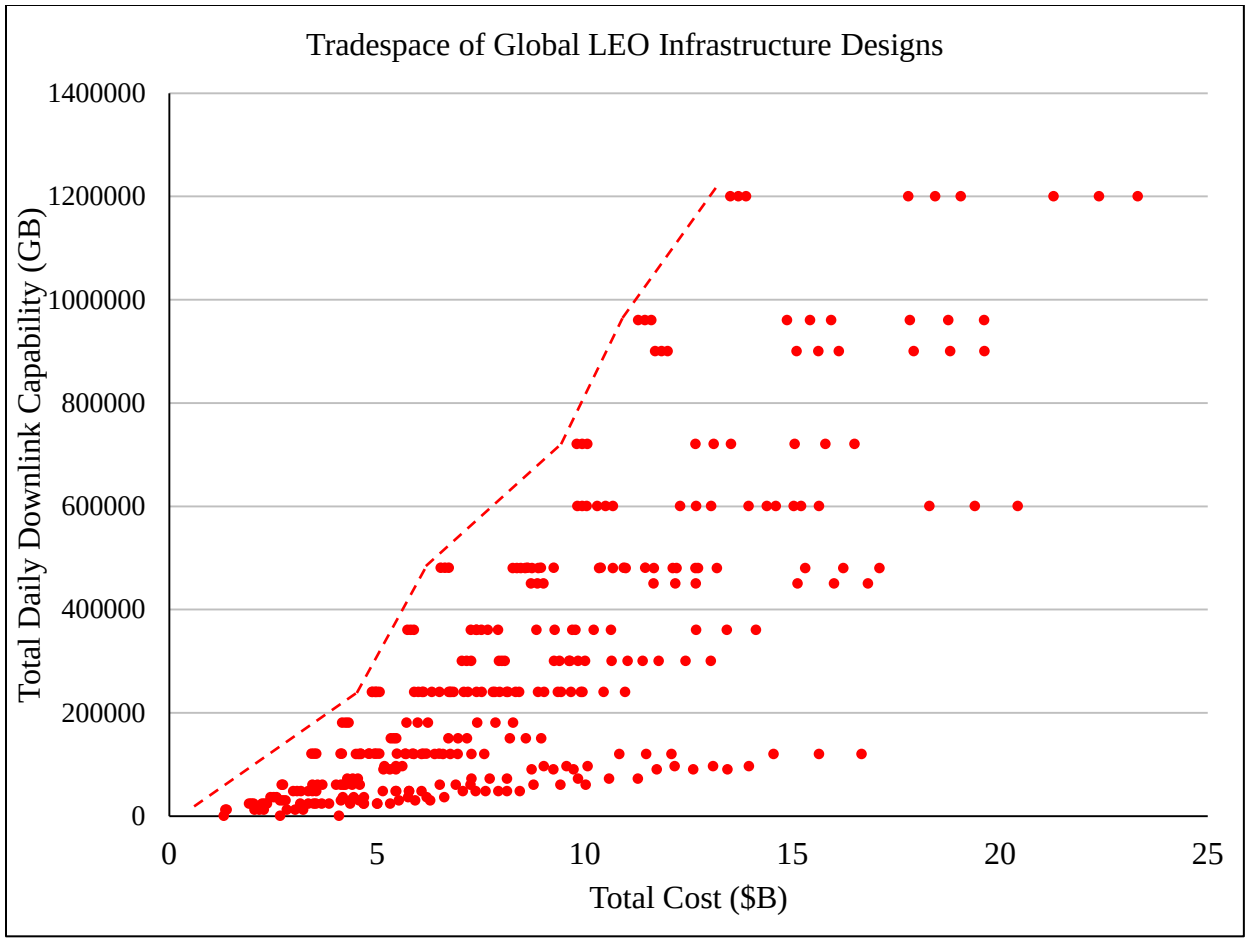


Figure 27. Tradespace of LEO infrastructure designs which satisfy the requirement for global coverage.

Design Number	Satellites per Plane	Number of Orbital Planes	Crosslink Data Rate (Mbps)	Interlink Data Rate (Mbps)	Maximum Crosslink Range (km)
12	32	10	10	10	2000
146	16	4	10	50	2500
43	24	8	10	50	2000
44	32	8	10	50	2000
147	24	4	10	50	2500
148	32	4	10	50	2500
217	8	4	10	10	3000
48	32	10	10	50	2000

Design Number	Cost (B)	Total Daily Downlink Capability (MB)	Satellite Mass (kg)	Satellite Power (W)	Notes:
12	6.399	1.200E+08	248	41	Largest smallsat constellation
146	3.491	1.201E+08	710	333	
43	7.267	3.601E+08	705	161	
44	8.577	4.799E+08	703	108	
147	4.252	1.802E+08	705	175	
148	4.981	2.402E+08	704	122	"Knee"
217	1.388	1.201E+07	255	312	Minimum infrastructure
48	10.314	5.999E+08	703	108	Largest smallsat constellation, with higher data rate

Table 3. Major solutions of interest in the global coverage tradespace; all but Design 12 exist on the Pareto front, indicating greatest performance to cost.

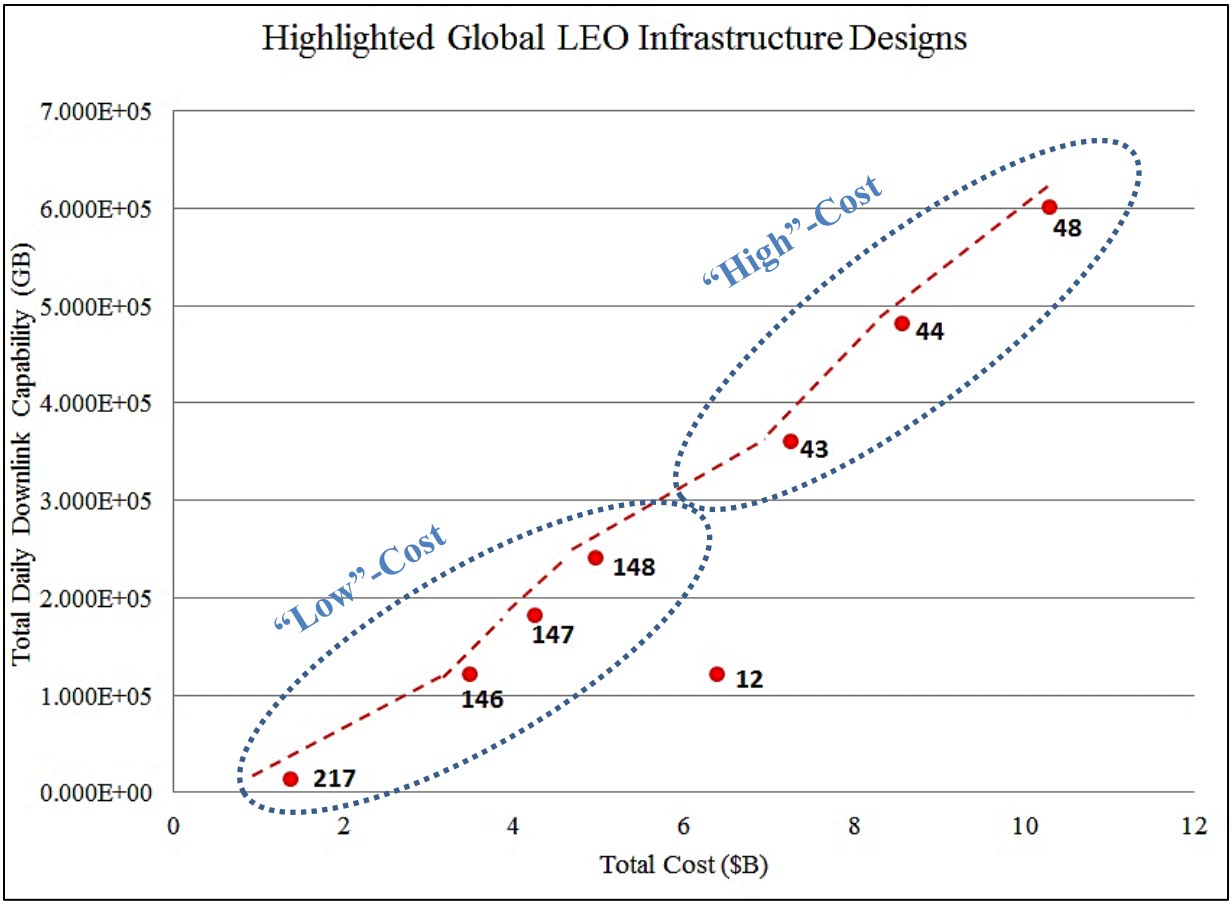


Figure 28. Closer look at the Pareto front of optimal designs, with the design numbers taken from Table 3.

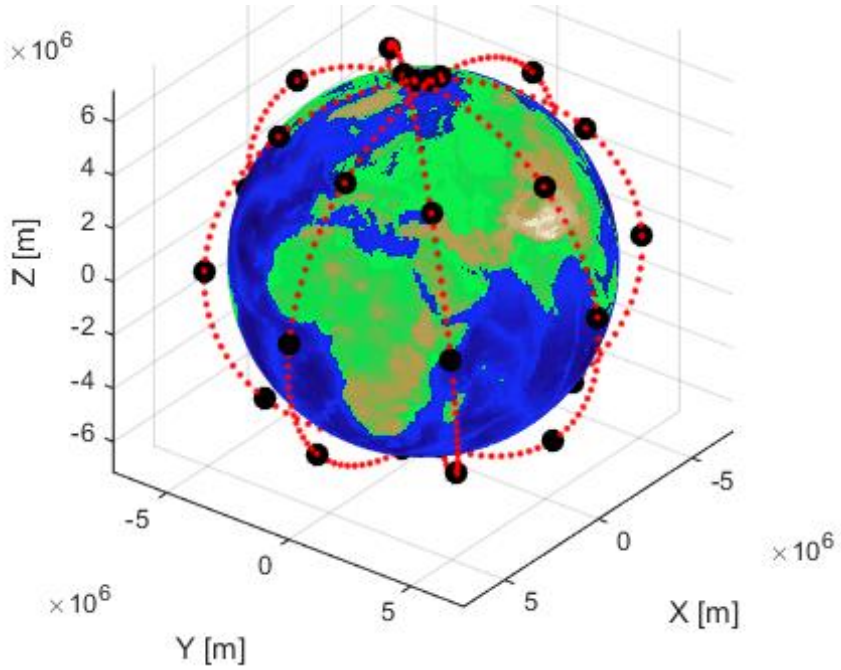


Figure 29. Global infrastructure design 217, which is composed of four orbital planes with eight infrastructure craft per plane (32 total spacecraft).

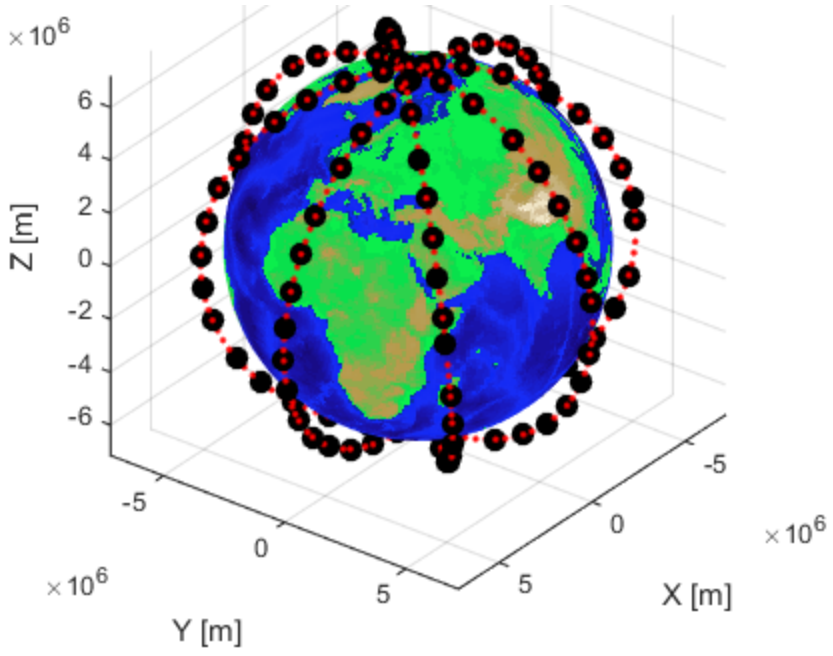


Figure 30. Global infrastructure design 148, which is composed of four orbital planes with 32 infrastructure craft per plane (128 total spacecraft).

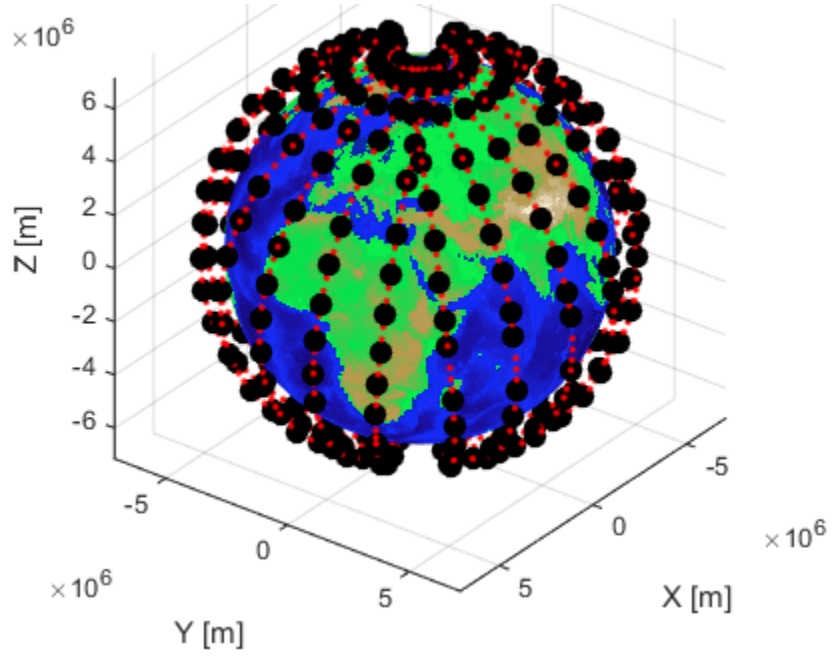


Figure 31. Global infrastructure designs 12 and 48, which are composed of 10 orbital planes with 32 infrastructure craft per plane (320 total spacecraft).

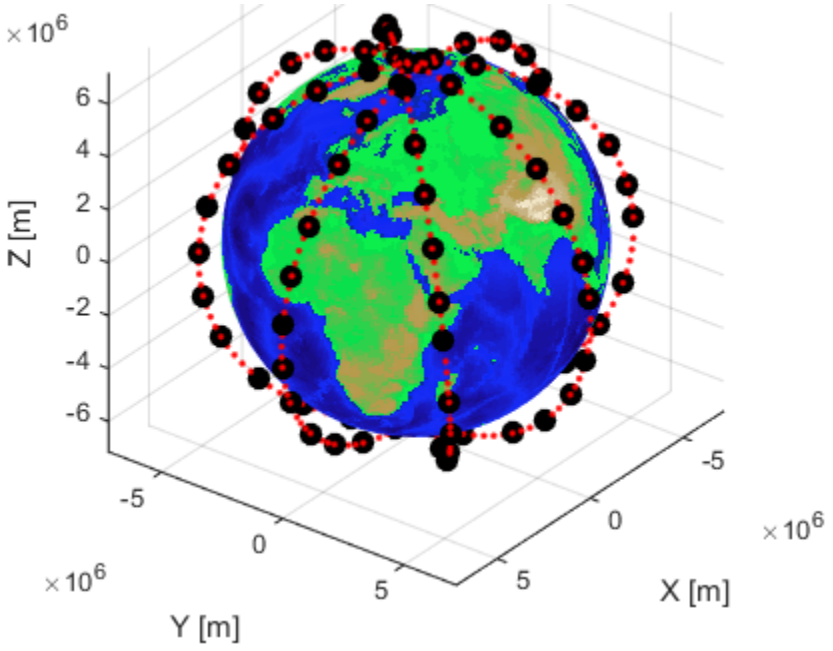


Figure 32. Global infrastructure designs 147, which is composed of 4 orbital planes with 24 infrastructure craft per plane (96 total spacecraft).

Discussion

The first major observation from this set of results is that there exists a set of relatively "low-cost" solutions, followed by a set of relatively "high cost" solutions (highlighted in Figure 28). The low-cost solutions include designs 217, 146, 147, and 148, and these all share the common design point of four orbital planes while increasing the number of spacecraft per plane (with design 217 being the smallest possible constellation size which satisfies the requirement for global coverage). All of these designs are estimated to cost less than $5 billion, and by increasing the number of spacecraft per plane, each individual craft can be made progressively smaller and less powerful. Then, after design 148, the next solution on the front costs over $2 billion more, which in addition to representing an increase in the number of orbital planes in the infrastructure also marks the beginning limitations to performance increases simply through the addition of more spacecraft. The last design on the front, 48, demonstrates exceptionally high downlink performance, but also costs over $10 billion.

Furthermore, all designs favored relatively low-data rate crosslinks to customer spacecraft, but then high interlinks between infrastructure spacecraft. This shows that, with the customer always in view, crosslink speed can be minimized without sacrificing a great deal of performance, as there is less of a "rush" to send all the data through the infrastructure. Meanwhile, interlinks can be designed for much higher data rates, both because infrastructure spacecraft transmitting through an orbital ring are always close to one another (thus less power is needed for a higher data rate), and because higher data rates in the interlinks means that more data can be transmitted to the ground at any given moment. Of course, there are limits to desired data rates in the interlinks; after all, an infrastructure spacecraft performing two-way communications with two customers on 10 Mbps crosslinks would theoretically never need more than a 40 Mbps interlink to handle all possible data (this tradespace considered interlinks of 10, 30, 50, and 100 Mbps). As these results indicate, though, a cost-effective, high-performance global infrastructure is likely to be one with relatively high network speeds and relatively low speed customer links.

The results also indicate that, under the established cost models and aggressive discounting used, the best solution may actually lie in the middle ground of designs. That is, the most optimal designs were those which used spacecraft which were moderate in size and power, and

constellations which were composed of neither very numerous sets of small spacecraft nor a relatively small number of large satellites. This is shown primarily through the distance of design 12 from the Pareto front, as this design is the most disaggregated, lowest powered solution analyzed (consisting of 10 planes of 32 satellites, for a total of 320 spacecraft). Indeed, this design did almost no better in total downlinked data than design 146 (hence why it is so far off the Pareto front), which in addition to being a fifth the size was also nearly half the total cost (see Table 3). Thus, infrastructure designers should not assume that greater disaggregation of assets will automatically equal increased performance for a given price point.

Ultimately, though, the most appropriate infrastructure design will depend largely on customer expectations and mission requirements. While this analysis indicates that there are several valid global infrastructure configurations at relatively low total costs, this does not mean that these designs are the perfect solutions to all situations. In fact, it is assumed that the final design for a LEO communications and data relay infrastructure will hinge primarily on the number of customers expecting to be serviced on a daily basis. The indicated "knee" of the front, design 148, could service at most 256 customers at a time (as there are four planes of 32 satellites, with each satellite communicating with two customers; however, this would require several customer spacecraft to be near the poles). If more or less customers are expecting to be serviced, though, then the design should be adjusted accordingly up or down the Pareto front. Furthermore, if it is expected that customers will demand higher data rates—for applications like real-time video links or high resolution imagery—then the infrastructure spacecraft will necessarily have to be designed with higher power and more robust communications capabilities than the 10 Mbps links called for in this analysis.

These results can be generalized to essentially any proposed constellation of LEO spacecraft, especially those that seek to achieve cost savings and enhanced performance through the use of many small, standardized spacecraft. While it may be tempting to believe that, as technologies progress and become increasingly small, then the performance of a system composed of many small spacecraft will still be comparable (or even better than) that of a few larger, monolithic satellites. Instead, for most systems there is likely a happy medium in constellation design, which captures the benefits of both disaggregation and miniaturization as well as recognizes the inherently higher performance of larger, more powerful spacecraft. In addition, recall that this

analysis considered only basic, point-to-point RF communications links, in addition to using industry standard cost models. This may lead the results to be less generalizable to cases where different technologies or cost structures are in place. For instance, performance would likely be far greater through the use of more complex RF transmitters and modulation schemes, or through the use of optical communications. Indeed, laser communications would likely be highly beneficial to overall data relay capabilities, especially between the stable, easy-to-acquire infrastructure spacecraft within the same orbital plane. Furthermore, even greater per-unit cost savings would likely be achieved through the use of commercial-off-the-shelf (COTS) components, or through greater automation of ground systems (Wertz J. , 2011). Such designs should be considered in any future research on similar LEO infrastructure concepts.

Case Study 2: Optimal LEO Communications and Data Relay Infrastructure Designs without Global Coverage

In this section, designs for a communications and data relay infrastructure are again analyzed, but without the requirement of global coverage. That is, in a scenario where customers accept a lack of worldwide, round-the-clock coverage, what are the optimal designs for a communications and data relay infrastructure, granting the greatest performance at the lowest costs? It was reasoned that an infrastructure which maintains its own inter-links but which is not always in communication with customer spacecraft may still be capable of generating significant performance enhancements for its customers, in the form of greater downlink capability and greater access to ground stations than the customers can achieve on their own.

Analysis Framework and Parameters

Unlike the global coverage case—in which the customer is always in contact with a ground station via the infrastructure—a non-global infrastructure case forces the metrics of percentage contact time, maximum outage time, and average outage time to become very important to determining overall utility. Thus, as a first step in the analysis, a small selection of infrastructure designs was used to establish a general performance baseline, based on the allowable communications range between customer spacecraft and the infrastructure constellation. This was done in order to determine which crosslink ranges made the most sense as design ranges for a larger variety of infrastructures. By knowing beforehand where diminishing returns occur (i.e. where one begins to see less technical and economic benefits from increasing the power and receptivity of the infrastructure), the tradespace exploration could be constrained to a more manageable computational load. This analysis used five reference infrastructures: one plane of 8 spacecraft, one plane of 12 spacecraft, one plane of 16 spacecraft, two planes of 8 spacecraft each, and 2 planes of 16 spacecraft each (these designs are displayed in Figure 33 below). All of the planes were at 63.5 degrees inclination and an 800 km altitude. The reference spacecraft used was Landsat-8, which orbits at 98.2 degrees inclination and a 705 km altitude (U.S. Geological Survey, 2014). Furthermore, to present an average case scenario, all of the infrastructure planes were offset from Landsat-8's plane by 45 degrees (i.e., the infrastructure's

visibility to the customer would not be at its worst or best, simply due to the beginning position of the customer's and infrastructure's orbital planes). The results of this initial exploration are shown below.

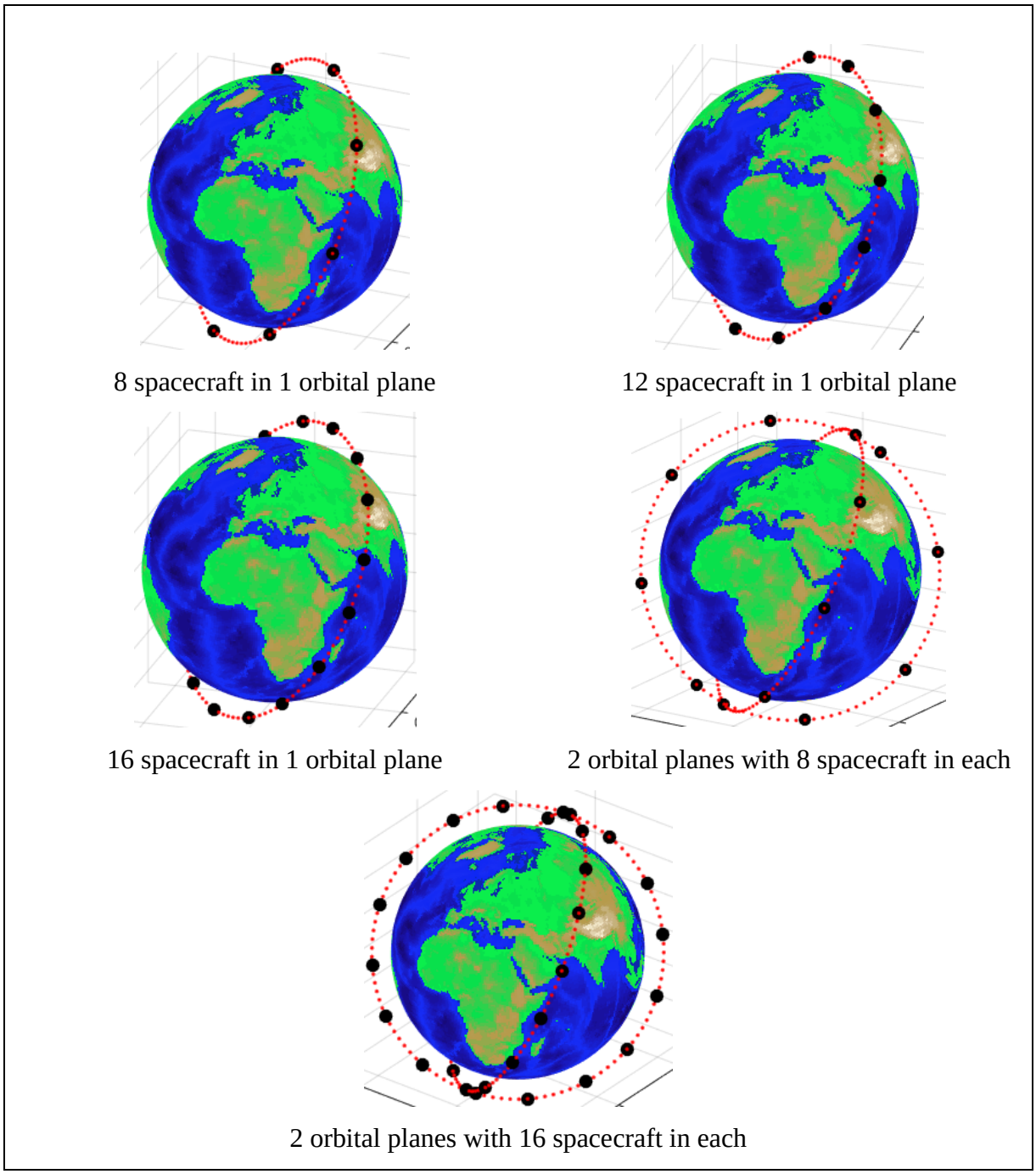


Figure 33. The five reference infrastructure configurations used for initial testing and validation of the tradespace.

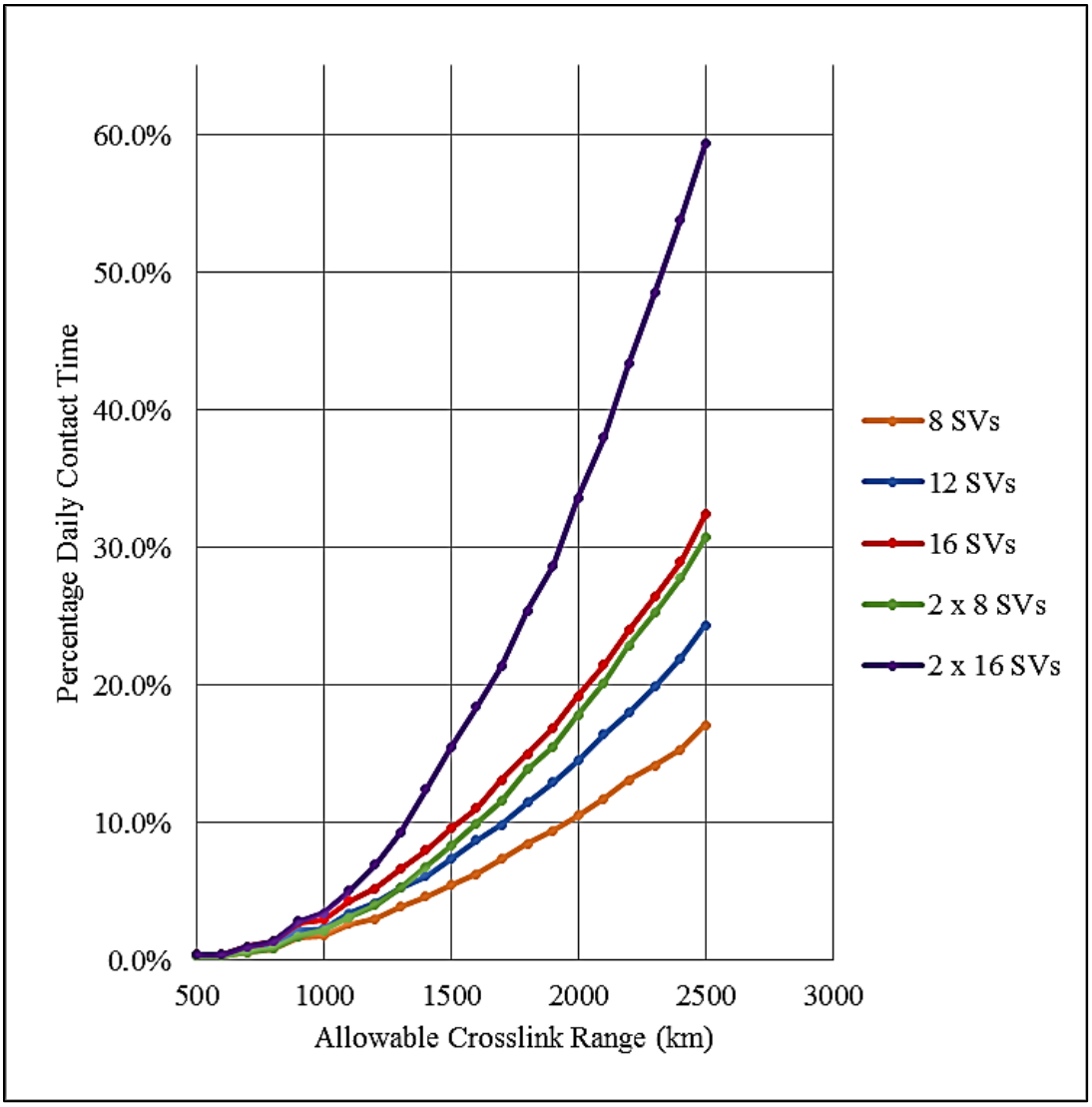


Figure 34. Percent contact time of various infrastructure configurations to a reference customer spacecraft.

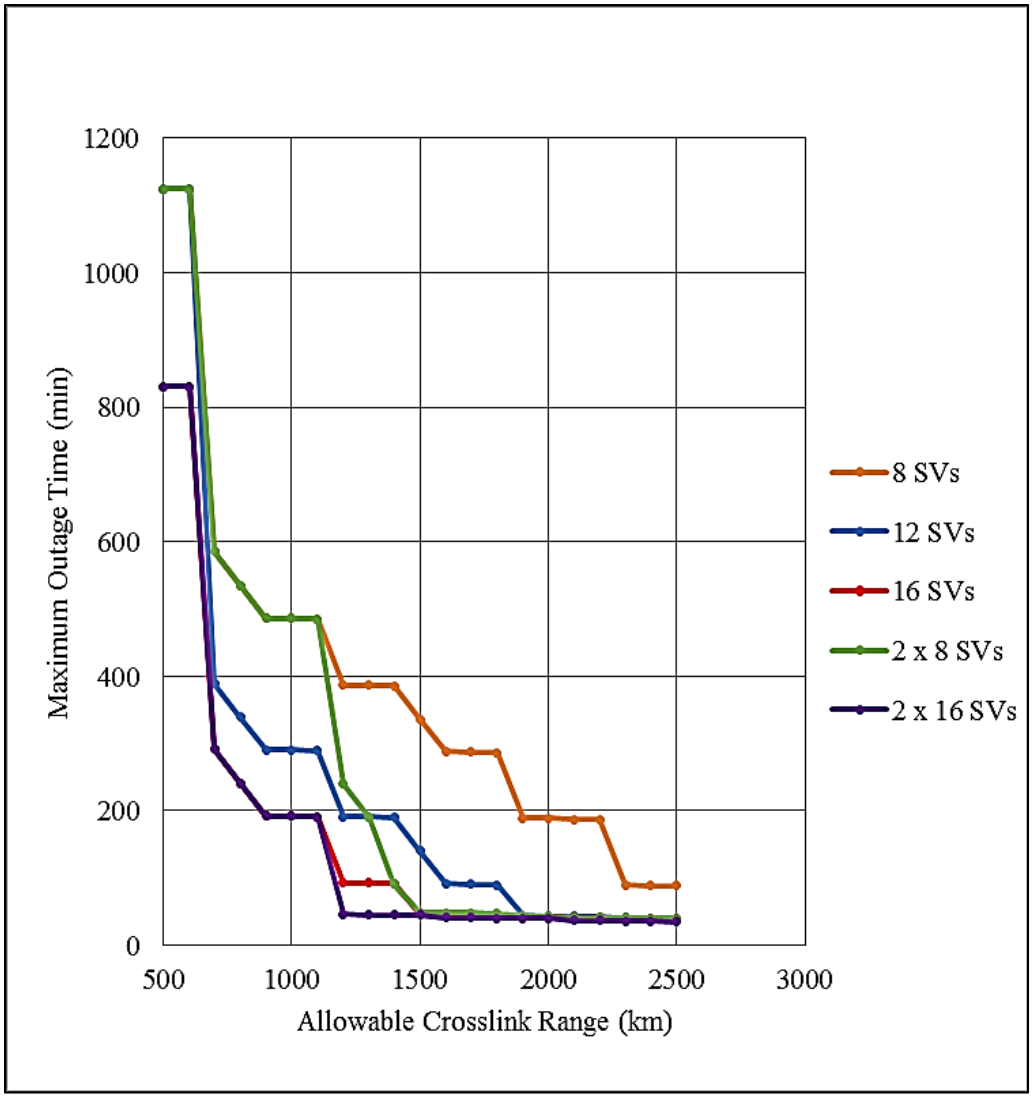


Figure 35. Maximum outage time of various infrastructure configurations to a reference customer spacecraft.

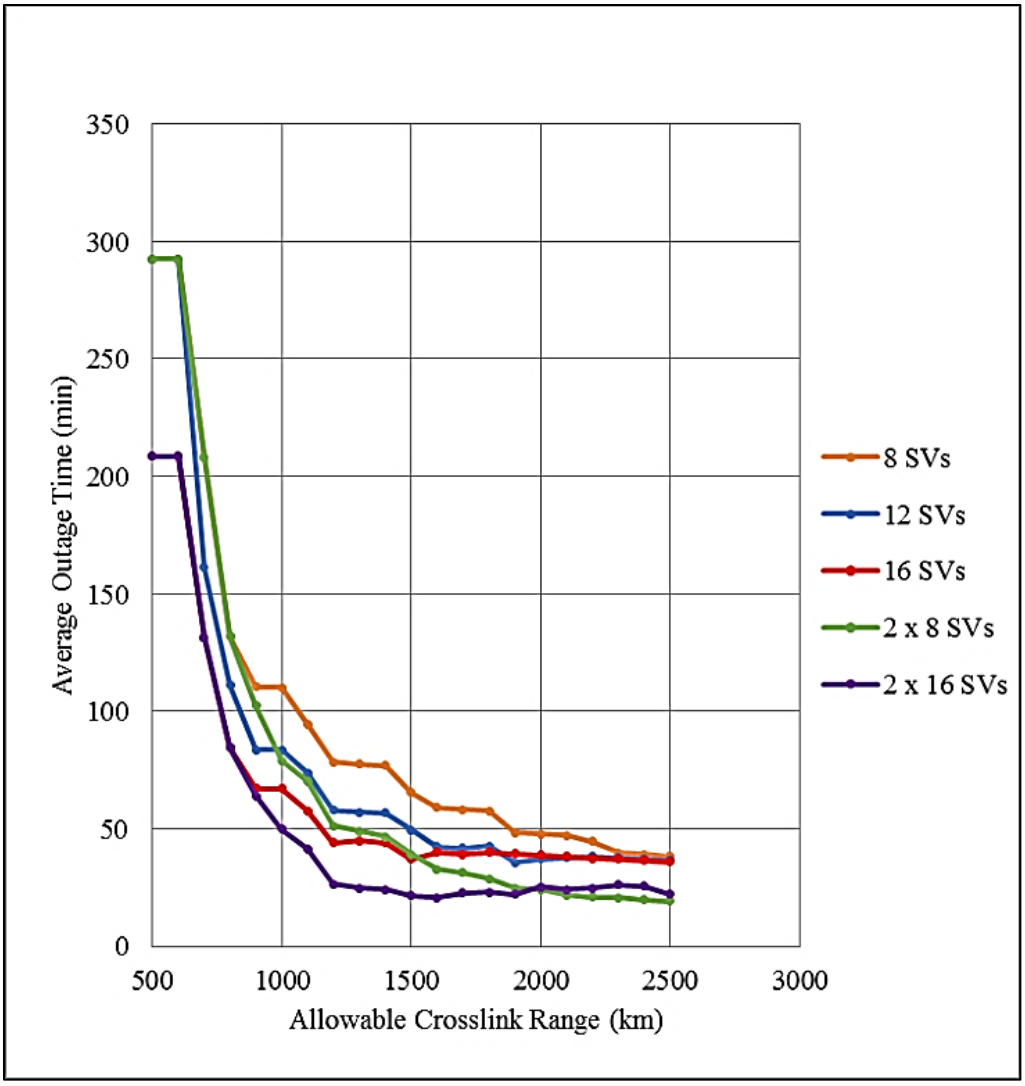


Figure 36. Average outage time of various infrastructure configurations to a reference customer spacecraft.

From this initial analysis, it is apparent that performance diminishes significantly after an allowable communications range of approximately 2000 km. While percentage contact time (shown in Figure 34) naturally increases steadily as crosslink range is increased, the average and maximum outage times do not significantly improve beyond this range (shown in Figure 35 and Figure 36). Knowing this, the tradespace of designs could be significantly reduced, as it is clear that designers would most likely be concerned with infrastructure-to-customer crosslink ranges of 1000 to 2500 km (knowing that designs below this range are of too poor performance and

designs above this range would become prohibitively power-needy and expensive, in addition to not offering significant benefits).

This initial analysis also provides some preliminary conclusions on the performance of infrastructure. From Figure 34, which shows percentage contact time between customer spacecraft and the infrastructure, it is evident that most infrastructures having crosslink ranges of 2000 km will be in contact with customers at least 10% of the time each day. Assuming that the infrastructure is always in contact with a ground station, this 2+ hour allotment of contact time represents quite an increase in contact time than is achievable by just the customer spacecraft, which may only pass over a ground station a few times each day (with LEO passes typically lasting less than 20 minutes (Wertz, Everett, & Puschell, 2011)). Additionally, the maximum outage times displayed in Figure 35 shows that, at a 2000 km crosslink range, maximum outage is typically less than an hour for moderately sized infrastructures, and just over 3 hours for the smallest infrastructure design. In addition, at this same range, average outage times are well below one hour for all sizes of infrastructure, as shown in Figure 36. Thus, it is easy to see that even if customers do not have constant communications with the infrastructure, they still enjoy very reliable and regular contact with the infrastructure, and by extension ground stations, throughout the day.

Performance and Cost Estimates

The next step was to complete the full analysis of orbital infrastructure's contributions to the capabilities of LEO space systems. As described in the Tradespace Exploration section of the Methodology, the utilities of over 10,000 infrastructure configurations were calculated with respect to three types of customers: a highly disaggregated constellation of very small imagery spacecraft, represented by Planet Labs Flock-1 (Niles, 2014); a moderately disaggregated constellation of small-to-medium sized imagery spacecraft, represented by SkyBox's SkySat constellation (Skybox Imaging, 2015); and a large monolithic imagery spacecraft, represented by NASA's Landsat-8 (U.S. Geological Survey, 2014). By comparing the utility gains of these three separate cases, greater insight was gained as to the best possible infrastructure design for all LEO customers.

The results of the tradespace exploration are shown in Figure 37 and Figure 38. A mask was applied to filter out designs producing less than 20% improvement in a participating spacecraft's total daily data downlinked (to enhance the visibility of results and filter out already poor designs). Also, the initial analysis allowed the tradespace to be constrained to crosslink ranges of 2000 km and below, since performance increases were shown to largely subside after this point. The blue points in the figures below correspond to designs which can satisfy the minimum performance improvements (20% greater data downlink) of all three customer classes (Landsat-8, SkyBox, and Planet Labs). The green points correspond to those designs which can satisfy both SkyBox and Planet Labs (the spacecraft of these two constellations are not equipped to return as much daily data as Landsat-8; therefore, the infrastructure designs which do satisfy their needs but not Landsat-8's are left unmasked). Finally, the red points correspond to those designs which only provide at least 20% improvement for Planet Labs spacecraft (due to the same reasoning, that the Planet Labs satellites are each not designed to return as much total data as SkyBox). In addition, a selection of the most optimal designs is displayed in Table 4.

Customer	Utility	Cost ($M)	Incl. (deg)	Alt. (km)	Total Satellites	Number of Planes	Crosslink Data Rate (Mbps)	Interlink Data Rate (Mbps)	Max Crosslink Range (km)	Total Daily Data (GBs)	Priority Data (GBs)
Landsat-8, SkyBox, and Planet Labs	0.4588	1295.7	95	800	12	1	90	10	1500	12110.6	232.5
	0.5463	1303.2	95	1000	12	1	90	10	1500	15153.8	277.1
	0.5682	1424.4	95	1000	14	1	90	10	1500	17717.3	324.0
	0.6127	1484.3	95	1000	16	1	70	10	1500	20244.1	370.2
SkyBox and Planet Labs ONLY	0.3786	748.3	95	800	8	1	20	10	1500	3622.6	154.8
	0.4651	753.6	95	1000	8	1	20	10	1500	3801.6	184.8
	0.4916	820.9	95	800	10	1	10	10	1500	2261.5	193.2
	0.5681	970.8	95	1000	10	1	30	10	1500	7119.7	230.8
	0.5903	950.6	95	1000	12	1	10	10	2000	2849.8	277.1
2-plane solutions	0.6466	1835.1	95	1000	10	2	30	10	1500	14346.9	930.0
	0.6686	2228.0	95	1000	14	2	30	10	1500	20113.3	1303.8
	0.6796	2420.1	95	1000	16	2	30	10	1500	22989.3	1490.2

Table 4. A selection of the most optimal infrastructure designs, with their orbital configurations and performance characteristics listed.

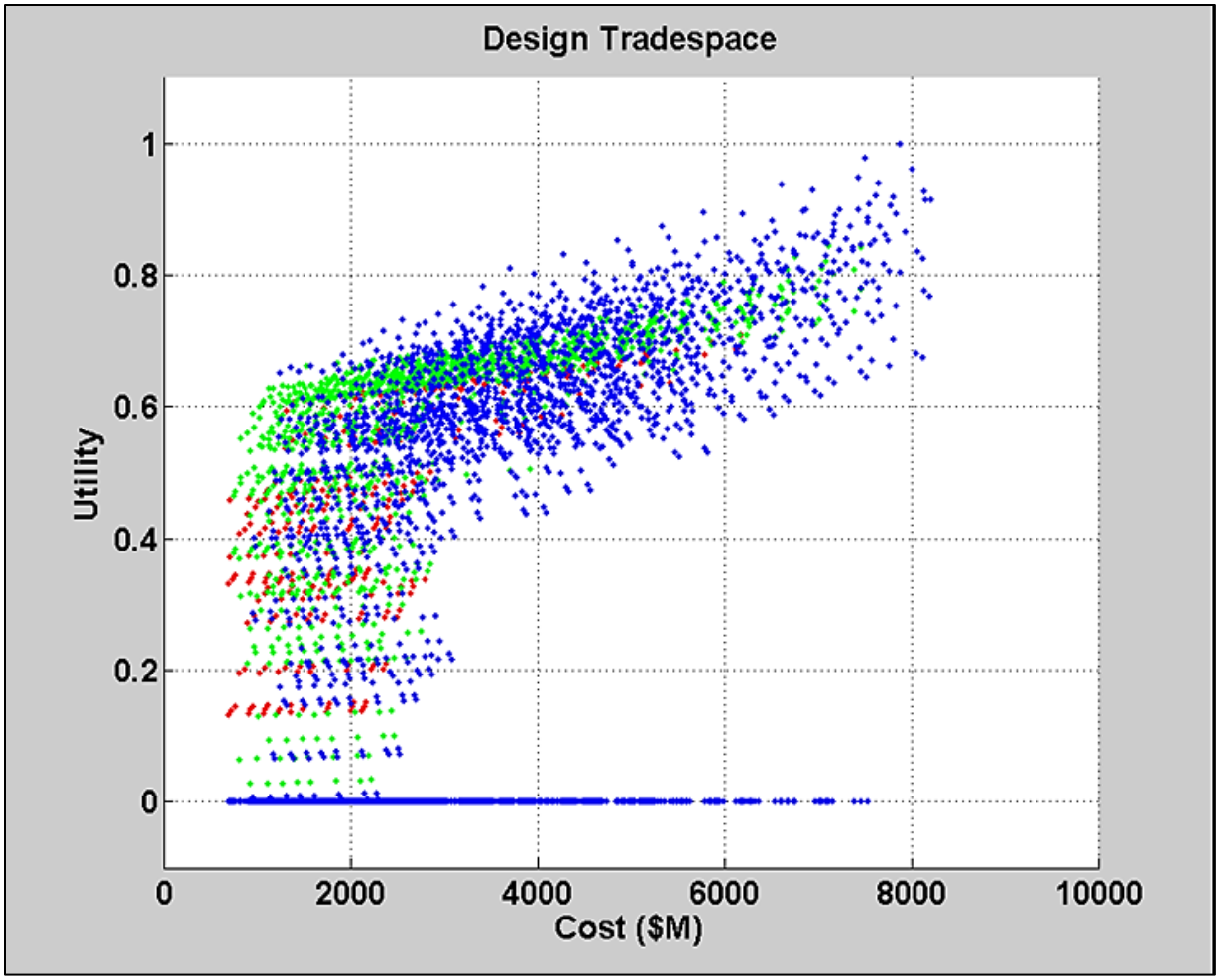


Figure 37. Full tradespace of infrastructure designs. Blue points represent designs which satisfy minimum performance requirements for all three cases (Landsat-8, SkyBox, and Planet Labs). Green points represent those which satisfy SkyBox and Planet Labs, and red points satisfy just Planet Labs.

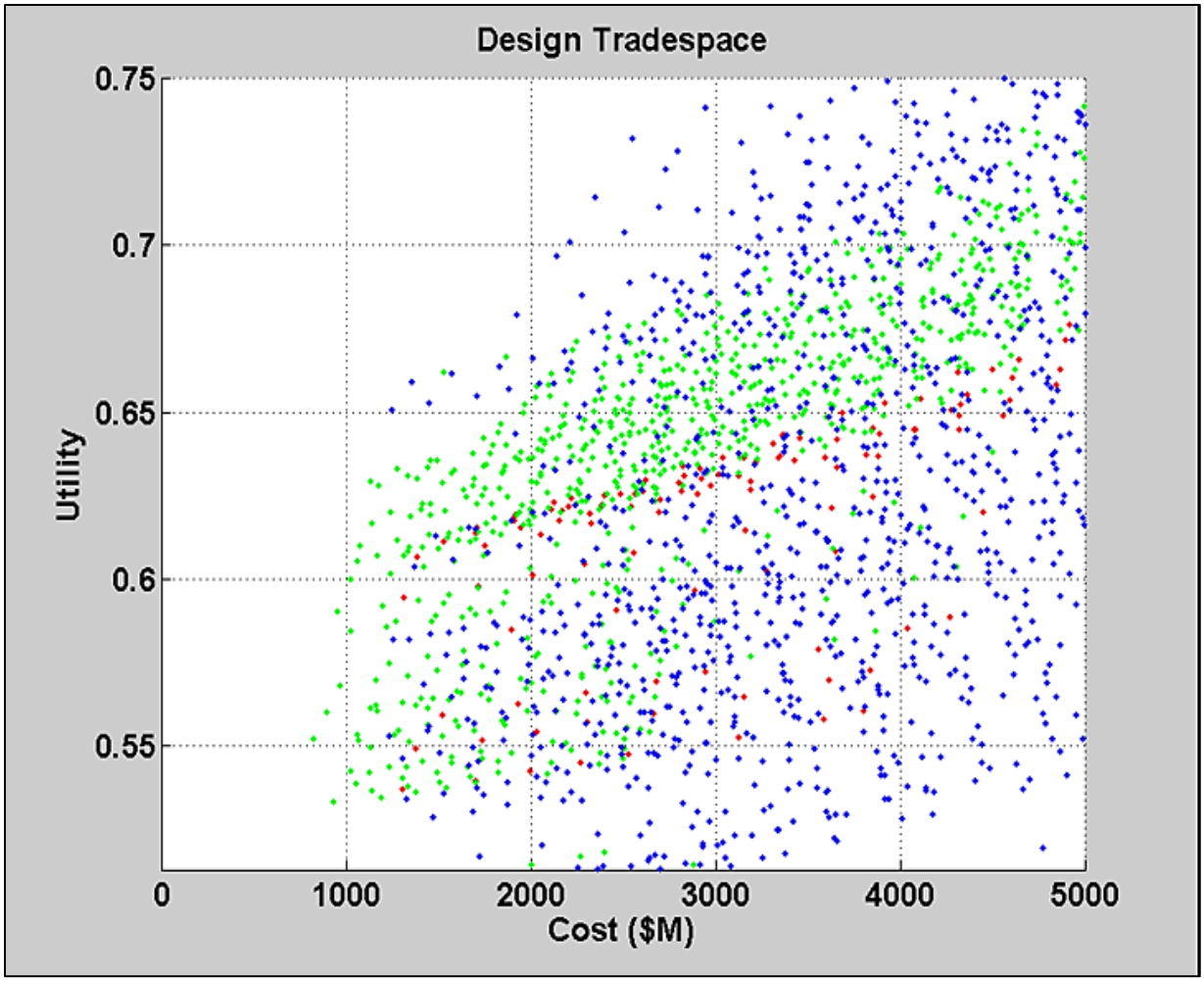


Figure 38. A closer look at the "knee" of the Pareto front of infrastructure designs. It is in this region that the optimal designs are found, based on both total utility and overall cost.

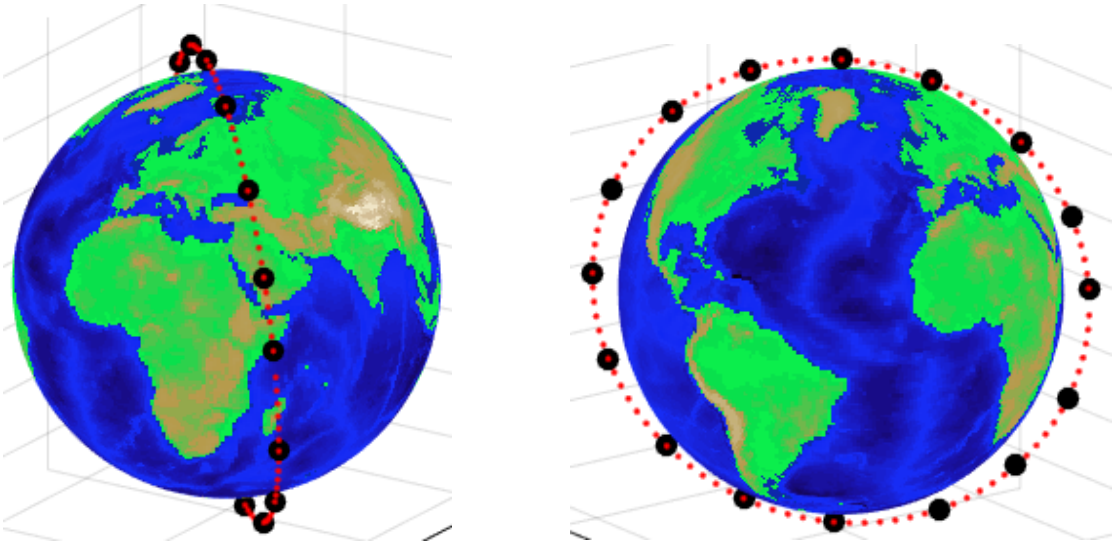


Figure 39. Infrastructure configuration of a single orbital plane containing 16 total satellites (both pictures display the same infrastructure, from different vantage points).

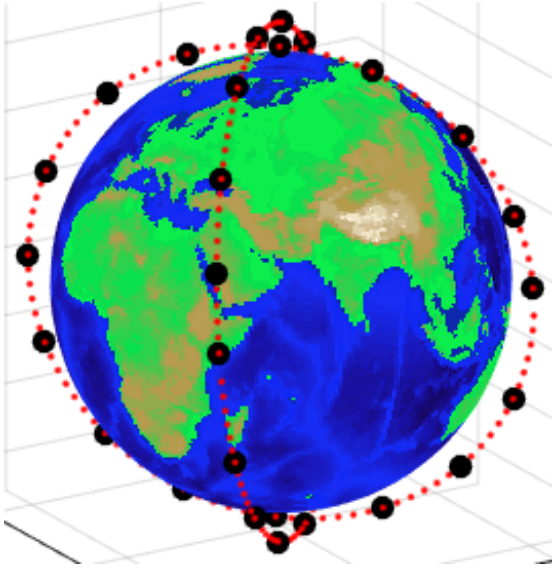


Figure 40. Infrastructure configuration of two orbital planes containing 16 satellites each, for a total constellation of 32 satellites.

Discussion

A closer look at the top-performing infrastructure designs revealed a few key trends. First, the best infrastructures were almost always in polar orbits. These orbits both maximized global coverage as well as yielded greater time in sunlight for the infrastructure spacecraft, allowing them to have decreased solar panel areas and power storage requirements (which subsequently reduced mass and cost). This would also be very likely to benefit any imaging spacecraft that are in sun- synchronous polar orbits, as the similar orbits of these customers and the infrastructure would lead to increased contact time (however, this was not a consideration in this analysis, as the results were intended to extend to more customers than simply those involved in optical imagery).

In addition, higher altitudes were favored, as this tended to increase the infrastructure's time in view to both ground stations and customer spacecraft. This result represents one of the major trade-offs involved in the analysis, in that higher orbits did lead to longer contact times, but also increased the communications distance, which increased the power and mass requirements of each infrastructure spacecraft. On the other hand, if contact time is increased, then the data rates required to transmit an equivalent amount of data are subsequently decreased, which then decreases power requirements. Within this analysis, the higher altitudes and their increased time in view to customers and ground stations ultimately won out over lower, faster data rate infrastructure designs.

The best designs also tended to have relatively low-powered, low-data rate inter-infrastructure links (typically with data rates of only 10 Mbps), but then would have much more powerful crosslinks to customer spacecraft. This is indicative of the fact that the infrastructure can transmit data within its own network continuously, and so is in less of a "rush" to send data through the infrastructure and down to ground stations. Once a customer spacecraft comes into view, though, there is a desire to maximize the total data transmitted, which requires a higher data rate and higher transmitter power. Also, recall that the model was built to oversize transmitters, such that if any infrastructure spacecraft fails then those in front of and behind the failed craft would still be able to bridge the communications gap. In addition, the model assumed that links would be transmitting 100% of the time, which would likely not be the case in real operations. Thus, the relatively low-powered links being favored in this analysis are likely

low-end estimates of the infrastructure's actual capabilities (i.e. the actual transceivers used by the infrastructure would exhibit higher performance than the baseline design values).

Another interesting result of this analysis was that an infrastructure configuration of eight satellites in one plane was comparable in price to one of 16 satellites in a single plane. This was due to the fact that the more infrastructure satellites there were in a plane, the less powerful each had to be to maintain inter-infrastructure links, with this reduction in power leading to corresponding reductions in satellite mass and cost. The cost model used for this analysis also included replication cost reductions for each additional craft constructed, following the assumption that all infrastructure craft were of the exact same design. In fact, given the conservative estimates used in the cost model, it may actually be overestimating the total cost required to build many smaller, simpler infrastructure spacecraft; indeed, even greater cost savings could be achieved through further disaggregation of the infrastructure (e.g. constellations of 30+ small spacecraft).

$$
\begin{aligned}
Design\ Utility\ = \ \frac{1}{9}\ [\ & (\%\ contact\ time\ with\ ground\ stations) \\
+\ & (Normalized\ average\ outage\ time\ with\ ground\ stations) \\
+\ & (Normalized\ maximum\ outage\ time\ with\ ground\ stations) \\
+\ & (Normalized\ total\ contact\ time\ per\ day\ between\ infrastructure\ and\ ground) \\
+\ & (Normalized\ maximum\ outage\ time\ between\ infrastructure\ and\ customer\ spacecraft) \\
+\ & (Normalized\ average\ outage\ time\ between\ infrastructure\ and\ customer\ spacecraft) \\
+\ & (\%\ of\ "priority"\ data\ transmitted\ each\ day) \\
+\ & (\%\ of\ data\ downlinked\ from\ customer\ to\ ground\ per\ day) \\
+\ & (\%\ of\ combined\ daily\ contact\ time\ with\ ground\ stations)\ ]
\end{aligned}
$$

Equation 1. Utility function used to compare utility of infrastructure designs. The "score" for each metric is a value between 0 and 1, extrapolated linearly from maximum and minimum values for each metric, and each metric is assigned an equal weighting in the final calculation of design utility.

The type of data being transmitted also represented a major trade-off analyzed by the tradespace exploration module. The utility function applied equal weighting to all metrics, and it

was found that optimal solutions almost always sacrificed "priority" data in favor of total daily data downlinked. In fact, the dominated solutions of the tradespace were those in which the quantity of priority data approached that of total data. This could change dramatically, though, if different weightings are applied in the utility function (displayed again in Equation 1 above); that is, if customers prefer the ability to downlink data instantaneously over simply downlinking more data each day, then the infrastructure designs which allow for this would become the optimal solutions.

Finally, one of the major results from this analysis involves the differences between large flagship missions and smaller, less-tested commercial architectures. This analysis shows that a LEO infrastructure may not be particularly valuable to large, flagship missions like NASA's Landsat-8, which already have all the necessary power and ground infrastructure to perform quite well on their own. However, there likely is a significant business case for an infrastructure meant to provide services to smaller firms; ones which do not have the resources of NASA and are attempting to use new, relatively untested technologies and architecture designs. For these smaller players, profitability resides in their constellations' ability to downlink large amounts of data daily, and an infrastructure which grants significant capability boosts in this area would likely be a very valuable aid to the firms' success. It is also worth noting that the infrastructures optimized to benefit these smaller customer classes also cost much less, with all of the designs in the "SkyBox and Planet Labs Only" section of Table 4 each costing less than $1 billion.

These results can easily be generalized to several similar space systems, especially those which are relatively new and untested. Rather than view infrastructure from the perspective of "all or nothing," these results show that such projects can have less than constant coverage and still exhibit relatively high performance over a customer's standalone capabilities. Indeed, many small, underserved customers would likely be quite satisfied by an infrastructure which ensures they can downlink data at least every hour. Like the first case study, though, it should be remembered that this analysis does not consider highly directional antennas or optical communications, which in addition to increasing system performance could shift the ways in which the infrastructure is utilized. For general constellation design, however, these results can certainly provide insights as to the optimal solutions, considering both cost and performance.

Case Study 3: Assessment of On-Orbit Servicing Effects

As discussed in the Methodology, a key assumption surrounding OOS is that it will significantly decrease the space industry's severe risk aversion, due to its ability to rescue failed spacecraft as well as refuel and/or upgrade spacecraft throughout their service lives. In other words, it is assumed that if OOS is implemented on a wide scale, then spacecraft designers would no longer feel the pressures associated with launching untouchable, difficult-to-modify satellites; instead, they could actually have a degree of access to their on-orbit assets, and thus be capable of repairing, refueling, and upgrading satellites as they see fit. With this new risk posture in place, it then becomes a question of how well do space architectures take advantage of the capabilities offered through OOS and whether the use of servicing actually leads to wider benefits and positive trends in the overall space industry. This analysis can then grant insight into the sustainability and feasibility of OOS in the long term, and help determine if OOS becomes a crucial fixture of future space architectures.

System Dynamics Parameters and Analysis Scenarios

To assess the long term effects of OOS incorporation into a reference customer space architecture, three main scenarios were devised and simulated. The first is a future scenario where OOS has no effect on overall spacecraft design trends; i.e., satellite mass continues to increase with each generation as architecture designers attempt to pack more and more performance-increasing components (more transceivers, larger fuel tanks, larger solar panels, etc.) on each new satellite. In this scenario, OOS is still modeled as easing the designed-for, required success rate of new satellites, as there is much less of a drive towards risk aversion from the occurrence of spacecraft failures, due to the ability to easily rescue or repair on-orbit failures (with this change being enacted in the Prospect Theory-based Risk Posture Structure described in the Methodology). However, this rescue and repair of failed spacecraft is considered to be the only function that OOS serves in this scenario, as the continued increase in satellite mass indicates that spacecraft operators are not conducting such OOS operations as periodic refueling or technology upgrades; i.e., OOS is used only sparingly, largely for emergency operations, and is not heavily incorporated into customer architectures.

The second scenario is one in which satellite mass is held constant after OOS is implemented. This may be the most realistic scenario, as spacecraft designers are likely to determine that any mass savings they achieve through the use of OOS would simply allow for the inclusion of more components. This scenario is supported by the case of all-electric spacecraft, which were originally pitched as a means of freeing up significant space for extra communications capability due to their lack of chemical propellant. In addition, while each all-electric spacecraft does weigh less than a chemically-propelled counterpart, the mass savings are somewhat negated by the time of launch, as the spacecraft operator chooses to simply launch more of them to increase overall constellation performance (Clark, 2015). Thus, in the mass-constant scenario, it is assumed that in addition to the rescue and repair of spacecraft failures, OOS is also used for limited refueling and upgrading operations (in other words, in a more operational and relied-upon manner than in the first scenario). Holding mass constant in this scenario was forced exogenously, in order to assess the resulting trends in other industry variables (such as SV design life, total costs, and industry experience) within a narrower future scenario.

In the third and last scenario, the implementation of OOS leads to steady mass decreases in satellite mass, as spacecraft operators rely more and more on the refueling and upgrade capabilities offered through OOS. In this scenario, not only would the spacecraft be carrying less fuel, but it is also assumed that several of the on-board components would be lighter and less enduring, as the satellite receives continuous, regular technology upgrades and improvements over its design life. Additionally, the spacecraft operators become even better over time at capitalizing on the capabilities of OOS, perhaps through such techniques as the modularized "satlets" concept currently being explored by DARPA, where a common satellite bus would be designed for periodic upgrades and the addition of new, standardized performance modules (Courtland, 2014). This is considered to be the best—though perhaps not the most realistic—scenario for OOS implementation, as the reference space architecture is considered to be making the greatest use of OOS and incorporating its capabilities heavily in long term operations. Indeed, in a mass-reduction scenario such as this, it is assumed that the constellation is critically reliant on OOS as a means of maintaining overall performance throughout satellite design lives. Like the previous scenario, these mass reductions were forced exogenously in order to assess long-term trends in other industry and architecture variables.

Finally, in order to model the potential costs of OOS, four basic cost structures were applied in each scenario. The "No Cost" case represents an infrastructure where use of servicing does not cost anything to the customer. This case is analogous to the GPS constellation, which provides all users with continuous navigation and timing services at zero cost (provided the customer already has the necessary technology to make use of the GPS signal). The "Low Cost" case represents a servicing fee of 10% of the normal cost of a spacecraft. In other words, the extra cost associated with using the infrastructure is priced at 10% of what it would cost to simply replace a spacecraft, applied across all spacecraft (as it is assumed that all will be designed to utilize the OOS infrastructure during their design lifetimes). The Low Cost case is similar in concept to toll roads, in that a small fee is incurred to use the infrastructure but it is low enough to be marginal compared to total costs. The "Moderate Cost" case represents a servicing fee of 50% of spacecraft cost, and it is meant to simulate a more expensive infrastructure funding structure, similar in nature to airport landing, refueling, and parking fees. Finally, the "High Cost" case represents a servicing fee of 100% of the cost of a spacecraft, which is meant to represent the extreme end of feasible servicing costs; any higher and it would no longer make economic sense to service a spacecraft rather than simply replace it. In all cost structures this fee is applied across all new spacecraft, as it is assumed that all satellites will be designed with the expectation of at least some servicing within their design lives. It should also be noted that, with the servicing fee being applied to the spacecraft in the system dynamics model, spacecraft failures can then be diverted directly back into the production pipeline; i.e., the cost of rescuing a failed spacecraft is covered in the architecture design loop, and the model assumes that all failures will be automatically rescued, such that the constellation continues to maintain its minimum capability requirements. Of course, in reality OOS may not be so assured and also may only be paid for as it is required, but this analysis assumes that servicing is completely integrated into the reference space architecture, and thus a fee across all spacecraft is reasonable.

In each scenario, the model was first simulated under normal conditions until satellite mass and design life reached that of present-day levels (at Month 360, or 30 years into the simulation, representing the progression of the space industry from roughly 1985 to 2015). This was done to ensure that the model is "primed" at the start to match current paradigms, and then must *adjust* from these present-day *anchors* as OOS effects take hold (Tversky & Kahneman, 1974). The

validity of this initial simulation was shown in the section titled "Using System Dynamics To Assess On-Orbit Servicing Effects" above, with the simulated satellite design life and mass tracking very closely with the average from the real world data. Then, from this point on the model evaluates the progression of large-scale industry and architecture trends, with all cases being compared to a case in which no OOS effects are enacted. It must be noted that **only** SV launch mass and design life were tuned to resemble present day values. Another key variable, that of Total Program Cost, is meant to track the progression of the sum of all the major costs for the launching of the reference constellation, including manufacturing, launch costs, any added servicing fees, or for the replacement of failed spacecraft (with spacecraft cost estimates being derived from cost models provided in *Space Mission Engineering: The New SMAD* (Wertz, Everett, & Puschell, 2011)). However, the actual monetary values generated for this variable are not meant to be reflective of real-world figures. Instead, what this model can show are the general, overall trends in costs which can be expected to result from OOS implementation.

For all three scenarios, the following key variables are shown over a future thirty-year life span: SV Launch Mass, SV Design Life, Total Program Cost, and SV Construction Learning Curve. Again, what is of most value from the model results are not the specific figures themselves, but rather the overall trends caused by OOS. This is due to the original research question, which asked for an evaluation of the necessary large-scale conditions in space architectures and the industry in general to ensure long-term viability of an OOS infrastructure.

Modeling Results and Trend Analysis

The key results of the first scenario, in which the implementation of OOS has no effect on overall spacecraft mass trends, are shown in Figure 41 below.

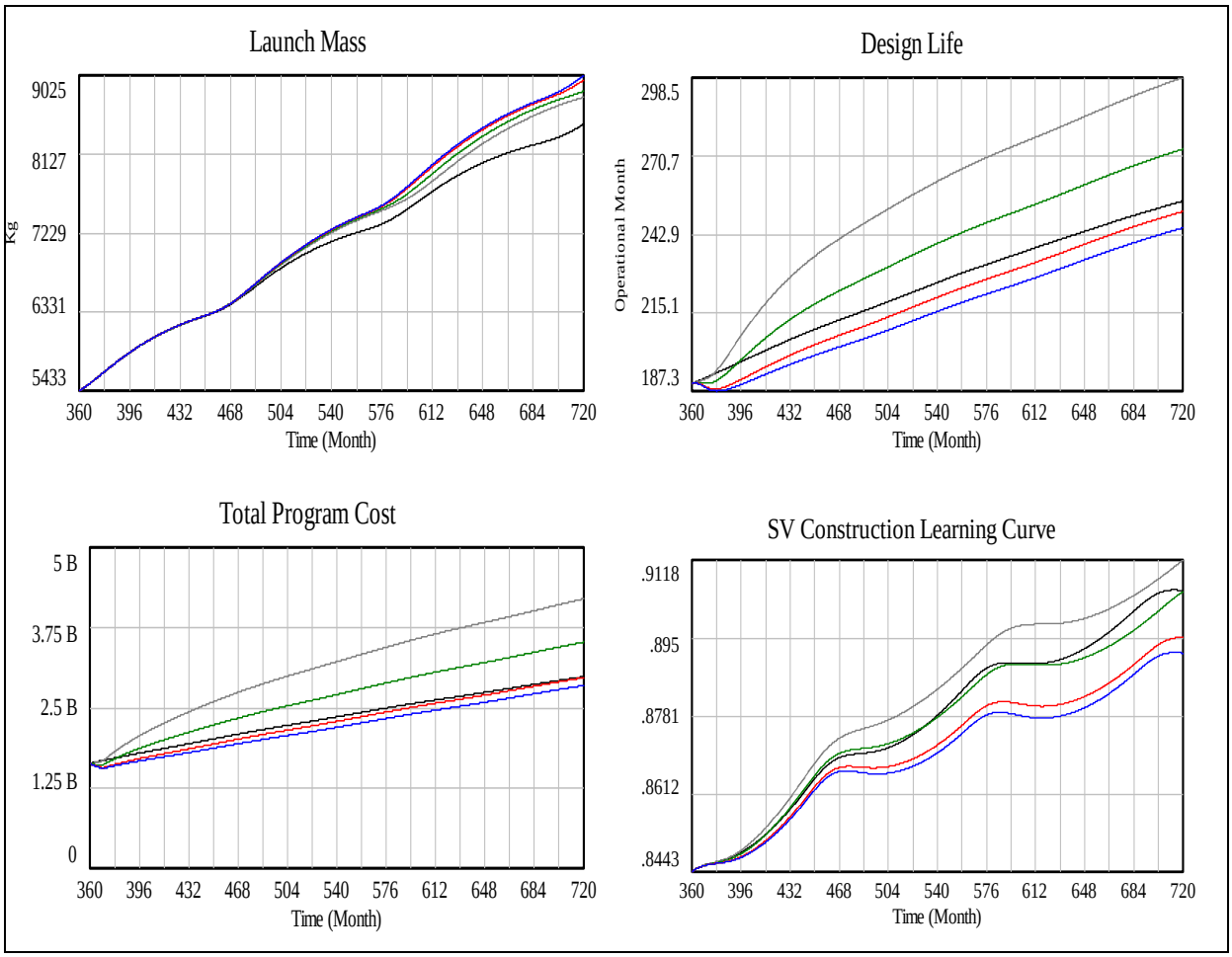


Figure 41. Results of the increasing spacecraft mass scenario. The blue line represents a No Cost OOS infrastructure, the red line represents a Low Cost infrastructure, the green line represents Moderate Cost, and the gray line represents High Cost. The black line displays the results of the Steady State results of the model without OOS implementation.

For the increasing mass scenario, each cost structure (No Cost, Low, Moderate, and High) is displayed along with a "Steady State" simulation; that is, the results of the model when OOS is not implemented at all and the model continues to run without any change in overall parameters. An initial interesting result of this simulation is found in the first graph, which displays "Launch Mass." When OOS is implemented, satellite mass actually begins to exceed that of the steady state case. This is due to the fact that, when design trends are linked to increasing the performance of each successive generation, OOS allows designers to pack more and more performance on each spacecraft, as the money saved on risk management goes towards

increasingly powerful spacecraft. Of course, it is expected that this transfer would occur to some extent in all three scenarios, but it is most evident in the case of unlimited mass increases.

Unfortunately, increasing satellite mass is also linked to increasing design life and total program costs, which in turn lead to decreased industry experience and a less effective learning curve. This is also displayed in Figure 41, and it shows how OOS implementation alone is not likely to be enough to reverse many of the potentially unsustainable trends currently evident in the space industry. Indeed, the added costs of OOS, especially in the Moderate and High Cost cases, are shown to be likely to cause even greater increases in spacecraft design life and total program costs, above that of the steady state case. Only in the cases of the No Cost and Low Cost OOS infrastructures does the model predict lower design life and total program cost, as well as a more effective industry construction learning curve. However, these benefits in the low cost cases are only a slight improvement over the base case, and do nothing to actually reverse overall industry trends—rather, the trends are just slightly dampened. Thus, the model predicts that, in a scenario where satellite designers remain stuck in old paradigms and continue to pack on more and more with each successive satellite generation, some small benefits to the overall industry can be achieved, provided that OOS is relatively inexpensive or free. In other words, in a scenario where OOS is only lightly incorporated into a reference architecture, overall industry trends are unlikely to be reversed, but there are still some benefits to be gained from an inexpensive OOS infrastructure.

Next, the results from the constant-mass scenario, in which spacecraft mass holds steady at its modeled value when OOS is implemented, are shown in Figure 42 below:

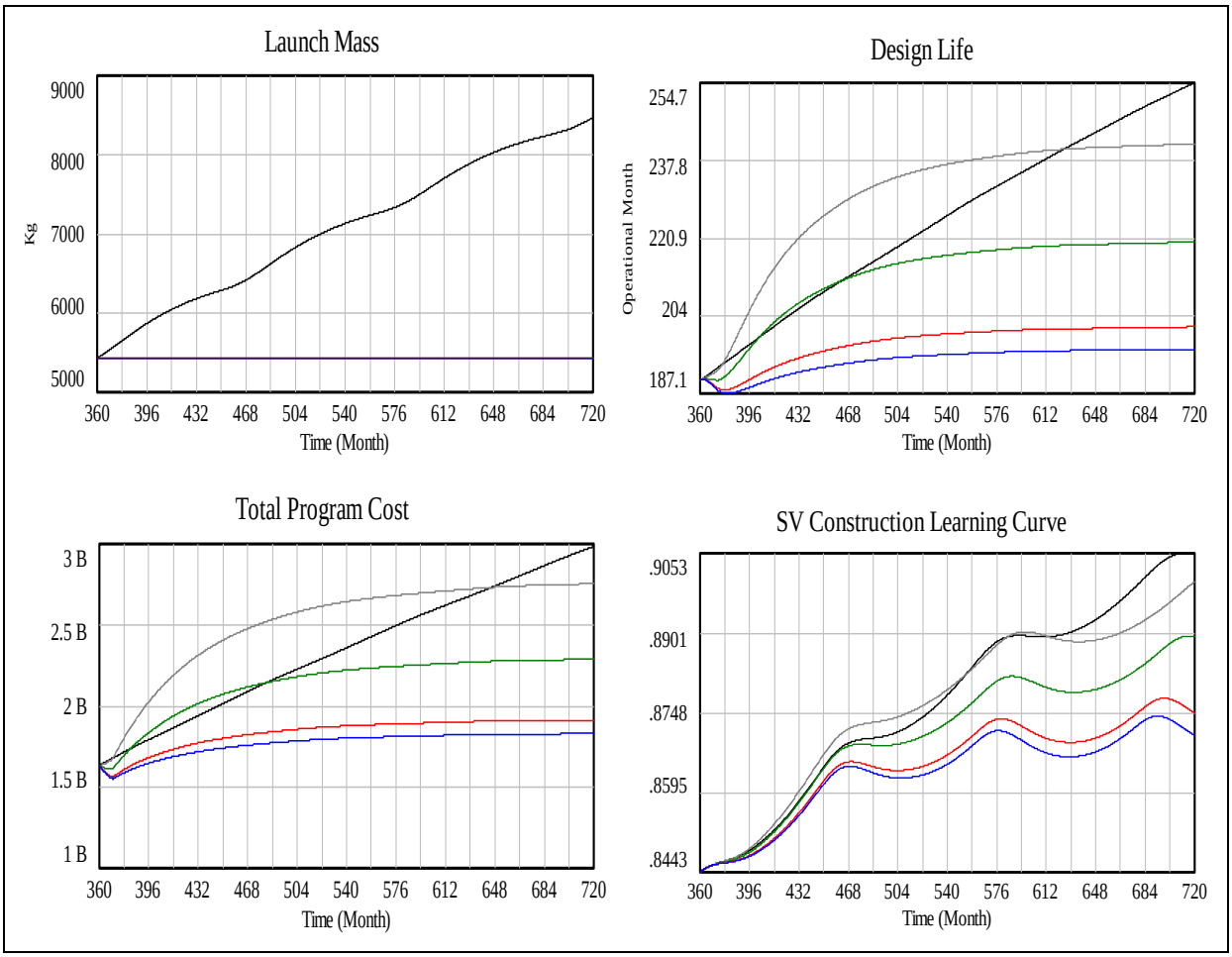


Figure 42. Results of the constant spacecraft mass scenario. The blue line represents a No Cost OOS infrastructure, the red line represents a Low Cost infrastructure, the green line represents Moderate Cost, and the gray line represents High Cost. The black line displays the results of the Steady State results of the model without OOS implementation.

To reiterate, in this scenario it is assumed that under an OOS paradigm, new spacecraft will be designed with the same launch mass that they had at the onset of OOS implementation. This is shown in the first graph displayed in Figure 42, which also displays the steadily increasing spacecraft mass of the steady state base case. As the other graphs clearly show, it is in this scenario that several benefits of OOS begin to take hold. In the No Cost and Low Cost cases especially, satellite design life and total program cost hold steady at levels well below that of the base case, and industry experience and learning curves are much more effective (recall that in this model, the learning curve acts as a multiplier to satellite cost and production time, such that a lower learning curve level translates to reduced costs and manufacturing delays). It is also

worth noting that even in the Moderate Cost case, total program costs remain largely on par with the base case during the initial decade of OOS implementation, until eventually steadying out at a much lower level. These positive effects can be attributed to the way in which OOS grants a degree of "breathing room" to the system; i.e., the reduced risk aversion and the removal of the need to continuously increase spacecraft mass allows costs and design life to level out, while also giving the industry time to accrue a greater amount of knowledge and capitalize on efficiencies gained when new generations of spacecraft need not last longer and longer to justify costs.

Of course, OOS implementation in this scenario is not without some drawbacks. As the High Cost case shows, if OOS is expensive to implement then many of these benefits are erased, especially at the onset of OOS implementation. Represented by the green line in Figure 42, when OOS fees are high then satellite design life and total program cost initially increase well above that of the base case, even in the much less risk averse setting provided by OOS, due to both the extra servicing costs and because satellite designs remain mostly unchanged. In the long-term, the reduced risk aversion and satellite mass do eventually lead to lower costs and less demand for increased design life, even in the High Cost case, but it is unlikely that satellite operators would be willing to tolerate such a large increase in price for many years just to get these small relatively small future benefits. Indeed, even the Moderate Cost case appears to offer very little in the way of cost reductions in the short term; with the total program cost in the Moderate Cost case almost identical to the base case for the first several years, OOS would likely need to demonstrate significant value on its own to justify the shift in architectural paradigms. Thus, an additional result from this scenario is that, for OOS to have the greatest chance of long-term sustainability and industry participation, it will need to be priced relatively cheaply in comparison to simply maintaining old architectural paradigms.

Finally, it is worth noting that in this constant-mass scenario, upward trends in satellite design life and total program cost are largely halted, but they are not reversed, even when the cost of using the infrastructure is very low or free. While this is certainly an improvement over the base case, it also means that satellite costs, design life, and production time remain historically quite high, while industry experience never recovers to peak levels. Of course, other factors besides OOS would likely impact these trends as well, but the results of this analysis suggest that OOS alone will not be enough to significantly reverse and shift overall industry

trends and paradigms (in a case where satellite launch mass remains largely unchanged over time).

In the third scenario, steady reductions in spacecraft mass result from OOS implementation, as servicing capabilities are incorporated more and more into the reference architecture. The results from this scenario are shown in Figure 43 below.

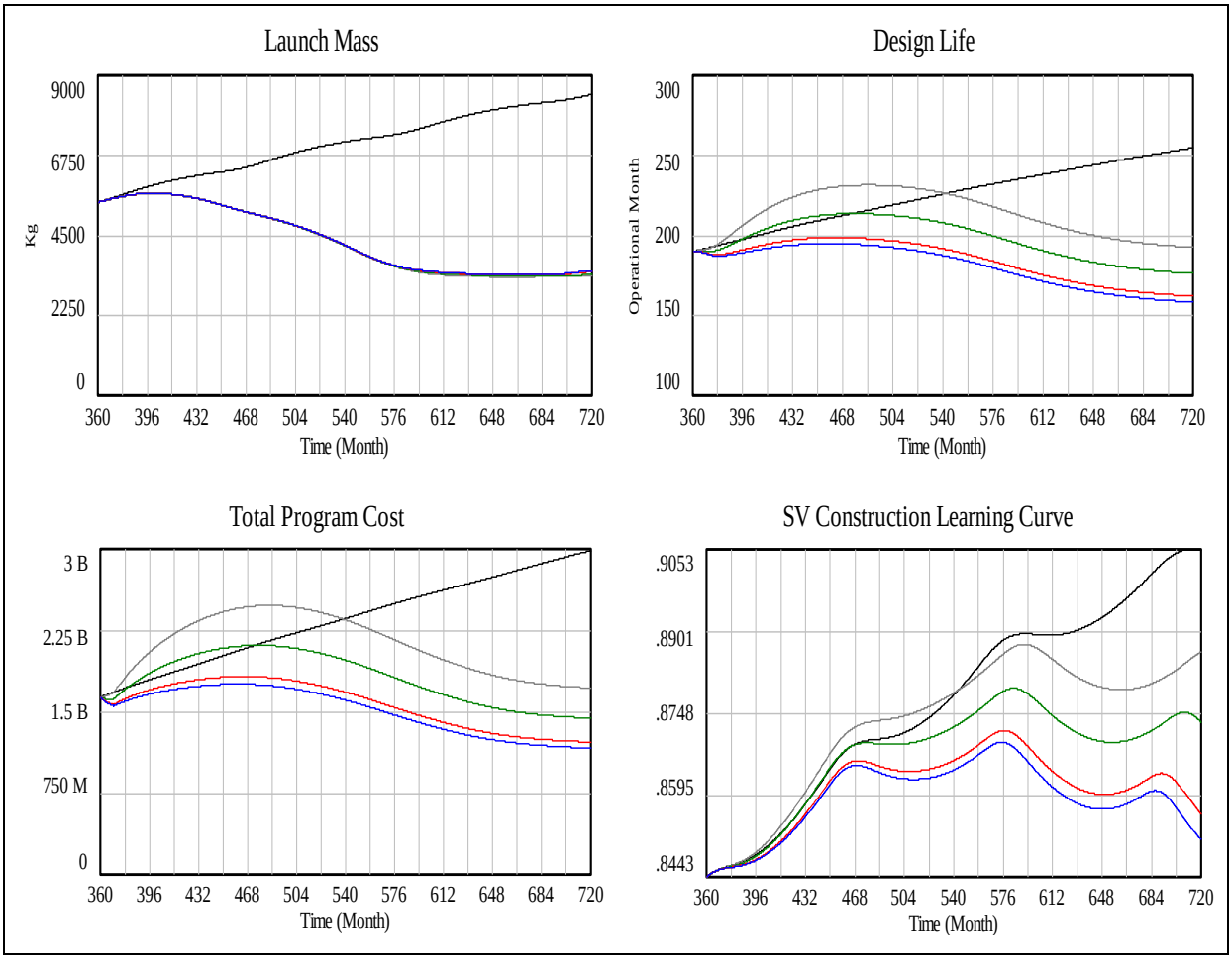


Figure 43. Results of the decreasing spacecraft mass scenario. The blue line represents a No Cost OOS infrastructure, the red line represents a Low Cost infrastructure, the green line represents Moderate Cost, and the gray line represents High Cost. The black line displays the results of the Steady State results of the model without OOS implementation.

As the results show, there are clear benefits to an architecture that heavily embraces OOS and capitalizes on its capabilities to drive significant changes in spacecraft design. In addition to

granting the space industry the necessary breathing room required to regain experience and halt trends in costs and learning curve effectiveness, an architecture which uses OOS as a means of steadily reducing spacecraft launch mass can expect to see sustainable long-term benefits, even if the costs associated with OOS are relatively high. This is evidenced by the fact that at the end of the simulation, spacecraft costs and design lives are substantially lower, and industry learning curves are significantly improved, under all three cost structures. Of course, benefits are greatest in the No Cost case, where cost and design life reductions are immediate and lead to substantial improvements to the overall system by the end of the simulation. Indeed, this No Cost, decreasing mass case—while perhaps unrealistic—nevertheless represents the best outcome achievable in the implementation of OOS.

As in the constant-mass scenario though, there are some concerns for OOS implementation raised in these results. The system is still quite sensitive to high costs, and while all the cost structures offer significant benefits in the latter half of the simulation due to the spacecraft mass reductions, it may again be unlikely that spacecraft operators would be willing to put up such large initial investments in OOS incorporation in order to realize the larger long term benefits much later. In fact, even the No Cost case experiences a brief period of slightly rising costs and design life increases near the beginning of OOS implementation, as the system adjusts to the new risk level and the industry continues to recover some of its experience and learning curve effectiveness. However, it is worth noting that this model still predicts costs to remain significantly lower than in the base case when OOS costs are low or free. Thus, in addition to relying more and more on OOS capabilities and steadily altering spacecraft designs to accommodate even more servicing over time, this analysis also predicts that spacecraft operators should be prepared for some growing pains associated with the long term integration of OOS into their architectures. It is also clear that OOS should be offered at as low a cost as possible, to further incentivize its acceptance into customer architectures.

If nothing else, this scenario clearly demonstrates that OOS offers the greatest benefits to customers when they fully embrace it as a vital component of their architectures and leverage OOS capabilities to make significant alterations in satellite design. By making such large-scale changes to new spacecraft and constellations, shown through a large reduction in satellite launch mass in this scenario, customers can expect to see positive downward trends in satellite design

life and total program cost, accompanied by the recovery of industry experience and learning curve effectiveness. However, this all assumes that OOS technology is effective and relatively guaranteed, and that the infrastructure is established and ready for customer use.

Discussion

The three scenarios analyzed above reveal a few consistent trends with regards to OOS implementation into customer architectures. First, the most intuitive conclusion holds firm: that for the greatest chance of sustainability, OOS will need to be offered at the lowest cost possible. While this is an obvious conclusion, this is not to say that OOS would not be found valuable at higher prices, or that this analysis captures all of the value provided to customer architectures through servicing. Long et. al., for example, demonstrate one way that OOS provides value in the form of increased system flexibility and other system ilities (Long, Richards, & Hastings, 2007). Indeed, even in the first scenario, where satellite mass, design life, and total program costs continue to increase over time, the No Cost and Low Cost cases were at least relatively similar in cost to the base case (if a bit cheaper), and this does not reflect the significant value to be found in cases where OOS would be rescuing failed spacecraft or allowing useful spacecraft to be refueled and extended in operational lifespan. Thus, a low cost OOS infrastructure, even one which is used sparingly and only in relative emergencies (as the first scenario envisions), still provides benefits to the overall industry; it just may not be addressing more of the core issues facing the industry.

This leads to the second conclusion, which is that, in order to spur wider benefits throughout the space industry, the capabilities of OOS will likely need to be incorporated heavily into customer architectures. In other words, for OOS to be successful in the long term and drive the industry towards more sustainable paradigms, it will need to be embraced in such a way as to cause fundamental shifts in design and constellation risk management. For example, to extract maximum value from the OOS infrastructure, spacecraft would likely have to be launched without a full lifetime's worth of on-board fuel, or with components are expected to be replaced periodically with upgraded technologies. This represents a significant departure from current practices of launching spacecraft without any plans for future changes or maintenance (beyond

software updates), and spacecraft operators may be reluctant to make such drastic changes to their architectures. However, this analysis does support the conclusion that, should space architectures make the necessary alterations to fully incorporate OOS, they can expect to see long term value and lasting benefits to the overall industry. It is also reassuring to note that NASA's On-Orbit Satellite Servicing Study similarly concluded that, as customers rely more and more on OOS capabilities, then they should expect to see cost reductions at the mission, program, and agency levels (NASA Goddard Space Flight Center, 2010).

Furthermore, the progression of the three scenarios shows that as OOS is incorporated more and more into customer architectures, then they are able to tolerate slightly higher prices than if OOS is only used sparingly. Of course, in all three scenarios, the High Cost cases were viewed as unlikely to drive long term sustainability; indeed, if OOS is very expensive, then spacecraft operators are likely to simply replace failed or under-performing satellites rather than pay to have them serviced. However, in the second and third scenarios, it was shown that customer architectures are likely to be willing to pay at least at the Moderate Cost level, compared to the cost of the base case which does not bring with it any of the other benefits associated with OOS. Recalling that the Moderate Cost cases applied a servicing fee of 50% of the base satellite cost, this does represent a relatively high cost tolerance in exchange for the capabilities offered through OOS. Once again, this does assume that servicing is both assured and crucially relied upon by customers, and this will almost certainly require effective policies in place to govern and support an OOS infrastructure (which is discussed in-depth in the following sections).

Along these lines, the results suggest that the determination of specific "break-points" beyond which OOS is no longer feasible will vary significantly with the level of OOS implementation envisioned. While greater incorporation of OOS capabilities is linked to an overall higher cost threshold, this analysis does not attempt to capture the value of OOS to more specific customers or space architecture types; e.g., a large constellation of standardized, long-lived communications spacecraft at geosynchronous orbit may be able to extract more value from OOS than a smaller, simpler, short-lived imagery spacecraft in low-Earth orbit. Thus it will largely be up to individual customers to determine their own willingness to pay for servicing; indeed, Saleh has already done much to quantify the value of OOS to customers, when considering the increase in options and system flexibility that servicing capabilities bring to

space architectures (Saleh, Lamassoure, Hastings, & Newman, 2003). However, if OOS delivers on its promises of significantly reducing spacecraft risk of failure, fuel depletion, and/or technological obsolescence, then this analysis does support the conclusion that customers would be willing to pay a significant amount as they incorporate OOS capabilities more and more into their architectures.

Finally, the small difference in benefits between the No Cost and Low Cost cases further supports the assertion that servicing fees are likely not as influential on long-term sustainability as is the level of overall incorporation of OOS capabilities. If the No Cost case was shown to be significantly more powerful in driving down program costs and satellite design lifetimes, then it could be said that customers are highly sensitive to any additional costs and would be unwilling to utilize an infrastructure that comes with any extra fees. However, from these results, it is suggested that customers would only become averse to OOS if it priced quite high (approaching the cost of launching entirely new spacecraft). This is especially true in the cases where the customers are actively making design changes to better incorporate OOS capabilities. Thus, while these results show that a low cost infrastructure is inherently more beneficial than a high cost one, this does not mean that OOS has to be developed and operated under regimes that offer it completely freely; rather, small fees can be applied without fear of significantly hampering long-term sustainability.

Perhaps the largest drawback of this analysis, and the greatest hindrance to its generalizability to other cases, is the fact that it is heavily focused on large, relatively standardized commercial communications satellites typically found at geosynchronous orbit. While these types of satellites are common and represent a large portion of all commercial space operations, they are far from being the only kinds of spacecraft to which an OOS infrastructure would render services. At the same time, however, the conclusions generated from these results are considered to be widely applicable to other infrastructure projects and space systems. For instance, to incentivize participation in any infrastructure, it is almost always necessary to offer its services at relatively low cost. Furthermore, simply lessening the level of risk aversion in the space industry, whether through OOS or another means, is beneficial but also unlikely to significantly alter dominant design trends.

Finally, as with most any model, the results become less and less reliable the farther one projects into the future. This model seeks to temper this somewhat by focusing on "big picture" variables, such as average satellite design life and launch mass and the accumulation of industry experience, and the results generated are meant to inform large-scale policy regimes, rather than more specific system design points. However, it is impossible to know for sure if a model applicable to today's space industry will remain relevant to the space industry 30 years from now. Indeed, any number of innovative technologies or shifts in stakeholder values could result in a much different commercial space environment than the one that was simulated by this model. However, within the context of analyzing the effects of general shifts in spacecraft costs and risks which may result from an infrastructure project like OOS, these results do provide a few meaningful conclusions which policymakers must consider in the development and launch of such a system.

Policy Recommendations

In addition to exploring the design options and long-term development possibilities for orbital infrastructure concepts, recommendations must also be made as to the most effective and logical policies enacted to guide such projects; policies which will lead to the largest customer benefits, the greatest chance of overall sustainability and viability of the infrastructures, and also the protection of stakeholder rights through proper management and clear regulations. In the following sections, policy considerations for infrastructure development will be discussed, drawing heavily from lessons learned in past projects and the results of the above analysis.

LEO Communications and Data Relay Infrastructure Policy Recommendations

Providing communications and data relay between spacecraft on orbit, and between spacecraft and ground stations, may not be the flashiest, most exciting space mission, but it is absolutely critical to many real-time operations and the timely delivery of needed data. As such, this mission has often been fulfilled by civil or defense agencies, but there is also likely a strong case for allowing private firms to take greater responsibility for this service. In addition, there are several political and legal considerations which must also be considered in the development and fielding of a large, highly disaggregated orbital infrastructure, and this is discussed in depth below.

Funding and Development Policies

Perhaps the most obvious policy regime for the funding and procurement of a communications and data relay infrastructure is to do so through a civil agency like NASA. The public space sector has decades of experience in orbital connectivity, and the public would have high confidence in its ability to operate and maintain such an infrastructure. Since it would also be developed as a publicly-funded, publicly-available infrastructure, one which is designed more or less as a "backbone" for other space activities, it would also be likely to have lower usage fees than if the infrastructure was required to return a profit. The main drawback to this, however, is

that NASA's budget is already stretched quite thin, and it may not have the necessary funds to field an entirely new data relay infrastructure (even one of the optimized, less costly single-plane designs shown in the second case study above). Agencies like NASA are also quite well served by TDRSS as is, and absent a congressional directive NASA may be unlikely to develop its own LEO infrastructure for use by new and untested space enterprises.

Launching a LEO data relay infrastructure as a military system is also an option, and this could borrow heavily from the policy regime used to develop and launch the GPS constellation. In this case, the infrastructure would be designed first and foremost to service defense assets in LEO, and then any additional capacity could be provided either free of charge or at low cost for civilian applications. This policy would benefit heavily from the deep experience of the U.S. military in space operations, in addition to likely having a greater chance of receiving all the necessary funding, due to the fact that such a venture would be viewed as a national security mission with large benefits to global military operations that are heavily reliant on real-time satellite data. The highly fractionated, small satellite-based designs highlighted in the results as optimal solutions also fit in well with the U.S. Air Force's current push towards more disaggregated, resilient space systems (U.S. Air Force Space Command, 2013). In addition, a LEO communications and data relay infrastructure could be touted and sponsored in much the same way as the interstate highway system: a publicly funded (data) transportation network, highly necessary for efficient military operations, but which nevertheless has extensive civilian applications (indeed, daily civilian use could eventually exceed military use). The main drawback of this regime is that, while it could be expected to provide the same high quality service displayed by GPS, it would remain a military asset; one which could deny service to customers at any time due to crisis or an increased need for military data. The Department of Defense may also be hesitant to allow outside actors access to a network specifically designed for the transfer of classified data to and from assets (although this concern could be alleviated somewhat through encryption or the provision of a separate signal for civilian use, as in the early days of GPS). Furthermore, the DoD may already feel that they are served well enough by current communications and data relay assets, and as such do not wish to field a system which may subsequently see more use by civilian assets than defense ones.

If neither of these publicly owned and operated options are chosen, there is also a great deal of promise in promoting a LEO communications and data relay infrastructure as a more private venture. Indeed, a great deal of interest has recently been shown in proposals for space-based internet services; ones which could rely heavily on large numbers of small satellites to connect users on the ground with the world wide web (Lapowsky, 2015). It would seem a natural extension of these internet services to connect LEO customers into such a global network (although necessarily with some necessary design revisions). It is also worth noting that the next generation of Iridium satellites (which make up the global constellation already providing ground-based customers with worldwide satellite telecom services) is also slated to provide limited LEO-to-LEO connectivity services (Iridium NEXT, 2015). Thus, it seems the private sector is ready and willing to take on ventures very similar to the kind of data relay infrastructure envisioned in this thesis.

One way that the public sector could aid in this enterprise, however, would be through the provision of funding and oversight, similar to how commercial crew and cargo services are now being procured through the Space Act Agreements. Under this policy regime, NASA provides funding to private firms as they meet developmental milestones and demonstrate adequate performance of their proposed space system designs (Anderson C. , 2013). Applied to a LEO data relay infrastructure, NASA or the DoD could solicit infrastructure designs from private actors, who could begin developing their individual systems knowing that the public agencies will act as initial customers. Of course, the ultimate objective of this policy would be to select one final firm to provide the entire infrastructure, but providing seed money in the initial stages to several competitors should eventually lead to better designs and lower costs than if only one firm is chosen outright. Another way this policy regime could work would be for the firms to compete to build the best "ring" of infrastructure spacecraft (as examined in the second Case Study above). Then the firm with the best design from this phase could subsequently use additional seed money or portions of their new profits to begin building the rest of the global infrastructure (by adding more orbital planes, as shown in the first Case Study). In this way, costs to build the completely global infrastructure would be much more spread out over time, and by building individual infrastructure rings, designers and policymakers could gain a great deal of early experience in LEO data relay, at a lower cost, before committing to a larger, more expensive global design.

Regardless of the policy used to distribute funding or select an ultimate infrastructure provider, it is almost certain that a public agency would be involved at least in a small way in its oversight or initial phases of development. Additionally, civil or defense architectures would be viewed as primary customers for any infrastructure project. Thus, policies should be enacted early in the development of LEO infrastructure, in order to encourage greater competition and innovative solutions, as well as to ensure the infrastructure has the necessary backing and support needed for long-term sustainability and continuous improvements.

Political and Legal Considerations

In addition to the various options for funding and developing a LEO communications and data relay infrastructure, there are also several policy considerations which must necessarily go into such a large scale infrastructure project. These regulations concern issues such as which electromagnetic frequencies the infrastructure uses to send and receive data, who holds ultimate liability and ownership over the system, how customer data is securely transmitted, etc.—similar to many of the regulatory concerns facing terrestrial telecom services.

Spectrum Management and Data Security

The most pressing regulatory concern for any communications and data relay infrastructure is electromagnetic spectrum management. Interference by the Earth's atmosphere causes there to be only a relatively small communications "windows" in the electromagnetic spectrum (centered around visible light as well as high frequencies between approximately 30 MHz to 30 GHz), and as such there are limited frequencies available for use by spacecraft operators (Wertz, Everett, & Puschell, 2011). These frequency "slots" are heavily regulated—in the U.S. by the Federal Communications Commission, and internationally by the International Telecommunications Union (ITU). The current regime for assigning these slots is a "first-come, first-served" basis, and each year international conferences and multi-lateral planning meetings are held to review and revise regulations and frequency assignments. Without such heavy regulation of the electromagnetic spectrum, it is likely that radio interference between spacecraft would be much

greater, degrading the operational environment for all actors, and there would also be little means of recourse for entities to settle disputes over spectrum use. Indeed, under the current regime, if there is any chance of interference from a proposed satellite or constellation, then the operators must negotiate with other nations or entities before launch to ensure that neither interferes with the operations of the other (Rendleman & Smith, 2011).

The LEO infrastructure analyzed in this thesis derives basically all of its value from its ability to reliably receive and relay data to and from customer spacecraft, as well as downlink data from the infrastructure spacecraft to ground stations on a round-the-clock basis. As such, any interference encountered due to spectrum mismanagement would immediately decrease the value and performance of the infrastructure. Therefore, in addition to selecting a set of frequencies which would be most conducive to a high-performance infrastructure, the developers of the infrastructure would also need to negotiate with other spacecraft operators to obtain the necessary rights to a frequency (a process which can take as long as three to five years (Rendleman & Smith, 2011)). Fortunately, given the fact that the infrastructure would enjoy constant inter-links between its own spacecraft as well as near-constant communications with ground stations, it infrastructure owners may find it easy to obtain a relatively low frequency slot; one which may not be as useful to other spacecraft operators due to its lower data-carrying performance but which could still yield acceptable performance for the infrastructure under a paradigm of constant data downlinking. In addition, the interlinks between infrastructure could actually use frequencies which are not useful for space-to-ground communications, as they would not be required to transmit through the atmosphere. Thus, obtaining the necessary ownership of frequency slots for an infrastructure project may not prove to be as challenging as some may think, due to its relatively unique operating parameters.

Also of concern, with regards to the data being transmitted to and from the infrastructure, is the security of those transmissions. Just as governments and private firms expect their terrestrial communications to be secure, so too must a space-based communications infrastructure have the necessary safeguards against intrusion and data breaches. To ensure this security, it will almost certainly be necessary that a minimum level of data encryption be required for all customer data, especially if the infrastructure expects to handle the sensitive data of government or military customers. Enforcing a high level of encryption within the interlinks between infrastructure

spacecraft (in addition to demonstrating the security of these interlinks on a practical basis) would also do much to boost customer confidence in the infrastructure's ability to securely transmit data. Furthermore, the ground stations being used to constantly downlink data from the infrastructure must also be secure, both in their cyber connections to outside networks as well as from a real physical standpoint. Indeed, there may be no faster way for an infrastructure to scare off customers and destroy its own long-term success than by demonstrating insecure data links.

Orbital Debris

Another major consideration, especially for the very large constellations explored in the above case studies, is the chance of increasing the already significant amount of orbital debris in LEO. Debris presents a major concern to spacecraft operators, as it can stay in orbit for decades and remain in dangerously close proximity to active space assets. As the average relative velocity in LEO between orbiting objects is 9 to 10 kilometers per second (and even faster relative velocities in more direct passes), even very small pieces of debris present a massive risk of damage to other orbiting assets (Wertz, Everett, & Puschell, 2011). The dangers of orbital debris were displayed fully in the 2009 collision between an active Iridium communications satellite and a deactivated Russian spacecraft—an event which, in addition to abruptly ending the operational life of an active asset, also created a vast cloud of over 3,300 catalogued pieces of debris which remain in Iridium's operational altitude regime to this day (Satellite Collision Leaves Significant Debris Clouds, 2009). Furthermore, as stated in the 1967 Outer Space Treaty, as well as the 1972 Liability Convention, spacecraft operators are held ultimately responsible for any damages caused by their spacecraft, for as long as those objects remain in orbit. While this provision has yet to be tested in an operational setting, all spacecraft operators agree that the utmost care must be taken to ensure that their space objects do not cause damage at any point between launch and deorbiting (Listner, 2011).

Knowing these concerns, any infrastructure project must consider the potential for debris creation in the launching and operation of the constellation, and take the proper steps to mitigate these risks. The US and UN have both published guidelines for debris mitigation, which all spacecraft operators must adhere to and which cover all aspects of the space mission, from

design to launch to operations and finally to end-of-life disposal (Wertz, Everett, & Puschell, 2011). For the LEO data relay infrastructure especially, the designers must coordinate with other spacecraft operators to ensure that the introduction of potentially several hundred new infrastructure craft in various orbital planes will not present an undue risk of collision with other satellites or constellations. Furthermore, the infrastructure must be designed for proper end-of-life disposal, most likely through intentional deorbiting of assets. This follows from the results of the optimized infrastructure case study, which showed that the best designs are positioned at a relatively high LEO altitude (spacecraft at this altitude would most likely require decades for their orbits to degrade naturally).

Optical Communications Considerations

If designers were to choose to use optical communications within the infrastructure (as this thesis suggests is an option worth exploring), then there is an extra consideration that arises from the use of this technology. Laser communications technology necessarily presents an additional hazard, in that the laser itself could cause harm to individuals on the ground or to other spacecraft if the beam is misdirected. As such, there are significant regulations in place for the use of optical communications. The US government and Department of Defense's Joint Space Operations Center currently maintain a Laser Clearinghouse, and anyone planning to use an optical communications system must obtain predictive avoidance open firing windows before conducting any laser activities (Rendleman & Smith, 2011). While certainly not an insurmountable obstacle, such regulations do require additional planning from spacecraft operators if choosing optical communications over other frequencies in the electromagnetic spectrum.

International Participation and Coordination

Finally, as space is an inherently international arena, consideration must be given to the rules that will govern international participation in the infrastructure. The desire of several national entities to field their own global satellite navigation systems is evidence of the fact that nations do not always trust other states to provide space-based services. Such distrust is likely to only be

heightened when considering the transmission of classified or sensitive data through another nation's infrastructure, regardless of the level of encryption. Therefore, policymakers would do well to decide at the outset what level of international participation is desired in the LEO data relay infrastructure, as this could have extensive ramifications on overall design and long-term development of the system.

Encouraging international use of the infrastructure seems rational at first glance. It would greatly increase the number of customers (thereby increasing revenues), and perhaps even contribute to an overall greater degree of cooperation and collaboration amongst nations' space operations. To achieve this, however, the infrastructure would almost definitely have to be offered as a largely public service (provided by a trusted commercial entity), with extensive collaboration between several international space agencies. This may then introduce increased inefficiencies to the system, as the design and planning for the system would have to be approved by several entities with possibly competing interests and intentions for use. Furthermore, this may act to dissuade a potentially primary customer of the infrastructure—U.S. defense assets—from using the infrastructure, due to the security concerns of transmitting data over a relatively unsecure network. Thus, the benefits to pushing international use and development of the infrastructure must be weighed against several mitigating factors.

Instead, the infrastructure could be developed with much more significant funding and oversight from a U.S. national entity such as NASA or the DoD, with the recognition that launching the infrastructure as a national asset could deter some space-faring nations from participating. This may be an acceptable regime, as the analysis showed that it would largely be U.S.-based start-ups and other small-scale customers that would benefit the most from the infrastructure. Furthermore, this may actually increase the likelihood that the infrastructure would be given the green light for development, as it could be designed for use with established U.S. assets first, with extra capacity provided for the types of customers identified in the analysis. This is not to say that the infrastructure would not allow for any international participation; much like the GPS constellation, it could be open to use by anyone with the necessary equipment. This also does not rule out significant commercial contributions, as it was determined above that some form of U.S.-based public-private partnership would most likely be utilized for the construction and operation of the project. However, policymakers should

recognize that any increase in national oversight of the infrastructure is likely to decrease the likelihood of participation by some international partners.

On-Orbit Servicing Infrastructure Policy Recommendations

An OOS infrastructure differs significantly from a communications and data relay infrastructure, in that this project necessarily involves a servicing spacecraft making physical connections to customer satellites very far away from controllers on the ground. It is also expected that a fleet of OOS spacecraft, involved in regular servicing operations and interacting with a variety of customer spacecraft, would require significantly higher investment and oversight than a LEO constellation which simply relays customer data. Thus, the policies required for the development and management of an OOS infrastructure differ significantly in key areas, and this is discussed in depth below.

Recall that the results from this analysis are aimed primarily at the very large, relatively standardized communications spacecraft typically found at geosynchronous orbit. These are considered to be important first customers of an OOS infrastructure, and are likely to have the most obvious benefits from servicing capabilities like refueling and upgrading of technologies. Thus, the policies described below are also directed primarily at the implementation of an OOS infrastructure at geosynchronous orbit, but it is expected that effective policies for this specific type of OOS infrastructure would also be largely generalizable to those designed for operations in other orbital regimes.

Funding and Development Policies

As shown in the system dynamics modeling results, for the greatest chance of long term viability and success, an OOS infrastructure should be offered at the lowest cost possible (that is, at the lowest cost to the customer). However, the costs to fully develop, launch, and maintain an effective OOS infrastructure are expected to be quite high. For example, NASA's On-Orbit Satellite Servicing Study concluded that a relatively simple mission, involving the autonomous capture and repositioning of 10 customer satellites in geosynchronous orbit, is estimated to cost

approximately $910 million (which includes all development, manufacturing, and programmatic costs). Meanwhile, a refueling mission capable of providing services for up to 25 customer spacecraft is estimated to cost over $1 billion, while an even more complex assembly mission is likely to cost over $3 billion (NASA Goddard Space Flight Center, 2010). Thus, for such an infrastructure to be developed and funded adequately—and then to have low enough prices to incentivize customers to actually use it—significant upfront investment by public agencies would likely be required; investments which may or may not actually be fully recouped by the investor, at least in strictly monetary terms. The need for significant upfront investment is especially cogent when considering the fact that there are as yet no readily available customers in line to pay for servicing operations, making for what many describe as a "chicken-and-egg" problem of needing both an established infrastructure as well as customers ready to use the infrastructure (Richards, 2006).

Luckily, agencies like DARPA and NASA have already performed much of the initial testing and experimentation required to demonstrate the technical feasibility OOS (described in the Background and Literature Review above). Following from this, there are a variety of policy regimes that could now take OOS from the experimental stage to a more operational footing. Perhaps the most straightforward policy would be to develop OOS as a civil project, launched and operated by an agency such as NASA. In this scenario, the infrastructure would be developed primarily for use by the agency, to conduct servicing operations of its own assets, with the option to have other entities pay for use of the infrastructure (similar to how TDRSS is currently operated). There are several advantages to such a regime; for example, maintaining the infrastructure as a civil project can help to ensure adequate funding (especially once the infrastructure is established), as well as allow for lower costs than if the infrastructure had to generate a profit. Potential customers would also be likely to have more confidence in the performance and safety of an infrastructure provided by such an agency as NASA, due to its expertise and relative transparency. However, the main drawback to this regime is that it may be quite difficult to secure the substantial funding required to turn NASA's servicing experiments into an established infrastructure, especially when considering the very cash-strapped operating environment the agency currently faces. One could also raise the argument that a purely public infrastructure may be lacking in the efficiency and cost-effectiveness that could result from a more private venture.

A similar policy regime to fund and develop an OOS infrastructure would be to initially launch the system as a defense architecture, with its primary mission being the servicing of national security assets. This would be operated and maintained similar to the GPS constellation, with the Department of Defense retaining ownership of the infrastructure but offering its services to others as its scheduling allows. Advantages to this policy are much the same as those of the civil project described above, in that the infrastructure could draw on the substantial defense budget to pay for its development and allow for lower usage fees, as well as benefit from the decades of experience in military space operations. The main drawback to this policy, however, is that the infrastructure would inevitably be viewed as a military project; indeed, the potential for OOS capabilities to be weaponized is already viewed as a major concern (Belcher, Freese, Laygo, & Osborn, 2014). Thus, customers may be more averse to participation (especially international entities), and objections could be raised on the basis of the Outer Space Treaty's strict prohibitions against weaponization. The likelihood of such a scenario is already being evidenced by other nations' development of their own satellite navigation constellations, even though the official policies governing GPS promote transparency and international interoperability (Clore, 2012). In summary, while operating an OOS infrastructure as a national defense asset offers many advantages in funding, oversight, and expertise, it is also likely to have more substantial political and legal impediments.

Another policy framework for OOS implementation is that of the public-private partnership. Under this regime, a public agency such as NASA or DARPA plays a large role in the management and overall guidance of the infrastructure, while allowing private firms to propose, develop, and operate their own systems. This framework also allows for large public investment in the initial development of the infrastructure, in addition to assuring private firms that civil or defense architectures will be available as initial customers. The Space Act Agreements currently governing NASA's procurement of commercial crew and cargo operations from private firms are a current example this sort of public-private partnership, which seeks to capitalize on the efficiencies and innovations of private industry while also drawing on the experience and oversight of public space agencies (Anderson C. , 2013). The largest drawback of such an arrangement is that private space systems may be viewed as slightly more risky and less proven than public systems, but this perception of risk can be largely mitigated through transparency and the demonstration of operational performance before more public funding is awarded. Public

agencies must also be willing to relinquish a significant amount of operational control of the infrastructure, and there must also be clear regulations for ownership and liability governing all servicing activities.

The public-private partnership framework appears to offer the greatest likelihood of success and real-world implementation, and this is evidenced by the fact that DARPA has recently promulgated a Request for Information (RFI) along these very lines. Specifically, the RFI states that "DARPA is particularly interested in establishing a public-private partnership that would make cooperative robotic servicing available to both military and commercial GEO satellite owners on a fee-for-service basis. The partnership would help develop near-term technical capabilities and significantly contribute toward the creation of a sustainable, commercially owned-and-operated space robotics enterprise" (DARPA, 2014). From this statement, it is clear that an OOS infrastructure is intended to not only undergo initial development as a public-private partnership, but it should also one day be spun off as a completely commercial venture (albeit with civil and defense entities as primary customers, as many commercial space architectures currently operate).

Finally, the results from this thesis' analysis do support the conclusion that a public-private partnership can ensure long-term sustainability of not only an OOS infrastructure but also many types of customer space architectures in general. With NASA and DARPA continuing to demonstrate the technical feasibility of satellite servicing, customers should have the necessary degree of confidence to begin making the required design changes to their architectures to begin to more heavily incorporate OOS capabilities. This confidence would be further bolstered by the near-term use of OOS in operational military constellations, as the DARPA RFI envisions, to show how OOS can become an integral component of a high-performance space architecture. Then, with the assurance that defense architectures will act as primary customers, coupled with the likely infusion of substantial public investment, the private operators of the OOS infrastructure should be able to overcome the initial growing pains shown in the system dynamics model, while still providing services at a relatively low cost. This should then lead to sustainable, long-term benefits to a variety of customer space architectures, provided that the continued performance of the infrastructure encourages greater and greater incorporation of OOS into daily space operations.

It is also worth mentioning that, while the system dynamics analysis presented in the Results above was directed primarily at large, geosynchronous communications constellations, these policy recommendations are meant to be applicable to a wide range of servicing operations. Few would argue that an infrastructure should not be offered at as low a cost as possible, or that it will require a high degree of customer participation, regardless of in which orbital regime the OOS infrastructure is based. Similarly, the political and economic dichotomies between implementing the infrastructure as a civilian or defense system, or between a public or private venture, will also be present regardless of if the infrastructure is based in LEO or GEO. Thus, the fact that the trends identified in the analysis did not necessarily apply to smaller, less powerful spacecraft based in LEO should not be considered as a barrier to using the policy regimes described above outside of a GEO infrastructure.

Political and Legal Considerations

Regardless of the policies governing the funding and development of an OOS infrastructure, there are in addition several political and legal considerations that would accompany such a project. In an extensive overview of the issues facing OOS implementation, Belcher et. al. identify six key policy areas which could impede future servicing projects:

1) **Responsibility and Ownership**

   Because OOS is necessarily accompanied by the risk of accidental damage to customer spacecraft, there must be clear and unambiguous rules for ownership and liability. Customers will be unlikely to participate in an OOS infrastructure that does not provide for clear rights in the case of damage or operational mishaps. In addition, such clear rights and lines of ownership should be extended into an international framework, agreed upon by all nations participating in the infrastructure.

2) **Insurance**

   The relatively new and unproven enterprise of satellite servicing means that insurance providers are unlikely to provide coverage to customers utilizing OOS. The availability

of insurance to cover possible damages caused by OOS may be crucial to encouraging spacecraft operators to begin using the infrastructure. While much more analysis and research will likely be needed before a suitable framework is established for servicing insurance, Belcher et. al. recommend that NASA and DARPA continue technical demonstrations and share their data with space insurance firms to aid in this process.

3) On-Orbit Operational Regulation

Again, owing to the fact that OOS is a relatively new and untested enterprise, a proper regulatory framework must be established for the overall conduct and operational planning of OOS activities. The fact that space is a common operating environment, in which all operators must pass overhead others' territory, means that it shares many similarities with international airspace or open waters. As such, a regulatory regime which functions similarly to the UN's Convention on the Law of the Sea is likely required, and it should be written so as to encourage cooperation, transparency, and mutual respect between all sovereign and/or private assets. Belcher et. al. recommend that the FAA issues and controls regulations for US OOS activities, as it has existing experience in the regulation of commercial space activities. However, it is also likely that an international organization, be it the UN Office for Outer Space Affairs or another entity, must be provided with the power to issue and enforce regulations on all OOS activities.

4) Assurance

As previously stated, it is widely acknowledged that OOS has strong weaponization potential, as any spacecraft which can maneuver close to and manipulate the components of another satellite can also disable or otherwise cause damage to selected targets. Thus, in addition to strictly following regulations and best practices for OOS operations, it may also be necessary to publish telemetry and a detailed schedule for servicing, perhaps even with punitive measures in place for those instances where an OOS operator deviates from their published plans without the proper notification. This is also a reason for why the US may not wish to launch an OOS infrastructure as a purely military system (in addition

to why DARPA makes liberal use of the word "cooperative" when describing future servicing operations in their recent RFI (DARPA, 2014)).

5) Spectrum and Slotting

For effective, assured OOS operations, it is crucially important that operators have sustained communications with both the servicing spacecraft and the customer satellite, and this will require clear and unambiguous allotment of radiofrequencies. While there are currently very stringent international regulations on spectrum use, there is the possibility of new challenges in OOS, where the spacecraft will be in extremely close contact with one another far away from Earth. This could increase the chance of radio interference, especially if the servicing spacecraft and its customer are operating on similar frequencies, and thus effective regulations governing OOS must include provisions for spectrum allotment and communications protocol.

6) Imaging

For decades, there have been strong prohibitions against direct imaging of spacecraft, both for national security reasons and for the protection of proprietary satellite designs. However, OOS will likely necessitate such on-orbit imagery. Regulations are already in place for terrestrial imaging, and in the U.S. the Department of Commerce and the National Oceanic and Atmospheric Association (NOAA) are responsible for the issuing of imagery licenses and regulating the use of terrestrial imagery. Belcher et. al. argue that NOAA already has the statutory authorization needed to regulate on-orbit imaging of other spacecraft, and this authority will need to be extended to OOS operations, in order to ensure there are clear guidelines on the use and promulgation of orbital imagery (Belcher, Freese, Laygo, & Osborn, 2014).

In essence, all of these policy considerations underscore the need for a clear, unambiguous policy regime in place to govern the development and conduct of OOS operations. This policy regime should take significant cues from the UN Convention on the Law of the Sea (UNCLOS), which was enacted in 1982 to administer international waters and make clear the expected conduct of all nations when utilizing or traversing the high seas. Among other provisions, the

UNCLOS treats international waters as a common resource and directs all nations to behave in a cooperative manner, in addition to making clear the demarcations between territorial waters and international waters (UN Convention on the Law of the Sea, 1982). Thus, it sets out clear guidelines for conduct and ownership, which an effective OOS regime must also do. Indeed, no one desires a "Wild West" development phase in the beginnings of OOS operations; rather, all space-faring entities would be much better served by clear, unambiguous regulations which direct servicing missions in the same way that ships are governed on the high seas. Of course, some nations do not follow all the guidelines set out in the UNCLOS, and there may be some disagreement over the interpretation of its provisions, but what is most important is that there is a widely-acknowledged and clear reference policy which all nations can turn to when settling disputes. To have the greatest chance of success and sustainability, OOS must have the same.

International Participation in an OOS Infrastructure

Finally, at the very outset of infrastructure development, policymakers must determine who will be allowed to participate and make use of servicing capabilities. While including international partners in the development of an OOS infrastructure would greatly increase the number of available customers as well as (possibly) promote a more cooperative and transparent operating environment, these benefits must be weighed against the increased security concerns associated with one nation's spacecraft performing services on those of another state. Policymakers must also take into account the likelihood for proliferation of separate, state-based servicing operations, which may occur if one nation excludes others from its infrastructure (similar to the way in which many national entities are developing their own global navigation systems, separate from GPS). Such a scenario could very well lead to an arms race in servicing capabilities, where several nations independently develop OOS infrastructures and are highly distrustful of other states' intentions.

Indeed, it would appear that DARPA anticipates such international squeamishness at the prospect of the U.S. having its own OOS infrastructure, as it calls for the development of completely commercially-owned and –operated servicing spacecraft (DARPA, 2014). The fact that the infrastructure would be fielded as a commercial venture would hopefully negate

widespread fears of any malicious intent for the infrastructure, as a private firm could operate largely outside of concerns for international relations and service customers on a much more neutral, non-securitized footing. Still, an OOS infrastructure utilizing DARPA technologies and operated by a U.S. firm would remain a U.S. asset according to the 1967 Outer Space Treaty. Thus, while this solution largely removes any explicit military overtones to the project, international actors in space may still be wary of the intentions of the infrastructure (especially one which provides regular services to defense spacecraft).

In summary, there are no clear guidelines for determining the level of international participation in an OOS infrastructure, beyond the obvious need for transparency and significant efforts to dispel fears of OOS weaponization. Such considerations only underscore further the need for an UNCLOS-like code for operational conduct for servicing operations—one which establishes clear, unambiguous rules for the conduct of nations involved in servicing and implements significant safeguards against the use of OOS for malicious purposes. Operating the infrastructure as a commercial venture is an appropriate first step, but significant regulatory action, characterized by a high degree of international agreement, will still be required to ensure long-term safety and sustainability of the infrastructure.

Conclusions and Future Work

This thesis provides several useful insights with regards to the design and development of orbital infrastructure. For a LEO communications and data relay infrastructure with the requirement of constant, global coverage for LEO customers, it was shown that an optimal design point exists between a relatively small constellation of large, powerful spacecraft and a very large, disaggregated constellation of small, much less powerful satellites. It was also shown that a conservative estimate for the total cost of such a large-scale infrastructure is less than $5 billion, and it was postulated that even better performing, more cost-effective designs could be achieved through the use of more advanced communications technologies (as this analysis only considered basic RF communications links). Then, it was subsequently shown that, if the infrastructure is not required to provide constant coverage, then there would still be substantial benefits to LEO customers. In this case, customers' ability to downlink data throughout an operational day is greatly increased, with this non-global infrastructure estimated to cost less than $1 billion. Finally, the system dynamics analysis of OOS implementation into a reference geosynchronous communications constellation showed that the viability and sustainability of the infrastructure depends greatly on the customer's willingness to incorporate servicing capabilities and make significant changes to their satellite design paradigms. Specifically, it was shown that decreases in spacecraft mass due to increased use of the OOS infrastructure lead to the greatest likelihood of lower costs, lower satellite design lifetimes, and increased industry experience. Furthermore, it was shown that a reference space architecture should be able to tolerate moderately higher servicing fees as it relies more and more on the OOS infrastructure. More in-depth summaries of the results for both infrastructure cases are provided in the sections below.

As the results have shown, a communications and data relay infrastructure, based in low-Earth orbit, could offer significant benefits to participating spacecraft, especially LEO imagery constellations composed of many small spacecraft which place a high value on delivering large amounts of timely data. These results are intended to be generalizable to a variety of LEO constellations, especially those which seek greater performance through disaggregation of assets and through the networking of many orbital assets. However, because this analysis does not consider more advanced communications technologies (such as optical communications), and because industry-standard cost models were used, these results may provide less insight as to the

design of more advanced, non-traditional constellation designs. Additionally, an on-orbit servicing infrastructure could radically alter current space architecture paradigms, leading to less expensive, less massive, shorter-lived spacecraft which, because of the capability to repair and upgrade on orbit, would grant satellites designers the chance to back away from the space industry's excessively risk-averse posture and explore new methods of constellation management. The results for the OOS infrastructure apply primarily to the case of very large, standardized communications spacecraft at geosynchronous orbit, which are considered to be an important first customer for such a project. However, this analysis is less generalizable to other types of space architectures, such as small spacecraft in LEO, which have not exhibited the same trends in increasing spacecraft mass, power, and design lifetime.

Finally, such infrastructures were shown to be very unlikely to come cheaply, and if they are to be implemented they will almost certainly require significant collaboration between public and private partners. The fact that much of the technical demonstration of the basic operations of an OOS infrastructure have been conducted by public agencies like NASA and DARPA shows that private industry is unlikely to undertake such large-scale projects on their own (or at the very least, private firms are unlikely to conduct the initial on-orbit testing and validation of technologies required before fielding a mature infrastructure, which is itself quite expensive and risky). Additionally, the communications infrastructure aiding NASA science and human exploration missions acts as a highly beneficial analogue to the development of a more publicly-available network of communications spacecraft. Furthermore, both types of infrastructure are accompanied by significant policy considerations, and the establishment of effective policy regimes to handle such issues as ownership, operational codes of conduct, spectrum management, and liability for damages will be crucial to the long-term success of these infrastructures. Much like the Interstate Highway System, National Airspace System, or the GPS constellation, these projects could be undertaken as a means of promoting both public and private welfare and enabling future economies that few would have originally imagined.

LEO Communications and Data Relay Infrastructure

For a LEO communications and data relay infrastructure, there are two key conclusions. First, for an infrastructure that is designed for global connectivity, there is a definite break-point in constellation design, where further miniaturization and disaggregation of assets begins to yield diminishing increases in system performance. This is to be expected; after all, there are physical limits to how small an effective communications satellite can be, and simply launching more low-powered spacecraft will not necessarily allow such a constellation to match the performance of one which is composed of a smaller number of more powerful spacecraft. However, it is also acknowledged that the ultimate design of a global LEO data relay infrastructure will depend largely on its intended usage and the number of customers expected to be serviced. For instance, if customers value high-speed connections with the infrastructure, rather than the low data rate crosslinks recommended in these results, then this must be taken into account in the final design, else the infrastructure will not provide the degree of performance customers demand.

For a LEO infrastructure not requiring global coverage, the main conclusion of this thesis concerns the types of customers who would benefit the most from an orbital communications and data relay infrastructure. It was shown that large flagship missions like NASA's Landsat program are already well-supported by a large infrastructure of ground stations, in addition to having spacecraft which are more than adequately sized to downlink large quantities of data, and so these types of programs would be unlikely to gain as much value from an orbital infrastructure. New imaging constellations, however, which are composed of smaller spacecraft and which may not have access to the resources of a large organization like NASA would likely see a great deal of value in the use of an infrastructure. These types of customers, represented by new firms like SkyBox and Planet Labs, were shown to have much greater increases in performance, and there is likely a strong business case for these customers to contribute to such an infrastructure. Furthermore, this result is supported by the fact that there is a burgeoning LEO economy, which is being driven in large part by new applications of small, low-powered spacecraft (Morring Jr., 2015).

Furthermore, should such a LEO infrastructure be implemented, its best possible configuration would be high altitude, polar orbiting, and composed of several relatively small spacecraft in just one or two orbital planes. In fact, it was found that 16 satellites in one plane

was actually comparable in price to eight satellites in one plane, due to the decreased power and mass requirements of each spacecraft in the larger constellation. Also, the type of data that customers want out of an infrastructure is an important design point; an infrastructure designed to maximize total data downlinked each day can be quite different from one that is designed to maximize instantaneous communications links. Based upon the infrastructure configurations found to be most optimal, it was found that an infrastructure utilizing "rings" of spacecraft in the same orbital plane does not pay any penalties for "store and forward" data downlinks. One infrastructure craft (and oftentimes two) is in contact with a ground station at or near one of the Earth's poles at all times, and thus the infrastructure can begin forwarding and downlinking customer data as soon as the customer spacecraft comes into range of any portion of the infrastructure ring. This increases the amount of "priority" data downlinked as well as simplifies data storage and forwarding issues.

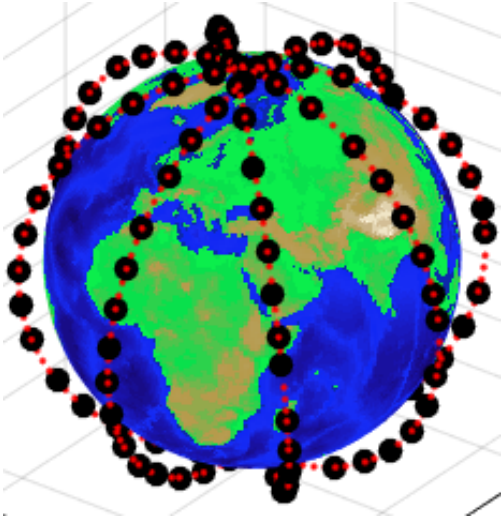


Figure 44. This constellation configuration was identified as one of the optimal solutions for providing constant coverage to LEO customer spacecraft.

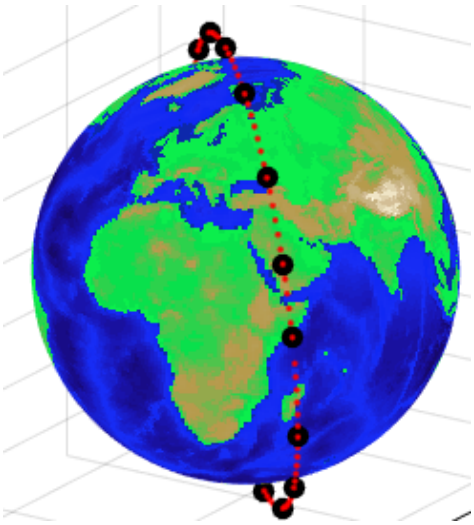


Figure 45. For a LEO infrastructure not requiring global coverage, it was determined that one orbital plane of 16 infrastructure spacecraft could still provide substantial benefits to LEO customers at signficantly lower costs.

An important factor of this analysis is that only RF communications links were considered. For future work, an analysis of an infrastructure utilizing optical communications technologies, either to the ground or between spacecraft or both, would represent another interesting and perhaps even more powerful means of providing infrastructure services. Optical links would allow for much higher data transmission rates, but also has the potential to introduce greater complexity to the infrastructure system. Similarly, an analysis of an infrastructure utilizing highly directional antennas could also lead to interesting results, again with the caveat that more consideration must be given to possible increases in complexity. Furthermore, infrastructure configurations beyond LEO should be considered, as well as more novel methods of implementing this type of system than through just dedicated infrastructure spacecraft. For instance, as Golkar et. al. envision, hosted payloads may be one means of implementing infrastructure (Golkar, 2013). In addition, the value of a LEO infrastructure should be examined through methods other than a performance-to-cost framework, as it is expected that the availability of an infrastructure would offer more than simple performance boosts to customer spacecraft. Rather, it is also highly likely to offer value in the form of system attributes like flexibility, survivability, responsiveness, testability, and similar "ilities" which are gaining much greater prominence in the field of systems engineering.

The policy implications of a LEO infrastructure are also substantial. In order to fund and develop such a project, a classical approach would be for a government agency such as NASA to complete the infrastructure largely on its own, and then provide the service as a public good with minimal to no fees. A similar approach would be to develop the infrastructure as a military system, which in addition to servicing national security assets could also be open for civilian applications. While such options have worked for infrastructure projects in the past, perhaps a better option would be to develop the infrastructure within a public-private partnership; one which would have significant public input, but would ultimately be designed, manufactured, and operated as a commercial product. Such arrangements have become increasingly prevalent and successful, as the space industry has shifted from an almost entirely public-funded endeavor to private enterprise. Finally, the implementation of any communications and data relay infrastructure necessarily includes several political and legal considerations, and this thesis identified electromagnetic spectrum management, orbital debris concerns, and issues surrounding optical communications as the major regulatory hurdles to a LEO infrastructure.

On-Orbit Servicing Infrastructure

The primary conclusion from the system dynamics analysis of the implementation of an OOS infrastructure is that, for the greatest chance of long-term sustainability and benefits from the infrastructure, then customers must be willing to make significant architectural design changes to fully incorporate OOS capabilities. It was shown that a shift in risk attitude alone from the use of OOS was not enough to reverse damaging upwards trends in satellite mass and design life in a system dynamics model simulating a theoretical constellation of geosynchronous communications satellites. Indeed, in such a case where designers continue to increase spacecraft mass generation after generation, as they have for the last several decades, then the added costs of OOS may actually exacerbate damaging industry trends and further degrade industry knowledge. More or less freezing spacecraft mass at current levels as OOS is implemented was shown to largely halt potentially unsustainable industry trends, but in order to actually reverse trends and return spacecraft design points to more sustainable levels, spacecraft designers must be willing to make significant alterations to architectures. Such changes would

likely include much smaller fuel tanks or less robust components, as constellations begin to rely on the OOS infrastructure for periodic refueling and technology upgrades.

It was also shown that OOS should be offered at as low cost as possible to the customer, as this allowed for the greatest benefits in design trends and overall sustainability. However, OOS does not necessarily have to be completely zero cost; indeed, as the customers incorporate more and more OOS capabilities into their architectures, the results suggest that they are then able to tolerate moderately higher servicing fees. Furthermore, cost of the infrastructure was shown to be a much less significant factor in the success of OOS than the overall level of incorporation of servicing capabilities into spacecraft design and operations. This result should be reassuring to proponents of an OOS infrastructure, as the tolerance for servicing fees means that there are several viable options available to policymakers for the ultimate development of the project.

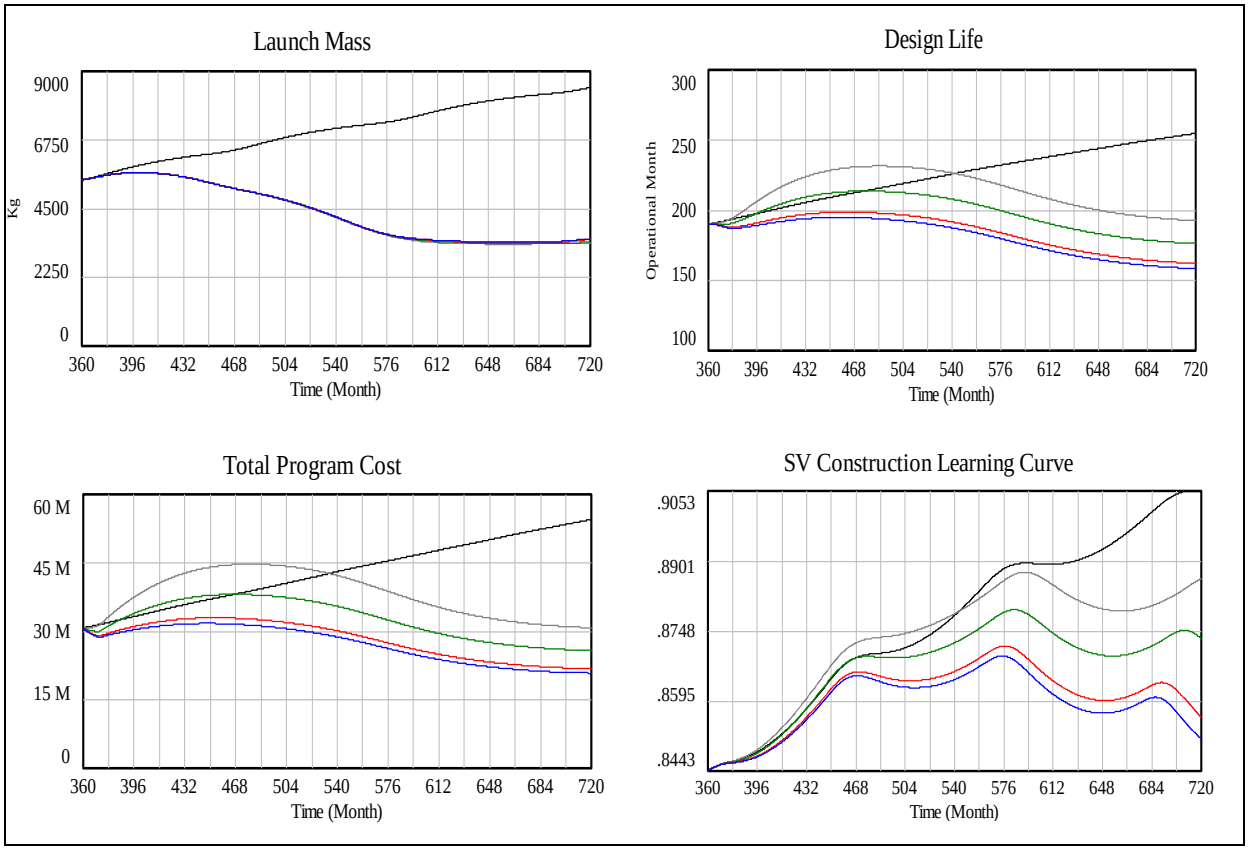


Figure 46. These graphs display the most optimal case for OOS implementation; one in which spacecraft mass is reduced as OOS is incorporated into the reference architecture, with the blue and red lines representing free and low cost servicing, respectively.

Along these lines, several options for the funding and development of OOS were outlined and discussed. Like the LEO communications and data relay infrastructure, the classic approach to such a project would be to have the OOS infrastructure developed as a NASA or DoD project, primarily as an asset for use with their own systems but which can still be made available for civilian applications. However, such policy regimes seemed less likely to be implemented—NASA or another public agency may not have the necessary room in their budgets for such a project, and the DoD may also face resistance to fielding such a system, due to weaponization concerns. Instead, there is a much greater likelihood that OOS will be implemented through a public-private partnership, where public agencies provide much of the funding and oversight, while a private firm conducts the manufacturing and operation of the infrastructure. Indeed, the groundwork for such an arrangement is already in place; NASA and DARPA have already conducted much of the initial development and testing of OOS technologies, and DARPA has recently issued a request for information from private contractors interested in the many aspects of an OOS project (DARPA, 2014).

Finally, several political and legal issues surrounding an OOS infrastructure were also discussed, with the six main areas of concern being responsibility and ownership, spacecraft insurance, on-orbit operational regulation, assurance of non-weaponization, spectrum and slotting, and imaging regulations (Belcher, Freese, Laygo, & Osborn, 2014). For the overall framework for OOS policy, it is asserted that regulation of the infrastructure can borrow heavily from the UN Convention on the Law of the Sea, which sets out clear rules and guidelines for the conduct of maritime vessels on international waters. Recognizing that space is the province of all nations, and that the long-term success of OOS will depend heavily on transparent, non-confrontational use of servicing technologies, policymakers can borrow heavily from the regulation of other international arenas as they establish the guiding policies for an OOS infrastructure.

Future work in the assessment of OOS concepts could focus on more specific price points for successful implementation. For instance, while this thesis analyzed the way in which OOS could affect general space industry trends, there are likely certain prices at which OOS would be most attractive to customers, depending on the type of spacecraft being serviced and the specific services being provided. While some argue that the cost of replacing a failed spacecraft is a

good guidelines to how expensive OOS can be (Benedict, 2013), it is likely that the actual cost tolerance of customers is more nuanced than this, and will depend heavily on the value being provided to specific types of customers. For instance, Saleh et. al. have examined the value OOS provides to space architectures in the form of flexibility (Saleh, Lamassoure, Hastings, & Newman, 2003), demonstrating that it is likely that the ultimate rationale for an OOS infrastructure will not lie in a strictly cost-to-performance basis. Furthermore, a more detailed timeline for the implementation of OOS, perhaps beginning with the DARPA RFI and extending to a fully developed infrastructure, could also be developed and used as a guideline for policymakers moving forward on this concept.

Vision for Orbital Infrastructure

Space and spacecraft technologies have become an increasingly vital component of global economies, and human utilization of space is only expected to increase in scope and significance. Along these lines, it is now high time for the implementation of orbital infrastructures. These systems can offer many benefits to current space enterprises as well as aid in the development of entirely new ventures. An orbital communications and data relay infrastructure based in low-Earth orbit, providing data downlink and connectivity services to customer spacecraft, and an on-orbit servicing infrastructure, providing repair, rescue, refueling, and upgrade services to several different types of spacecraft, are concepts which could yield incredible value to space architectures, in addition to potentially ushering in an entirely new paradigm of spacecraft designed to make the greatest possible use of these infrastructures. The ultimate vision for a data relay infrastructure concept would be a LEO environment in which spacecraft operators have near-constant access to their assets, granting them the ability to downlink far more data as well as enable massively networked space enterprises—creating systems which truly capitalize on the benefits of miniaturization and disaggregation. Furthermore, the vision for OOS is an entirely new architectural paradigm, in which spacecraft would no longer be launched without any prospects for future access and physical modifications. Rather, an OOS infrastructure will be available to provide regular upgrades and improvements on existing assets, eliminating technological obsolescence and allowing spacecraft adapt to new missions throughout their orbital lifetimes.

An ideal path for implementation of these orbital infrastructures would follow largely from the discussion of policy options. For the LEO data relay infrastructure, it is envisioned that a national entity such as NASA or DARPA would solicit private industry for possible design solutions (borrowing heavily, of course, from the configurations identified in this thesis). Then, a public-private partnership would be established, where private industry takes on the vast majority of manufacturing and operational tasks, while public entities provide funding and also aid in the regulatory oversight of the infrastructure. The infrastructure's performance could first be demonstrated through the networking and downlinking of data from public space assets, and then private industry could be brought into the infrastructure through collaboration with public missions. Finally, new missions designed specifically for use of the infrastructure would be launched, and these missions would experience greatly increased connectivity and sharing of resources between assets, in addition to the downlinking of far more data to customers on the ground.

For the OOS infrastructure, much of the vision for implementation has already been realized. NASA and DARPA have successfully conducted the initial technical demonstrations of OOS capabilities, with more tests planned for the near future. Furthermore, DARPA has issued an initial request for information from industry players for the design of a geosynchronous OOS project, and this request explicitly called for the implementation of OOS as a public-private partnership (DARPA, 2014). From here, private industry will take over the construction and operation of the infrastructure, with public support in the form of funding and regulatory oversight. The initial viability of the infrastructure will then be demonstrated on defense and government assets, followed by a string of requests from private operators to have their own spacecraft receive services. Finally, geosynchronous orbit would enter an era marked by the regular provision of services to a variety of customer spacecraft, and space architectures would benefit significantly from the launching of less massive and more technologically-agile satellites.

It is expected that the implementation of these infrastructures would lead to an operational environment in space which is radically different than that of the past or present. Massive constellations of small spacecraft could be networked across the entire LEO regime, and rather than waiting for ground station passes to check up on assets, LEO operators could monitor and issue new commands to their constellations in real-time. Spacecraft in geosynchronous orbit or

elsewhere would no longer be untouchable islands, but rather would become operational platforms receiving regular replenishments and upgrades throughout their service lives. Furthermore, in addition to providing these substantial increases in performance, the space environment that infrastructure engenders should also be far more sustainable, accessible to new entrants, and available at lower costs than that experienced at any other point in the history of space utilization.

Technical feasibility of these concepts has long been proven. With sound policies for funding and development, and clear regulation to ensure fairness and transparency, these orbital infrastructure concepts can become reality, ensuring space remains a breeding ground for new technologies and new endeavors for centuries to come.

Bibliography

50th Space Wing Public Affairs. (2011, June 16). *50 SW completes GPS constellation expansion*. Retrieved January 2015, from Schriever Air Force Base: http://www.schriever.af.mil/news/story.asp?id=123260251

Akcam, B. K., Guney, S., & Cresswell, A. M. (2011). Major Issues in Mixed Use of Grounded Theory and System Dynamics Approaches in Qualitative Secondary Data. *The 29th International Conference of the System Dynamics Society.* Washington, DC.

Anderson, C. (2013). Rethinking public-private space travel. *Space Policy*, 266-271.

Anderson, P. R., & Bartamian, L. (2008). Growth Trends in Communication Satellites and the Impact on Satellite System Architecture. *26th International Communications Satellite Systems Conference (ICSSC)*.

Batt, H. W. (1984). Infrastructure: Etymology and Import. *Journal of Professional Issues in Engineering*, Vol. 110, p. 1-6.

Beeferman, L., & Wain, A. (2012). *Infrastructure: Defining Matters.* Cambridge, MA: Pensions and Capital Stewardship Program, Harvard Law School.

Belcher, D., Freese, S., Laygo, K., & Osborn, D. (2014). Analysis of United States Policy and Legal Impediments to On-Orbit Satellite Servicing Activities. *64th International Astronautical Congress*.

Benedict, B. (2013). Rationale for Need of In-Orbit Servicing Capabilities for GEO Spacecraft. *AIAA Space 2013*.

Center, N. G. (2001). *NASA Facts: Three Newly Designed Tracking and Data Relay Satellites to Help Replenish Existing On-Orbit Fleet.* Greenbelt, MD.

Clark, S. (2012, March 19). *Electric propulsion could launch new commercial trend.* Retrieved 2015, from Spaceflight Now: http://spaceflightnow.com/news/n1203/19boeing702sp/

Clark, S. (2015, March 1). *Boeing's first two all-electric satellites ready for launch.* Retrieved 2015, from Spaceflight Now: http://spaceflightnow.com/2015/03/01/boeings-first-two-all-electric-satellites-ready-for-launch/

Clore, R. E. (2012). *International Developments in Global Navigation Satellite Systems.* Washington, D.C.: U.S. Department of State: Office of Space and Advanced Technology.

Cohen, N., Richardson, G., Martinelli, S., & Betser, J. (2011). Extending Satellite Lifetimes in Geosynchronous Orbit with Servicing. *AIAA Space 2011*.

Courtland, R. (2014, December 10). *DARPA Prepares to Launch "Satlets"*. Retrieved 2015, from IEEE Spectrum: http://spectrum.ieee.org/tech-talk/aerospace/satellites/darpas-satellite-revival-program-gears-up-for-first-launch

Coyle, G., & Exelby, D. (2000). The validation of commercial system dynamics models. *System Dynamics Review, 16*(1), 27-41.

Curley, S., Yates, J., & Abrams, R. (1986). Psychological Sources of Ambiguity Avoidance. *Organizational Behavior and Human Decision Processes, 38*, 230-256.

Danchik, R. J. (1998). An Overview of Transit Development. *Johns Hopkins APL Technical Digest, 19*(1), 18-26.

DARPA. (2014, September 3). *Wanted: Insights to Guide Creation of Robotic Satellite-Servicing Capabilities in Geostationary Earth Orbit.* Retrieved 2015, from http://www.darpa.mil/NewsEvents/Releases/2014/09/03.aspx

de Selding, P. B. (2014, November 5). With One Satellite to Spare, Iridium Counting on SpaceX to Maintain Launch Tempo. *SpaceNews.com*.

de Weck, O., Roos, D., & Magee, C. (2011). *Engineering Systems: Meeting Human Needs in a Complex Technological World.* Cambridge: MIT Press.

de Weck, O., Ross, A., & Rhodes, D. (2012). Investigating Relationships and Semantic Sets amongst System Lifecycle Properties (Ilities). *Third International Engineering Systems Symposium, CESUN 2012, Delft University of Technology*.

Ellis, R. E. (2001). *The Use of System Dynamics in a Large-Scale System Engineering Effort: Case Study of the Deepwater Project.* McLean, VA: JAYCOR, INC.

Emmons, D., Bitten, R., & Freaner, C. (2007). Using Historical NASA Cost and Schedule Growth to Set Future Program and Project Reserve Guidelines. *IEEE Aerospace Conference*.

Escobedo Jr., V. M. (2014, October 21). *Robotic Refueling Mission (RRM)*. Retrieved January 2015, from International Space Station Research & Technology: http://www.nasa.gov/mission_pages/station/research/experiments/778.html#publications

European Space Agency. (2014, December). *EDRS.* Retrieved 2015, from Communications & integrated applications: http://www.esa.int/Our_Activities/Telecommunications_Integrated_Applications/EDRS/(print)

Federal Aviation Administration. (2006). *Advisory Circular 150/5190-7: Minimum Standards for Commercial Aeronautical Activities.* U.S. Department of Transportation.

Federal Aviation Administration. (2014). *Federal Aviation Regulations/Aeronautical Information Manual (FAR/AIM).* Aviation Supplies and Academics, Inc.

Federal Aviation Administration. (2015, February). *A Brief History of the FAA*. Retrieved from About FAA: https://www.faa.gov/about/history/brief_history/

Figliola, P., Behrens, C., & Morgan, D. (2006). *U.S. Space Programs: Civilian, Military, and Commercial.* Congressional Research Service.

Flock 1 Imaging Constellation. (2014). Retrieved from Earth Observation Portal: https://directory.eoportal.org/web/eoportal/satellite-missions/f/flock-1

Forrester, J. (1968). *Principles of Systems.* Cambridge: Productivity Press.

Fox, B., Brancato, K., & Alkire, B. (2008). *Guidelines and Metrics for Assessing Space System Cost Estimates.* RAND Corporation.

Galileo Services. (2015). Retrieved from European Global Navigation Satellite Systems Agency: http://www.gsa.europa.eu/galileo/services

Golkar, A. (2013). Federated Satellite Systems (FSS): A Vision Towards an Innovation in Space Systems Design. *IAA Symposium on Small Satellites for Earth Observation*.

Golkar, A. (2013). Federated Satellite Systems: A Case Study on Sustainability Enhancement of Space Exploration Systems Architectures. *64th International Astronautical Congress, Beijing, China*.

Golkar, A., & Lluch, I. (2013). Satellite-to-Satellite Coverage Optimization Approach for Opportunistic Inter-Satellite Links. *IEEE Aerospace Conference*.

Guier, W., & Weiffenbach, G. C. (1997). Genesis of Satellite Navigation. *John Hopkins APL Technicall Digest, 19*(1), 178-181.

Hashimoto, M. (2009). *Public-Private Partnerships in Space Projects: An Analysis of Stakeholder Dynamics.* SM Thesis.

Hegarty, C., & Chatre, E. (2008, December). "Evolution of the Global Navigation Satellite System (GNSS)". *Proceeding of the IEEE, 96*(12), 1902-1917.

History of the Interstate Highway System and Interstate FAQ's. (2015). Retrieved January 2015, from U.S. Department of Transportation, Federal Highway Administration: http://www.fhwa.dot.gov/interstate/history.htm

Homer, J. B. (1997, Winter). Structure, Data, and Compelling Conclusions: Notes from the Field. *System Dynamics Review*, 293-309.

International Space Station: Research and Technology. (2015, March). Retrieved from NASA: https://www.nasa.gov/mission_pages/station/main/

Iridium Global Network: Satellite Constellation. (2015). Retrieved from Iridium Communications Inc.: https://www.iridium.com/about/IridiumGlobalNetwork/SatelliteConstellation.aspx

Iridium NEXT. (2015). Retrieved from Iridium Communications Inc.: https://www.iridium.com/about/IridiumNEXT.aspx

Jones, P., & Spence, B. (2011, August). Spacecraft Solar Array Technology Trends. *IEEE A&E Systems Magazine*, 17-28.

Lane, D. (1994). Social Theory and System Dynamics Practice. *International System Dynamics Conference*, (pp. 53-66).

Lane, D., & Oliva, R. (1998). The greater whole: Towards a synthesis of system dynamics and soft systems methodology. *European Journal of Operations Research, 107*, 214-235.

Lapowsky, I. (2015, January 22). *The Startup That Could Beat SpaceX to Building a Second Internet in Space*. Retrieved from Wired. com: http://www.wired.com/2015/01/greg-wyler-oneweb/

Listner, M. (2011, October 17). *Revisiting the Liability Convention: reflections on ROSAT, orbital space debris, and the future of space law*. Retrieved 2015, from The Space Review: http://www.thespacereview.com/article/1948/1

Long, A., Richards, M., & Hastings, D. E. (2007). On-Orbit Servicing: A New Value Proposition for Satellite Design and Operation. *Journal of Spacecraft and Rockets, 44*(4), 964-976.

Mai, T. (2014, October). *Near Earth Network*. Retrieved 2015, from NASA Space Communications and Navigation: http://www.nasa.gov/directorates/heo/scan/services/networks/txt_nen.html#.VOeDkeZ4pKA

Mai, T. (2014, October 3). *Near Earth Network (NEN)*. Retrieved from NASA Space Communications and Navigation (SCaN): https://www.nasa.gov/directorates/heo/scan/services/networks/txt_nen.html

Meyer, R. (2014, Jan 14). *Silicon Valley's New Spy Satellites.* Retrieved from The Atlantic: http://www.theatlantic.com/technology/archive/2014/01/silicon-valleys-new-spy-satellites/282580/

Morring Jr., F. (2015, April 1). *Investors Bullish on Pending LEO Economy.* Retrieved from Aviation Week and Space Technology: http://aviationweek.com/space/investors-bullish-pending-leo-economy?eid=forward

NASA Goddard Space Flight Center. (2010). *On-Orbit Satellite Servicing Study: Project Report.* Greenbelt, MD.

NASA Goddard Space Flight Center. (2012). *Space Network Users' Guide (SNUG): Revision 10.* Greenbelt MD.

NASA-Goddard Space Flight Center. (2012). *NASA Facts.* Retrieved 2015, from Tracking and Data Relay Satellite: Continuing the Critical Lifeline of Communications: http://tdrs.gsfc.nasa.gov/assets/images/Publications/TDRS%20FS%20full.large.pdf

National Space Policy of the United States of America. (2010, June). Retrieved 2015, from https://www.whitehouse.gov/sites/default/files/national_space_policy_6-28-10.pdf

Niles, L. (2014, February). *Largest Flock of Earth-Imaging Satellites Launch into Orbit From Space Station*. Retrieved 2015, from Space Station Research & Technology: http://www.nasa.gov/mission_pages/station/research/news/flock_1/

OECD. (2014). *The Space Economy at a Glance 2014.* OECD Publishing.

Office of the Under Secretary of Defense for Acquisition, T. a. (May 2003). *Report of the Defense Science Board/Air Force Scientific Advisory Board Joint Task Force on Acquisition of National Security Space Programs.*

Orbital Express On-Orbit Mission Updates. (2007, July). Retrieved January 2015, from DARPA Tactical Technology Office: http://archive.darpa.mil/orbitalexpress/mission_updates.html

Pace, S., Frost, G. P., Lachow, I., Frelinger, D. R., Fossum, D., Wassem, D., et al. (1995). *The Global Positioning System: Assessing National Policies, Appendix B: GPS History, Chronology, and Budgets.* RAND Corporation.

Parks, C., & Berger, B. (2006, October 31). *NASA's Mission to Service Hubble in 2008 Will Cost $900 Million*. Retrieved from Space.com: http://www.space.com/3057-nasa-mission-service-hubble-2008-cost-900-million.html

Phoenix Goals. (2015). Retrieved January 2015, from DARPA Tactical Technology Office: http://www.darpa.mil/Our_Work/TTO/Programs/Phoenix/Program_Goals.aspx

Pruyt, E. (2006). *What is System Dynamics? A Paradigmatic Inquiry.* Vrije Universiteit Brussel, Business School Solvay, .

Rendleman, J. D., & Smith, M. (2011). Law and Policy Considerations. In J. R. Wertz, D. F. Everett, & J. J. Puschell, *Space Mission Engineering: The New SMAD* (pp. 337-352). Hawthorne CA: Microcosm Press.

Richards, M. (2006). *On-Orbit Serviceability of Space System Architectures.* Cambridge: MIT.

Robotic Refueling Mission. (2015, January). Retrieved from NASA Satellite Servicing Capabilities Office: http://ssco.gsfc.nasa.gov/robotic_refueling_mission.html

Ross, A. M., & Hastings, D. E. (2005, July). The Tradespace Exploration Paradigm. *INCOSE International Symposium.* Rochester NY.

Saleh, J. (2008, January). Flawed Metrics: Satellite Cost per Transponder and Cost per Day. *IEEE Transactions on Aerospace and Electronic Systems, 44*(1), 147-156.

Saleh, J., Hastings, D., & Newman, D. (2002, March-April). Spacecraft Design Lifetime. *Journal of Spacecraft and Rockets, 39*(2), 244-257.

Saleh, J., Lamassoure, D., Hastings, D., & Newman, D. (2003). Flexibility and the Value of On-Orbit Servicing: New Customer-Centric Perspective. *Journal of Spacecraft and Rockets*.

Saleh, J., Torres-Padilla, J., Hastings, D., & Newman, D. (2006, January-February). To Reduce or to Extend a Satellite Design Lifetime? *Journal of Spacecraft and Rockets, 43*(1), 207-217.

Satellite Collision Leaves Significant Debris Clouds. (2009, April). *NASA Orbital Debris Program Office: Orbital Debris Quaterly News, 13*(2), pp. 1-2.

Skybox Imaging. (2015). *Technology*. Retrieved 2015, from Skybox Imaging: http://www.skyboximaging.com/technology

SkySat Imaging Program of SkyBox Imaging. (2014). Retrieved from Earth Observation Portal: https://directory.eoportal.org/web/eoportal/satellite-missions/s/skysat

Space Segment. (2014, December 12). Retrieved January 2015, from GPS.gov: http://www.gps.gov/systems/gps/space/

Sterman, J. (1989, March). Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decisionmaking Experiment. *Management Science, 35*(3), 321-339.

Sterman, J. (1994, Summer-Fall). Learning in and about complex systems. *System Dynamics Review, 10*(2-3), 291-330.

Sterman, J. (2000). *Business Dynamics: Systems Thinking and Modeling for a Complex World.* Boston: Irwin/McGraw-Hill.

Stone, C., & Fershtut, A. (2010). Increased Government Involvement in Developmental Test and Other Lessons Learned from Responsive Space. *AIAA Space 2010*.

System Dynamics for Academia. (2014). Retrieved from System Dynamics Society: http://www.systemdynamics.org/sd-for-academia/

Systems Toolkit (STK) version 10. (2015). Retrieved from Analytical Graphics, Inc. (AGI): http://www.agi.com/products/stk/stk-10/

Team Hubble: Servicing Missions. (2015). Retrieved January 2015, from Hubble Site: http://hubblesite.org/the_telescope/team_hubble/servicing_missions.php

The Tauri Group. (2014). *State of the Satellite Industry Report.* Satellite Industry Association.

Tracking and Data Relay Satellite Project. (2012, October). *Tracking and Data Relay Satellite: Continuing a Critical Lifeline of Communications.* Retrieved January 2015, from NASA Goddard Space Flight Center: http://tdrs.gsfc.nasa.gov/assets/images/Publications/TDRS%20FS%20full.large.pdf

Trautmann, S., Vieider, F., & Wakker, P. (2008, May). Causes of ambiguity aversion: Known versus unknown preferences. *Journal of Risk and Uncertainty, 36*, 225-243.

Tversky, A., & Kahneman, D. (1974, September). Judgement under Uncertainty: Heuristics and Biases. *Science, 185*(4157), 1124-1131.

Tversky, A., & Kahneman, D. (1979, March). Prospect Theory: An Analysis of Decision Under Risk. *Econometrica, 47*(2), 263-291.

U.S. Air Force Space Command. (2013). Resiliency and Disaggregated Space Architectures.

U.S. Geological Survey. (2014, November 10). *Landsat Ground Stations*. Retrieved 2015, from Landsat Missions: http://landsat.usgs.gov/about_U_S_ground_stations.php

U.S. Geological Survey. (2014, September). *Landsat-8 History*. Retrieved 2015, from Landsat Missions: http://landsat.usgs.gov/about_ldcm.php

(1982). *UN Convention on the Law of the Sea.* United Nations General Assembly.

UN Office for Outer Space Affairs. (1967). *Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, Including the Moon and Other Celestial Bodies.* United Nations.

Union of Concerned Scientists. (2014, August 1). *UCS Satellite Database.* Retrieved from http://www.ucsusa.org/nuclear_weapons_and_global_security/solutions/space-weapons/ucs-satellite-database.html#.VF03kSJ4pKA

United States Policy. (2013, October 21). Retrieved January 2015, from GPS.gov: http://www.gps.gov/policy/

Vilar, E., & Austin, J. (1991). Analysis and Correction Techniques of Doppler Shift for Non-Geosynchronous Communication Satellites. *International Journal of Satellite Communications, 9*, 123-136.

Walker, J. (2007, March). Orbital Express. *Fact Sheet: Defense Advanced Research Projects Agency*. Arlington, VA.

Weeden, B. (2014). *Space Situational Awareness Fact Sheet.* Secure World Foundation.

Weingroff, R. M. (2011, April 7). *"When did the Federal Government begin collecting the gas tax?"*. Retrieved January 2015, from Ask the Rambler, Federal Highway Administration: http://www.fhwa.dot.gov/infrastructure/gastax.cfm

Wertz, J. (2011). Introduction. In J. Wertz, D. Everett, & J. Puschell (Eds.), *Space Mission Engineering: The New SMAD* (pp. 1-28). Hawthorne, CA: Microcosm Press.

Wertz, J. (2011). Reducing Space Mission Cost and Schedule. In J. Wertz, D. Everett, & J. Puschell (Eds.), *Space Mission Engineering: The New SMAD* (pp. 355-398). Hawthorne, CA: Microcosm Press.

Wertz, J. (2013). *Reinventing Space: Dramatically Reducing Space Mission Cost — Reducing Launch Cost.* SpaceNews.com.

Wertz, J. R., Everett, D. F., & Puschell, J. J. (Eds.). (2011). *Space Mission Engineering: The New SMAD.* Microcosm Press.

Winternitz, L. M., Bamford, W. A., & Heckler, G. W. (2009, August). A GPS Receiver for High-Altitude Satellite Navigation. *IEEE Journal of Selected Topics in Signal Processing, 3*(4), 541-556.

Wolstenholme, E. (1983, Winter). System Dynamics: A System Methodology or a System Modelling Technique. *DYNAMICA, 9*(II), 84-90.

Yin, R. K. (2014). *Case Study Research: Design and Methods (5th Edition).* Thousand Oaks, CA: SAGE Publications Ltd.