Massachusetts Institute of Technology
Engineering Systems Division

Working Paper Series

ESD-WP-2003-01.01-*ESD Internal Symposium*

INCORPORATING UNCERTAINTY INTO CONCEPTUAL DESIGN OF SPACE SYSTEM ARCHITECTURES

Daniel E. Hastings

Annalisa L. Weigel

Myles A. Walton

MAY 29-30, 2002

Incorporating uncertainty into conceptual design of space system architectures

Daniel E. Hastings, Annalisa L. Weigel, Myles A. Walton

MIT Engineering Systems Division &
Department of Aeronautics and Astronautics

Abstract

The environment in which space systems are developed and operated can be classified as nothing less than dynamic. However, it is clear that the methods and tools relied on in conceptual design are based on static assumptions and leave little room for anything more than snapshots of the product and its environment. This paper introduces an approach to challenge that model and instead quantify and compare space system architectures around the central theme of uncertainty, with emphasis on policy uncertainty, as well as, technical and market uncertainty. Two cases of implementation are presented and three generalized principles are proposed that flow from the analysis: 1) engineering systems must be designed with uncertainty as one of the central organizing principles, 2) since engineering systems have management and social dimensions and thus involve human interactions, there is an irreducible uncertainty associated with these dimensions that will affect the design of the system, and 3) uncertainty in use may allow the engineering system to satisfy quite different missions from the original one intended.

Introduction

There has been much debate about what constitutes an engineering system and if there are any fundamental principles underlying the design and operation of these systems. In this paper we examine a particular class of engineering systems as a paradigm for what can be learned about the broader notions of engineering systems. The class of engineering systems that we will consider is military and commercial space system architectures. We will take the point of view in this paper that engineering systems are complex systems that are technologically enabled (i.e. depend for their existence on products created by technology and using technology). They are also systems that involve interactions at the enterprise or societal levels.

The term space system architecture is used here to cover all the aspects of the space system and their interactions and interfaces that allow the function to be delivered to the users. For example, the Global Position System (GPS) [1] space system architecture delivers real time precise position and timing information to users on a worldwide basis. It is composed of 24 satellites in Medium Earth Orbit (MEO) that emit timing signals, a set of distributed ground stations that communicate with the satellites, a set of command and control centers that periodically inform the satellites of their positions and the user handsets that calculate time and position based on the signals from four GPS satellites. We choose to focus on military and commercial space systems rather than scientific space systems (e.g Hubble or Galileo) because they have the following properties. They are technologically enabled, complex (literally “rocket science”), global in scale (literally in the sense that the functions of such systems can be delivered all over the globe) and involve direct interactions with large numbers of people (in the military case with large numbers of dispersed forces, both friendly and hostile, in the commercial case with large markets needing communications or remote sensing). Typically, scientific systems have direct interactions with small numbers of scientists and the interaction at the societal level is through the support of national governments. For example, Hubble operations are funded through NASA that is funded by the US Congress but the direct recipients of the functions of Hubble are a small number of astronomers.

An excellent example of a complex space system architecture can be seen by looking at the systems that provide worldwide military communications. The national military strategy relies on information superiority and the need for worldwide communications has increased dramatically in recent years. In Desert Storm, the total data rate required was 100 Mbps while Allied Force, which deployed a force only 10% of the size of that used in Desert Storm forces, required 250 Mbps. This increased data rate was provided largely by commercial geo-synchronous satellite communication architectures (over 75% by the end of Allied Force) supplemented by highly protected military communication satellites. The highly protected military satellite communications architecture right now is several MILSTAR satellites [1]. These are designed to provide worldwide communications to the National Command Authority in the event of a nuclear conflict. Thus, the architecture consists of some very complex and expensive satellites along with ground stations combined with less expensive commercial satellites with more capacity but owned by different stakeholders.

The DoD also makes heavy use of other complex satellite architectures. The Defense Support Program (DSP) [1] provides missile warning at both the strategic and tactical level. This consists of a set of geo-synchronous infrared sensing satellites, a set of communication links, and a distributed set of information fusion centers as well as trained operators who interpret the data. The Global Positioning System as discussed above provides precision navigation and timing services worldwide.

DSP, MILSTAR and GPS are all examples of (military) space system architectures. They have a number of characteristics that are of interest from the viewpoint of engineering systems. On the positive side, they are very high performance systems. In the case of DSP they give the US a strategic warning capability that only the Soviets matched. In the case of GPS, they have helped create a whole new industry based around knowing one's position (Hertz NeverLost, plowing with GPS etc) as well as leading to a new way of fighting wars where bombs can be aimed as accurately as bullets. However, they have proven to be expensive to build (MILSTAR costing over a billion per satellite) and/or operate (GPS costing hundreds of millions a year to operate even though the satellite cost is only approximately 30 million per unit).

These space system architectures traditionally define "cutting edge technology". However, while most often these architectures, once constructed, deliver excellent performance, it is extremely rare that they deliver the initially promised performance on the initially proposed cost and schedule. There are several reasons for this but a large amount of the blame lies in the fact that there is a little understanding of how to incorporate uncertainty into the design process.

The development of space systems is subject to not only cost, technical and market uncertainties, but also to uncertainties from the policy domain. This paper introduces an approach to quantify and compare space system architectures under uncertainty, with emphasis on policy uncertainty as well as technical and market uncertainty. We then use the results to draw larger conclusions with respect to engineering systems.

The structure of the paper is as follows. We first define the major sources of uncertainty associated with military and commercial space system architectures. We then describe how we construct simulations of space system architectures in a way that allow both structured and unstructured uncertainty to be included and allow us to see large tradespaces of architectures and not just point designs. We then describe two specific architectural tradespaces. One is for an ionospheric mapper that will allow improvement in the ability to predict the effect of the ionosphere on GPS signals. The second is for a broadband communications architecture that will deliver space based T1 links to portable users on a worldwide basis. We use these two architectures as examples to illustrate the effects of market based, policy and technical uncertainty on the design of these types of systems. Finally we wrap up by considering the broader lessons for engineering systems.

Major Sources of Uncertainty

It is not trivial that this paper's content is focused on uncertainty, rather than risk. Uncertainty in this context is defined as our inability to deterministically predict an architecture's value to the stakeholders of the system, i.e. company, customer, shareholders, users etc. This is in contrast to the term risk that almost always reflects a negative meaning of the probability of loss or injury. The delineation is important, as it opens the research to aspects of uncertainty that may in fact be positive.

Table 1: Uncertainty Categorization

<table>
<tr><th>Development Uncertainty</th><th>Operational Uncertainty</th></tr>
<tr><td>Political Uncertainty- uncertainty of development funding instability</td><td>Political Uncertainty- uncertainty of operational funding instability</td></tr>
<tr><td>Requirements Uncertainty- uncertainty of requirements stability</td><td>Lifetime Uncertainty - uncertainty of performing to requirements in a given lifetime</td></tr>
<tr><td>Development Cost Uncertainty- uncertainty of developing within a given budget</td><td>Obsolescence Uncertainty – uncertainty of performing to evolving expectation in a given lifetime</td></tr>
<tr><td>Development Schedule Uncertainty- uncertainty of developing within a given schedule profile</td><td>Integration Uncertainty – uncertainty of operating within other necessary systems</td></tr>
<tr><td rowspan="2">Development Technology Uncertainty- uncertainty of technology to provide performance benefits</td><td>Operations Cost Uncertainty – uncertainty of meeting operations cost targets</td></tr>
<tr><td>Market Uncertainty-uncertainty in meeting demands of an unknown market</td></tr>
<tr><td colspan="2">Model Uncertainty</td></tr>
</table>

The first step in any uncertainty analysis should be to develop a holistic view of uncertainties of potential architectures that enumerates all of the primary sources of risk over the lifecycle of the space system. The uncertainty structure that was developed is presented in Table 1. This characterization helps to both encompass the various types of uncertainty but also serves as a framework for discussion.

From an aerospace perspective, the life-cycle view on uncertainty is significant because its operational existence is as significant as the development perspective. The reason this is typically overlooked is that the contractors and buyers are imminently interested in delivery of the product within time and fiscal constraints. The operational context is therefore a secondary priority. However, this framework gives us the opportunity to focus on the life-cycle uncertainty as life-cycle value has evolved to a design driver and decision criteria.

Quantifying Uncertainty in Space System Architectures

Risk and uncertainty are major decision criteria in the pursuit of space system design, and yet the ability to quantify and provide uncertainty information is not satisfactory. From interviews with space architects and policy decision makers, uncertainty and risk analysis in conceptual design in the space industry at present can be characterized as qualitative, expert driven and point based. Moreover, uncertainties are evaluated individually, assessed and addressed as unique and any calculations of these uncertainties are not embedded in the end models of the designs. Finally, they are usually accounted for after a point design of an architecture has been chosen.

A more complete approach to design would provide a method for enabling the quantification and aggregation of uncertainty, as well as an approach to integrate that information into the design models in the earliest stage of the design. We present two examples of incorporating uncertainty into the conceptual design of space systems in this paper. Both are enabled by a conceptual design technique that provides for the exploration of potential architectures that we briefly describe in the next section.

Modeling of Space System Architectures: GINA and MATE

GINA and MATE analysis techniques enable mathematical modeling of many architecture candidates, and evaluation of those candidates in various cost-performance tradespaces. GINA is the Generalized Information Network Analogy methodology, which had its beginnings at MIT. MATE is the Multi-Attribute Tradespace Exploration methodology, which also had its beginnings at MIT and built upon the GINA work.

The GINA methodology [2, 3], allows for the rapid comparison of space systems by mathematically modeling them as information transfer networks. GINA "is a hybrid of information network flow analysis, signal and antenna theory, space systems engineering and econometrics, and specifies measurable, unambiguous metrics for the cost, capability, performance and adaptability of any space system" whose mission is communications, navigation or remote sensing. GINA specifies satellite system attributes as either part of a "design vector" or a "constants vector." Attributes in the design vector vary across a given range, and distinguish one space system from another. Attributes in the constants vector remain unchanged across all space systems under consideration.

Rapid increases in computing power allowed GINA to expand over the years. Further work by the Space Systems Laboratory extended GINA analysis to examine thousands of architectures simultaneously, whereas Shaw [3] examined hundreds simultaneously. Jilla et al. [4] worked on incorporating optimization algorithms into GINA to efficiently search a very large tradespace enabled hundreds of thousands of architectures to be analyzed simultaneously.

MATE builds upon the GINA research, and generalizes space system performance modeling from an absolute cost-per-function scale to a scale based on the concept of utility in economic theory. As in GINA, MATE separates architecture attributes into two categories: Design vector and constants vector. The performance measure for MATE is not cost-per-function like GINA, but a utility of the space system. Overall utility is measured on several component dimensions by using certainty equivalent lotteries with the space system customer. While utility is not an absolute measure, it is a useful relative measure for comparing how well different space system candidates satisfy a user's needs [5].

MATE has been applied to a series of space system designs projects, including B-TOS and C-TOS. B-TOS is a terrestrial observer swarm of symbiotic distributed satellites whose mission is to map the ionosphere. BTOS is the first example case that is discussed using the GINA/MATE framework. C-TOS undertook the detailed design of the space part of the architecture defined in B-TOS.

B-TOS Case Study

B-TOS is a space-based atmospheric mapping mission to characterize the structure of the ionosphere using topside sounding techniques. Accomplishment of this mission on a global basis enables more precise corrections to GPS. The three primary goals of B-TOS are:

- Measurement of the ionosphere topside electron density profile
- Measurement of angle of arrival of signals from ground-based beacons
- Measurement of localized ionospheric turbulence

To accomplish these goals, the B-TOS space system uses a swarm architecture of distributed small satellites in multiple collaborating clusters. B-TOS is required to maintain at least a minimum altitude for topside sounding, operate at a frozen orbital inclination of 63.4 degrees, and use the Tracking and Data Relay Satellite System operated by NASA for communication with the ground. This is shown in Fig 1.

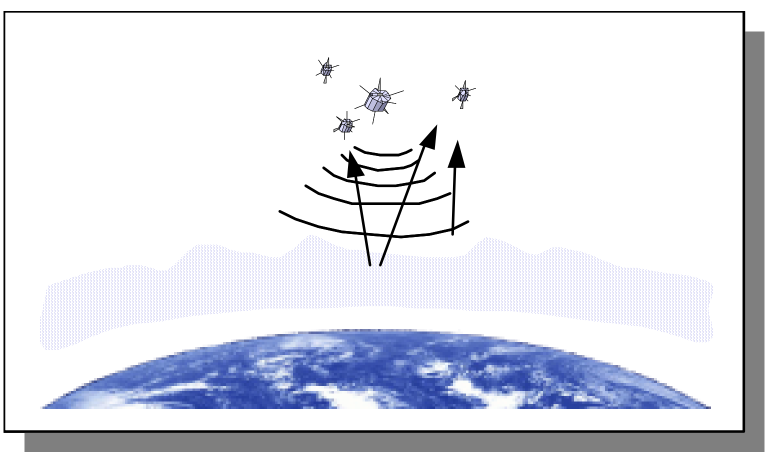


Figure 1. Conceptual rendering of a swarm of mother and daughter satellites performing topside sounding.

GINA and MATE analysis techniques were used to develop the B-TOS mission architecture candidates. The B-TOS GINA/MATE design vector and resulting tradespace enumeration is shown in Table 2. The completely enumerated B-TOS tradespace encompassed over 4000 unique architectures for the mission. Each architecture was evaluated for how much utility it provided the end user, as well as how much the architecture cost. The mapping of these two measures into utility-cost space produces a Pareto optimal frontier of B-TOS architectures as shown in Figure 2. A Pareto optimal solution is one that can't be improved in one dimension without sacrificing "goodness" in another dimension. Five architectures along the frontier are especially interesting, as they represent places on the frontier where its slope changes. These are the points labeled A, B, C, D, and E on Figure 2 and their corresponding design vector values are shown in Table 3 and Table 4 [6].

Table 2. B-TOS design vector variables and values.

Design Vector Variable	Values
Circular orbit altitude (km)	1100, 1300
Number of orbital planes	1, 2, 3, 4, 5
Number of swarms per plane	1, 2, 3, 4, 5
Number of satellites per swarm	4, 7, 10, 13
Radius of swarm (km)	0.18, 1.5, 8.75, 50
Payload capability	5 configurations of number of sounding antennas and capability, short and long range communication capability, and on-board data processing capability

Table 3. B-TOS Pareto optimal frontier architecture attributes.

Point	A	B	C	D	E
Altitude (km)	<-- 1100 -->				
Num of Planes	<-- 1 -->				
Swarms/Plane	1	1	1	1	2
Satellites/Swarm	4	7	10	13	13
Swarm Radius (km)	0.18	1.5	8.75	50	50
Functionality Study	<-- #5 -->				

Table 4. B-TOS payload functionality attributes for Pareto optimal architectures A, B, C, D, and E.

Functionality Study	5	
Spacecraft Type	Mother	Daughter
Number	1	3+
Payload (Tx)	Yes	No
Payload (Rx)	Yes	Yes
Processing	Yes	No
TDRSS Link	Yes	No
Intra-Swarm Link	Yes	Yes

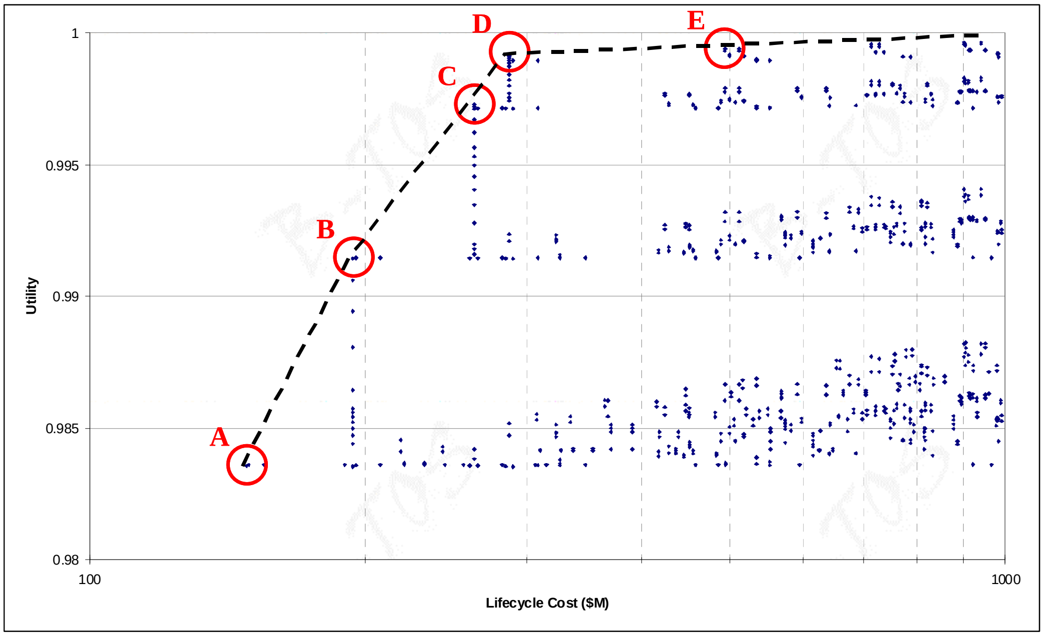


Figure 2. The B-TOS architecture tradespace plotted in cost-utility space.

In the discussions that follow, architecture candidates will frequently be represented in cost-utility space as they are in Figure 2. Each dot on Figure 2 represents a single and unique space system architecture that can accomplish the mission, with its corresponding cost and utility. Note that the representation in Fig. 2 does not take uncertainty into account. Each point represents an architecture that is assumed to deliver its nominal performance at its nominal cost. Plots like this are most useful for illustrating how easy it is to define a bad point design (i.e. one that deliver low utility at high cost) and for finding the optimal front of architectures (i.e., A,B,C,D,E). We will return to the BTOS case study to discuss policy uncertainty and its impacts, but first we introduce a second case study and an uncertainty analysis approach that explores the potential of portfolios of solutions as a means to manage uncertainty.

Broadband Case Study Example

The struggle of delivering broadband infrastructure has been the focus of a number of recent commercial endeavors, ranging in implementation concepts from wired options like cable and DSL to wireless delivery options either through ground, air or space based sources. The most successful implementations thus far have been through ground-based systems; however, there are also companies seriously exploring the capabilities a space-based platform provides. The primary benefits of a space broadband system over that of any ground based system is that space systems have less reliance on any preexisting ground infrastructure and can serve changing and/or rapidly growing markets more effectively through the repositioning of satellites and adding more capacity to the systems through increasing the complement of space assets or satellite upgrades. Locations where satellite based services have advantages over land-based systems include economically developing nations with little pre-existing infrastructure, sea based platforms and air based platforms, and remote locations that have little access to land based systems. Space based broadband systems also have the potential to compete even in markets where infrastructure is widespread and competitors already serve customers. This phenomenon can be seen in the satellite TV industry where satellite based TV broadcast customers represent a significant share of the overall market. Through competitive pricing strategies and product differentiation, DirecTV and others have proven that space based systems are viable competitors with other platforms.

This case study explores the systems analysis of such a space based broadband architecture. This commercial venture allows the demonstration of the uncertainty analysis framework in a context that includes aspects of model and market uncertainty. Numerous examples of the effects of market uncertainty can be seen on the space industry, ranging from uncertainties in launch vehicle capacity to meet the evolving needs of low earth satellite delivery to market uncertainties that defined bankruptcies in the case of Iridium and GlobalStar space systems. Where the major decision criteria for a complex system is market driven, market uncertainties should always be considered

The goal of the systems analysis is to explore the tradespace of potential architectures that satisfy a recognized need for a broadband communications infrastructure. The major feature of the architectural concept consists of a satellite network complemented by ground stations. While we have chosen to model a space system to service this market, we have not defined the details of the architecture and have left them open for defining the tradespace. Six tradable parameters in the design vector define the boundaries of the tradespace. These are altitude, inclination, satellites per plane, number of orbital planes, payload power, and the area of the phased array antenna. These characteristics and their possible values are given in Table 5.

Table 5: Design vector for the Broadband Communication Satellite System

Name	**Description**	**Potential Values**
Altitude	Altitude for a defined circular orbit	LEO(1500km), MEO(20184km), GEO(35786km)
Inclination	The inclination of the circular orbits.	0-90˚
Satellites per Plane	The number of satellites in each of the occupied planes	1-8
Number of Planes	The number of orbital planes that the satellite constellation occupies	1-10
Payload Power	Downlink power from an individual satellite	1kW-10kW
Phased Array Area	Area in square meters of the total phased array antenna area	1-5m^2

In addition to the six elements in the design vector, there are also a number of variables that are held constant for all architectures in the tradespace or computed as intermediate variables through the GINA approach.

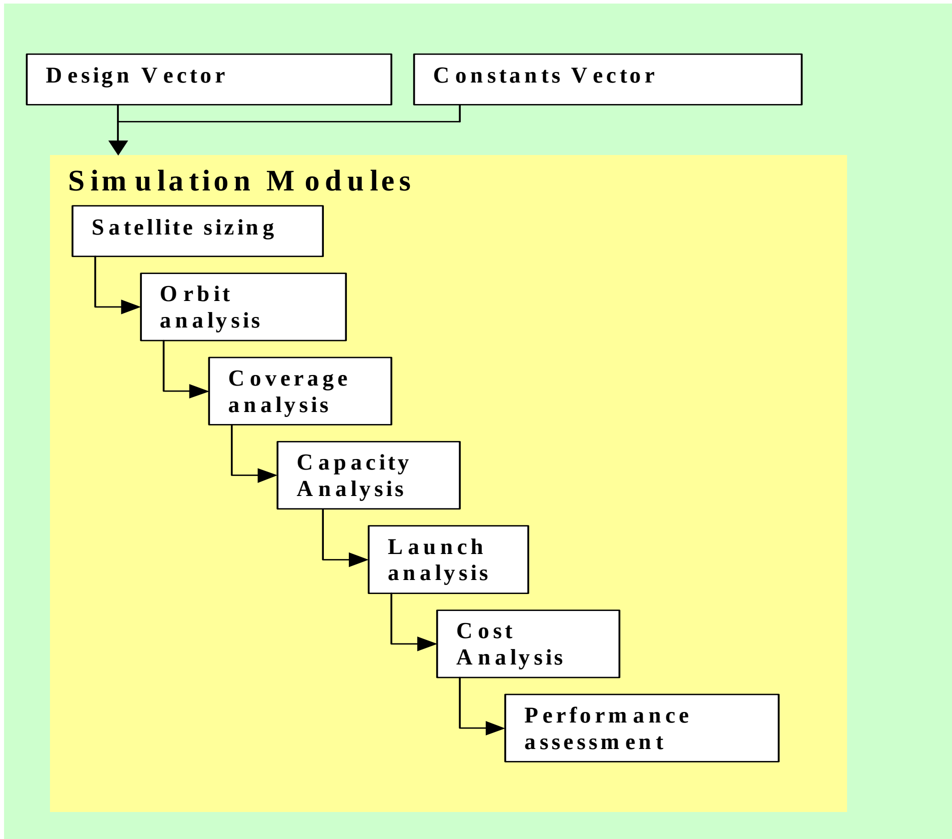


Figure 3: Systems Simulation Flow

Figure 3 describes the simulation flow that was employed in this case study, based on work by Kashitani [7]. The model is initiated with the definition of a constants vector that contains parameters of the designs that should remain constant across all of the architectures that are being evaluated. Examples of constants in the Broadband model are scientific constants, such as the earth's radius, and conversion factors. Other constants that are included in the Broadband model are market constants such as market size and distribution, satellite sizing ratios, and launch vehicle performance. Each box in Figure 3 represents an individual software module and each arrow represents the flow of inputs and outputs.

Model Results

From the broadband GINA model, thousands of architectures were evaluated. A set of Pareto optimal architectures in terms of subscriber hours and total system cost were found. In Fig. 4 only the architectures on the optimal front are shown. The architectures in the upper right of the plot are LEO based systems (such as Teledesic) while the ones in the lower left are GEO based systems such as Spaceway.

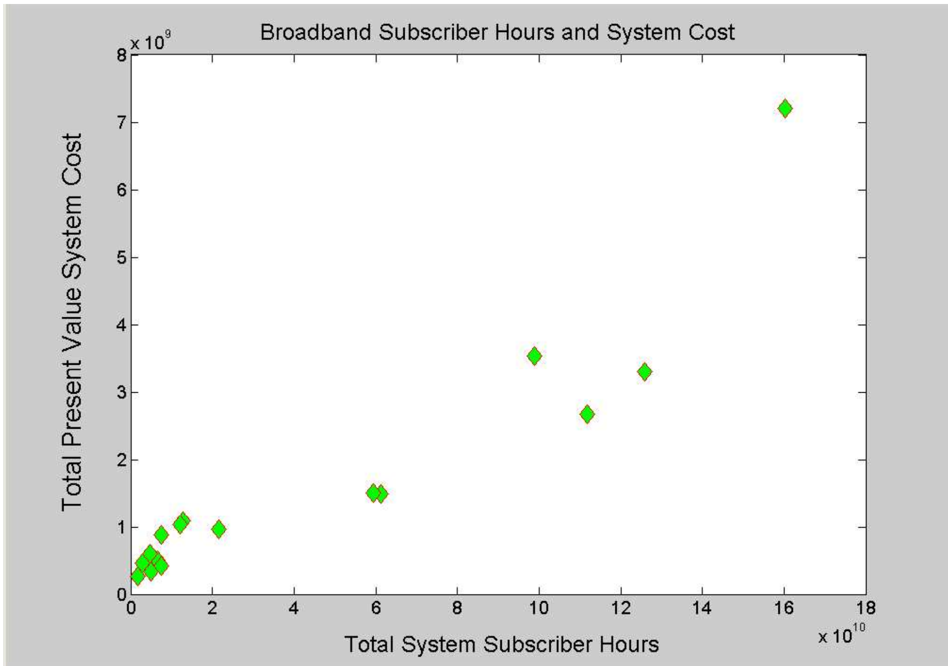


Figure 4: Commercial Broadband System Pareto Optimal Front

Market and Model Uncertainty

We have introduced two case studies. We first look at market and model uncertainties associated with the broadband case and then consider policy uncertainty associated with the mapping mission.

Table 6 presents the various sources of uncertainty that were considered in the Broadband case study. Because the Broadband GINA model is relatively coarse, a good deal of the uncertainty we are quantifying arises from the rules of thumb that are being used in the model simulation to generate results. However, because of the commercial nature of the case, market uncertainties are also introduced.

Table 6: Sources of uncertainty considered in Broadband Case

Total Market Size
Market Capture
Payload Power per Unit Mass
Mass Fraction of the Payload with respect to Dry Mass
Fraction of Dry Mass in Wetmass
Density of Satellite
Discount Rate
Theoretical First Unit Cost per Kilogram

The broadband system analysis affords the opportunity to introduce market uncertainty into application. Specifically this market uncertainty is arising from the estimation of three main parameters: 1.) total market size of broadband customers, 2.) percent market capture for this project, and 3.) the discount rate used in the cash flow analysis. These three sources serve as representative examples of market uncertainty. Others could have been included such as uncertainty in market geographic distribution or competition scenarios.

Uncertainty in total market size is modeled using a lognormal distribution that is consistent with previous market analysis of the broadband market potential. A lognormal distribution is used for the obvious reasons that the market has a lower bound of zero, but a less constrained upper bound. Figure 5 represents the market distribution that was used in the analysis. Percentage of market capture is also modeled as a lognormal distribution while discount rate is modeled as a normal distribution around 30%.

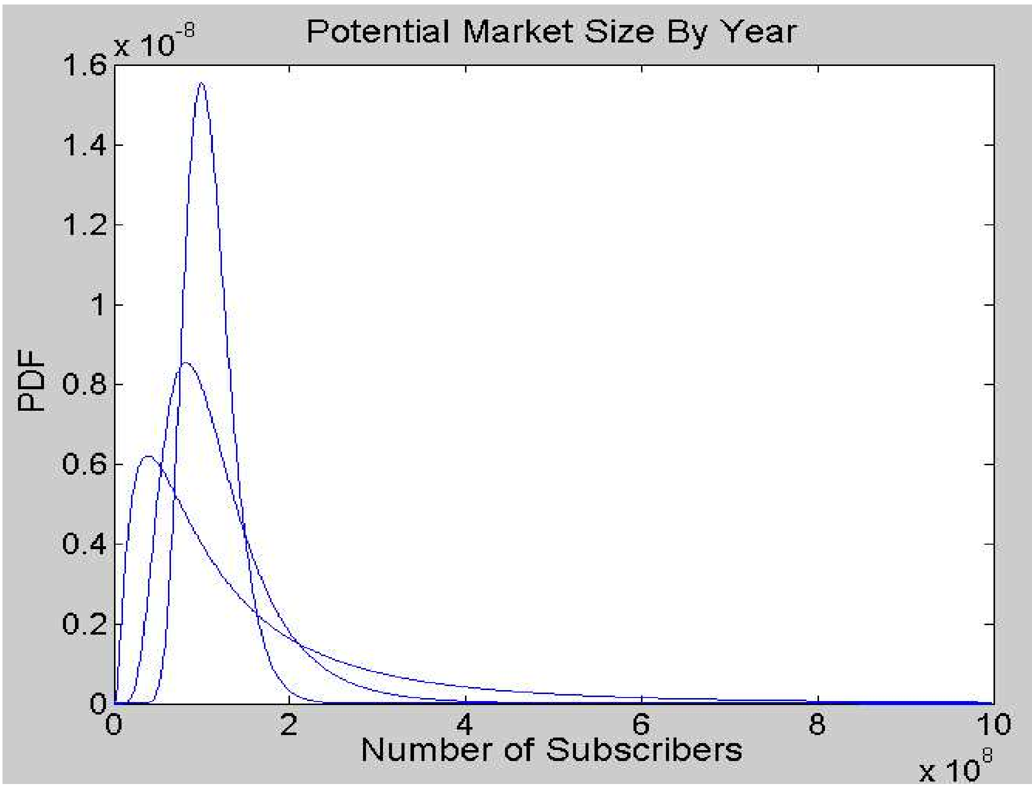


Figure 5: Uncertainty in potential market size per year

Although we describe market uncertainties in the Broadband case, by no means are market uncertainties isolated to commercial ventures. Military and civil systems also suffer from market uncertainties in a number of ways, ranging from competition to demand for the system, as we mentioned in the introduction in the case of the military communications architecture.

After the uncertainty analysis is complete, the evolved tradespace with the inclusion of uncertainty can be represented by Figure 6. Notice that there is an explicit visualization of the uncertainty surrounding the individual Pareto optimal architectures through the use of ellipses to represent the standard deviation in distribution of expectations.

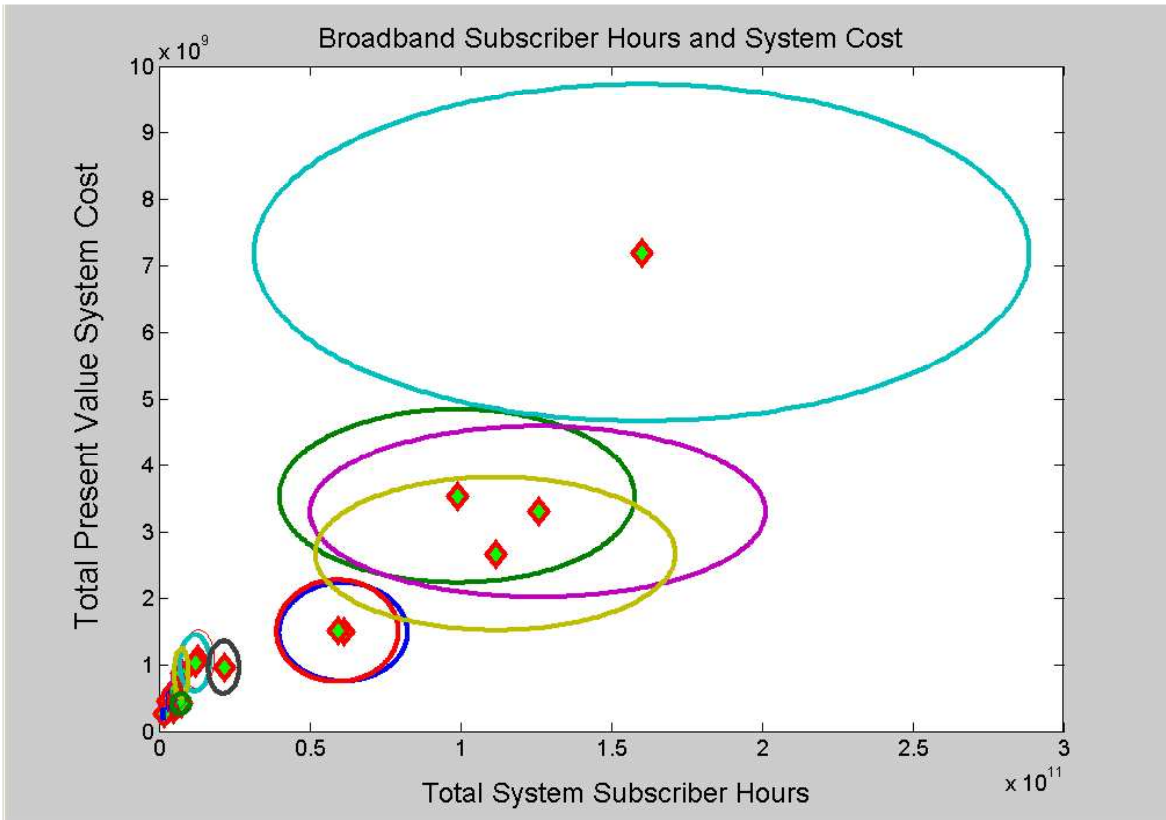


Figure 6: Broadband tradespace with the inclusion of uncertainty

From this inclusion of uncertainty, we can conclude that the highest return architectures have the highest uncertainty in the service they can support. Furthermore many of the Pareto optimal architectures have large overlap in uncertainty making them hard to distinguish in terms of overall value. Therefore, it is not straight forward that uncertainty information alone can provide the decision maker any clear strategy without some way to codify the uncertainty information and form trade-offs. The method we introduce to accomplish this task is portfolio theory and portfolio optimization. Having roots in economics and finance, portfolio theory has evolved to change mental models in investing in liquid assets and we believe it has the potential to change mental models in investment strategies in conceptual design as well.

Portfolio Theory Applied to Space Systems

Based around the central premise of maximizing return subject to a given level of risk aversion, portfolio theory has evolved since its introduction by Markowitz to a central paradigm in investment [8]. We briefly explain and apply portfolio theory to the broadband commercial space systems previously described to point out the potential that such an approach can have on the design of space systems.

Once the outcome distributions for individual architectures have been captured, portfolio theory can be directly applied to the problem of identifying an optimal investment strategy for a decision maker to pursue. [By investment strategy here we imply investment of resources (time, money, etc.) in architectural designs to develop]. Equation 1 presents the portfolio optimization algorithm, where r represent the expected value from the architecture, k represent the risk aversion coefficient, Q represents the covariance matrix and w is the vector that contains the relative investment of each architecture in the tradespace.

$$
\begin{aligned}
&\max r^T w - \frac{k}{2} w^T Q w \\
&S.T. \sum_{i=1}^{n} w_i = 1 \qquad (1) \\
&S.T. w \geq 0
\end{aligned}
$$

Equation 1: Portfolio Optimization Algorithm

The power of portfolio optimization is the ability to allow trade-offs of portfolios along the line of value and uncertainty, subject to a given decision makers level of aversion to risk. Further, the approach incorporates the idea of uncorrelated behavior under uncertain conditions which in the end allow for diversification of investment that can result in higher returns for a given level of uncertainty than would otherwise be possible.

Figure 7 represents the value/uncertainty tradespace for the broadband communications space system. That is, the value of the system is plotted against the uncertainty associated with it. In the broadband case, we have chosen subscriber hour/$ as the overall value criteria, while uncertainty is the standard deviation around the expected subscriber hour/$. The tradespace forms an efficient frontier denoted by the concave line. Along this line are portfolio mixes whose return cannot be exceeded without accepting a higher degree of uncertainty. This tradespace and the program created to navigate the tradespace allow the decision maker to investigate the relative position of a portfolio along the efficient frontier (on the left graph), but also the composition of the portfolio (as shown by the table on the right). We believe this type of analysis provides both a visual understanding of uncertainty in the tradespace, but also an approach to effectively manage it.

In this case, the LEO based systems dominate the upper right segment of the function per cost and uncertainty tradespace, implying higher returns and higher uncertainty than MEO or GEO based missions. Thus Teledesic (& the narrowband Iridium) initially chose LEO. However the uncertainty, and in this case risk, associated with these systems is so high that the use of GEO based designs, while returning lower performance, can be a better choice for a risk averse decision maker. This is the path that most other commercial systems have chosen to take. Even the commercial systems that initially went with more risk have been forced by the market realties of raising money to focus on the MEO and GEO systems. More generally, this case illustrates that the best way to consider the design may be to consider portfolios of architectures and carry balanced diverse sets of designs as long as the possibility exists to diversify uncertainty.

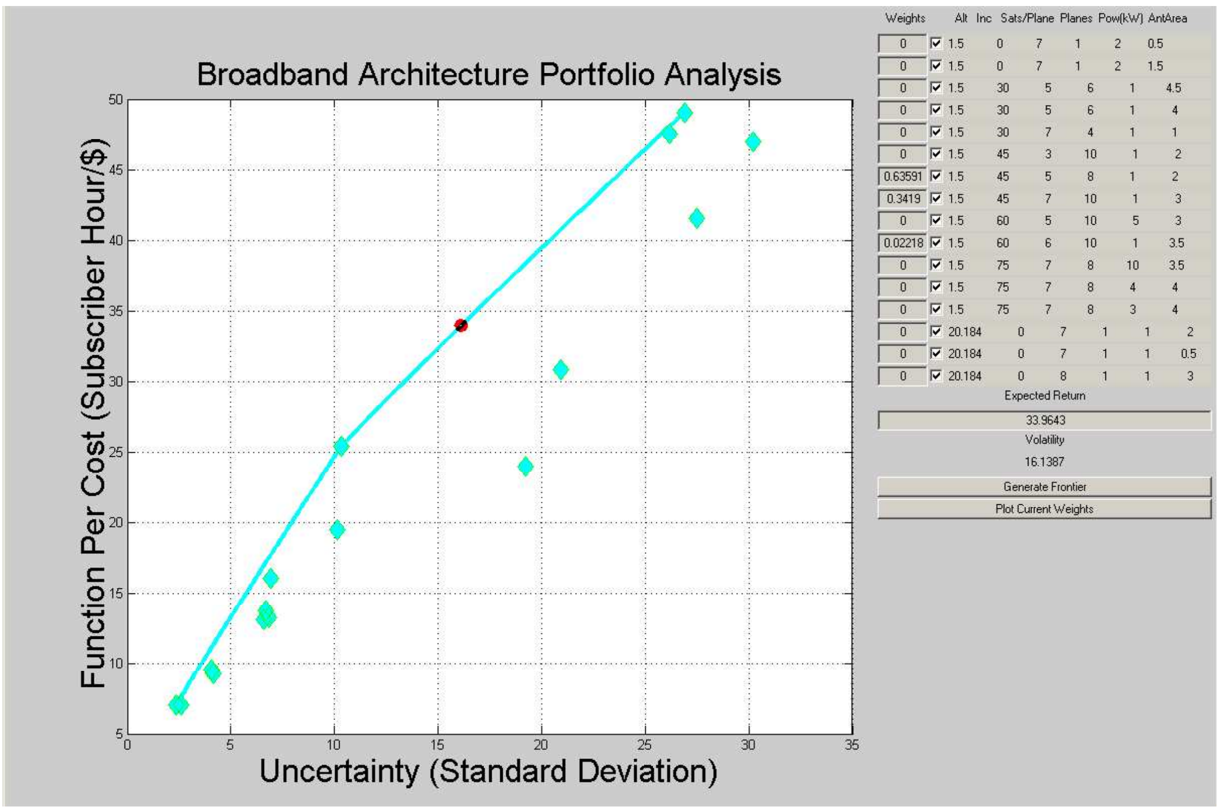


Figure 7: Portfolio Tradespace of Broadband Communications System

Policy Uncertainty

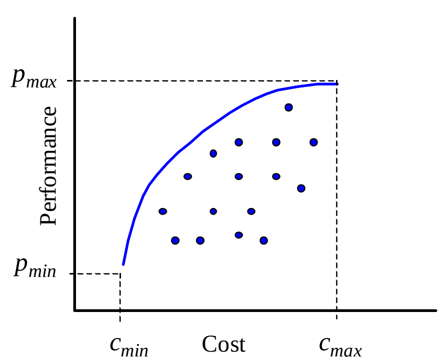
In the last section, we analyzed market/model uncertainties. In this section, we consider a different class of uncertainties namely policy uncertainties but model them with the same conceptual approach.

Budget changes are the most frequently reported action taken by policy makers, and it is not hard to understand why. Government and military space system programs are subject to budget approvals each year by their own agencies as well as the Congress. Each year, a program's budget can – and frequently does – change. For fiscal years 1996-1998, 32% of defense programs experienced a budget reduction, 53% experience a budget increase and only 15%

received the budget they requested. Hence a space architect can conclude that the probability that the budget will be changed is much larger than the probability that it will stay on the nominal plan. Thus, budget uncertainty is perhaps the most pervasive policy uncertainty facing government and military systems today. This is due to the nature of the competing stakeholder demands that are balanced and adjudicated through the budget resolution process each year. Since these are constantly shifting, the budget for any given system will be subject to this kind of uncertainty. Thus this uncertainty is irreducible based on human behavior and the nature of the budget process. Unlike technical uncertainty, which can often be reduced this kind is unlikely to change except in dire emergencies.

Potential effects of budget uncertainty can be examined during conceptual design of space system architectures, and these effects can better inform architecture choices that will be robust to budget changes. After space system architectures are modeled with GINA and MATE techniques, the effects of budget changes can be explored by varying the yearly budget allocation for the program. When the yearly budget level is pushed below the nominal yearly budget level for a given architecture, that architecture's schedule will need to be extended to accommodate the lower yearly budget level. Using historical data relating program schedule extension data to resulting program cost increases, the extra cost of extending the program schedule can be calculated. This is the cost a program bears if it should fall subject to a budget reduction.

A government space system program manager contemplating which of many Pareto optimal architectures to choose should be aware of how each of those Pareto architectures behaves relative to varying yearly program budget levels. Program budget increases or appropriations equal to requests will not in theory inhibit a program from meeting its original cost and schedule goals. However, budget reductions are likely to do this. If an architecture is initially selected, and later the Congress chooses to reduce the budget available each year for that program, the program will likely incur a cost overrun as well as a schedule slippage according to historical data. These effects in turn draw increased Congressional oversight to a program, increasing its chances for being cancelled altogether in future years. This is one vivid example of why a program manager should be concerned about budget uncertainty on his or her program, and in particular, about downside risks.



Critical stage transition point variables:

- p_{max} and p_{min} defines the range of acceptable performance for the program
- c_{max} represents the cost of the architecture on the pareto front with p_{max}
- c_{min} represents the cost of the architecture on the pareto front with p_{min}

Figure 8. Parameters associated with Pareto front set of architectures in cost and performance space.

We can describe in generalized terms how a budget reduction affects a program's total budget, and identify three distinct stages of the behavior of the Pareto optimal front of a set of architecture candidates. First, we identify the architectures that lie along the Pareto front in a given architecture cost–performance tradespace, and call this set of architectures i. In this set, there is a minimum acceptable performance (p_{min}) and its associated cost (c_{min}), and there is a maximum achievable performance (p_{max}) and its associated cost (c_{max}). These will tend to be at the extremes of the Pareto frontier, as illustrated in Figure 8. Each architecture in i will have some nominal program duration in years, d_i.

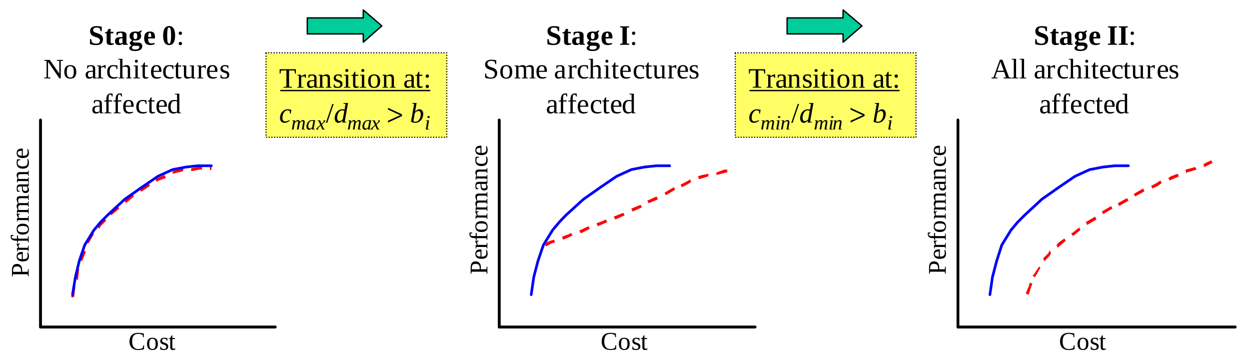


Figure 9. General behavior of the Pareto optimal front of a set of system architectures being affected by budget reductions. The solid line represents the nominal case and the dashed line represents the reduced budget case.

If we overlay the budget reduction cost–performance tradespace onto the nominal cost–performance tradespace, we see that they are identical for a yearly budget level, b_i, greater than c_{max} / d_{max}, and we call this behavior Stage 0. This is shown graphically in Figure 9. For a yearly budget level less than c_{max} / d_{max} but greater than c_{min} / d_{min}, the Pareto fronts of the budget reduction case and the nominal case will appear to diverge at some points when they are overlaid in the cost–performance tradespace. This separation of the Pareto fronts indicates that some architectures are being affected by the reduced budget level while others are not. We will call this behavior Stage I, and can identify the critical yearly budget level at which the Pareto front of architectures transitions into this Stage as c_{max} / d_{max}. Finally, for a yearly budget level less than c_{min} / d_{min} , the Pareto fronts of the budget reduction case and the nominal case completely separate, indicating that no architecture in the Pareto optimal set is unaffected by a reduced budget level. We will call this behavior Stage I, and can identify the critical yearly budget value at which the Pareto front of architectures transitions into Stage II as c_{min} / d_{min}.

These Stages identified above are guidelines for a program manager. A range of likely yearly program budgets can be examined to determine if they put the program architecture Pareto front into Stage 0, I, or II. If a program falls in Stage 0 for all likely yearly budgets, then a program manager can feel more confident in selecting any architecture candidate since they are all robust to the anticipated range of budget uncertainty. If a program falls in Stage I, then a program manager may want to consider robustness to budget uncertainty as a criteria in selecting a final system architecture, since cost and schedule overruns that can result from budget uncertainty will increase oversight and the probability of cancellation later on. Lastly, if a program falls in Stage II, then none of the architecture candidates are robust to budget uncertainty, and a program manager will want to take some action to avoid potential repercussions in the program's future. These might include (but are not limited to) generating new budget-robust architecture candidates to choose from, seeking increased protection from budget uncertainty from agency directors, or making a stronger case to appropriators for a larger budget.

Case study application: B-TOS

The budget uncertainty analysis was applied to the B-TOS mission. The performance and cost attributes of the Pareto optimal set of five architectures is shown in Figure 10, along with the critical values that describe the transition points between Stages 0, I, and II. The five Pareto optimal architectures in cost–performance tradespace are shown in Figure 2, and are labeled A, B, C, D, and E. Figure 10 (a) – (d) graphically shows the overlay of the nominal and the budget reduction cases for decreasing levels of yearly program budgets.

We find that the Stage 0 to Stage I transition occurs at $80.3M/yr, and the Stage I to Stage II transition occurs at $18.6M/yr. Since B-TOS is a small military mission, the likelihood of its initially receiving or sustaining an $80M/yr budget is probably very small (typically such missions stay under $20-50M/yr). Thus, a program manager concerned with budget uncertainty may not want to choose architecture candidate E, even though it is the highest performing architecture, because it is adversely affected by yearly budgets below $80M/yr. If the program manager foresees that a yearly budget of perhaps $25M/yr is far more likely for a small military mission, then architecture candidate A would be a good choice, because it is robust to budget uncertainty. But if the yearly budgets for small military missions are more likely to be in the $5M - $10M /yr range, then none of these Pareto front architecture candidates for B-TOS would be robust to budget uncertainty. So what is the B-TOS program manager to do now? Unfortunately, a small military mission is unlikely to warrant the agency director making a special effort to fence this one particular program's budget, and it is also unlikely to attract much support from appropriators who have larger concerns. If these are the case, the program manager and the program team may then wish to return to the drawing board and come up with less expensive architecture alternatives, or perhaps descope the program's requirements in an effort to create a budget robust architecture choice.

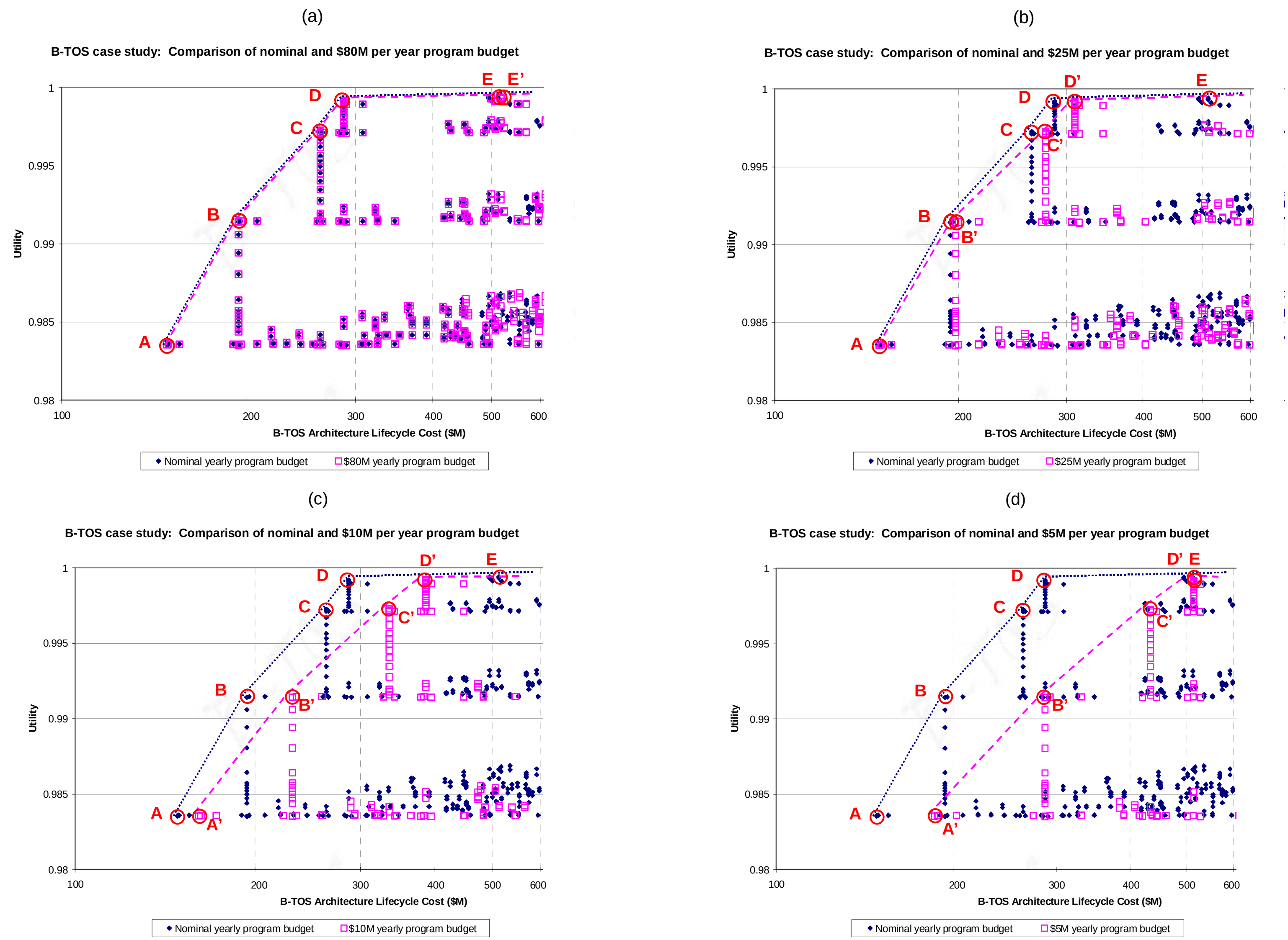


Figure 10. B-TOS case study results for nominal and budget reduction cases for decreasing yearly budgets.

Conclusions

In this paper, we have examined the effect of uncertainty on several classes of space system architectures. These are used as models of engineering systems. We have shown that there are many types of uncertainty. These range from technical uncertainty (does the propulsion system deliver the specified thrust with the specified efficiency) to market based uncertainty (what will be the response of people in London to the space based delivery of T1 links compared to procuring the service through fiber) to policy uncertainty (will the Congress keep funding the development of this system at the same level as last year?).

When these uncertainties are taken into account in the design of these complex space system architectures a number of interesting conclusions follow. The first is that the design points may be quite different with uncertainty incorporated from the beginning as compared to not incorporating the uncertainty. This is clearly seen in the broadband case where the LEO based systems have clear performance advantages over all other systems. Thus Teledesic (& the narrowband Iridium) initially chose LEO. However the uncertainty and in this case risk associated with these systems is so high that the use of GEO based designs, while returning lower performance, is the path that most commercial systems have chosen. More generally, this case illustrates that the best way to consider the design may be to consider portfolios of architectures and carry balanced diverse sets of designs as long as possible.

The second conclusion is that some of the uncertainty is caused by human behavior that is endemic to the nature of the way that stakeholders balance their interests associated with these systems. This was seen in the cost capping analysis of the ionospheric mapper. The likelihood that there will be budget changes in the design and construction of these systems is much larger than the likelihood that they will get exactly what they request. This is due to the dynamic nature of the political process by which decisions are made and policies are decided. Given that it flows from the nature of human behavior, it is a kind of irreducible uncertainty (unlike many types of technical uncertainty). We showed that it was possible to consider this kind of irreducible uncertainty in the design of the system and actually make choices knowing one is subject to this uncertainty.

A third observation flows from consideration of how some of the commercial and military space system architectures have been used in practice. GP S was originally designed for guiding long-range nuclear bombers to their targets (which accounts for the very low power signals). DSP was originally designed for finding strategic ballistic missile launches and relaying information on those launches to the National Command Authority in Washington. The primary civilian use of GPS is now helping hikers not get lost & providing timing signals to cell phone networks while the primary military use is in close air support. DSP is now used primarily to find short-range tactical ballistic missiles and relay those results to forces in theater. Both of these substantially different uses arise from the fact that the original architectures had enough uncertainty in their use (a kind of flexibility) that the interaction with creative humans led to new ways of thinking about and using the systems. This indicates that uncertainty is not a synonym for risk. If the architecture of DSP had been so tightly specified that it could not be used in any other way than finding strategic ballistic missiles then it would have fulfilled it's original mission and be unable to fulfill the subsequent missions (which where not envisioned when it was first flown). In a similar manner, the development of the large commercial market associated with GPS was a complete surprise to the original designers of the architecture but the architecture was robust enough to accommodate this kind of use.

We now generalize to engineering systems and try to draw some analogous conclusions. Thus the three conclusions about uncertainty and engineering systems that flow from this analysis are

1) Engineering Systems must be designed with uncertainty as one of the central organizing principles.
2) Since Engineering Systems have management and social dimensions and thus involve human interactions, there is an irreducible uncertainty associated with these dimensions that will affect the design of the system.
3) Uncertainty in use may allow the engineering system to satisfy quite different missions from the original one. Thus, uncertainty and risk may not be correlated; indeed it is humans interacting with the uncertainty that allows the flexibility to be creatively used.

Acknowledgements

The authors would like to acknowledge support from the Department of Defense in undertaking this research. Useful discussions were had with Profs. Ed Crawley, Earll Murman, Sheila Widnall and Tom Allen and Dr. Joyce Warmkessel.

References

1. Air Force Space Command Online Library, http://www.spacecom.af.mil/hqafspc/library/default.htm

2. Shaw, G., Miller, D., and Hastings, D. E., "Generalized Characteristics of Communication, Sensing and Navigation Satellite Systems", Journal of Spacecraft and Rockets, Vol 37, No. 6, pp801-811, November-December 2000

3. Shaw, G., Miller, D., and Hastings, D. E., "Development of the Quantitative Generalized Information Network Analysis (GINA) Methodology for Satellite Systems", Journal of Spacecraft and Rockets, Vol 38, No. 2, pp257-269, March 2001

4. Jillla, C. D., Miller, D and Sedwick, R. J., "Application of Multidisciplinary Design Optimization Techniques to Distributed Satellite Systems", ", Journal of Spacecraft and Rockets, Vol 37, No. 4, pp481-490, August 2000

5. Keeney, R. a. H. R. (1976). Decisions with Multiple Objectives. New York, Wiley and Sons.

6. Diller, N., Qi Dong, Carole Joppin, and S. K. Sandra Jo Kassin-Deardorff, Dan Kirk, Michelle McVey, Brian Peck, Adam Ross, Brandon Wood (2001). B-TOS Architecture Study:Second Iteration of the Terrestrial Observer Swarm Architecture. Cambridge, MA, Massachusetts Institute of Technology.

7. Kashitani, T. (2002). Development and Application of an Analysis Methodology for Satellite Broadband Network Architectures. Proceedings of the 20th AIAA International Communications Satellite Systems Conference, Montreal, Canada.

8. Markowitz, H. (1991). Portfolio Selection: Efficient Diversification of Investments. Cambridge, MA, Basil Blackwell.