REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE 8 Sep 97	3. REPORT TYPE AND DATES COVERED

4. TITLE AND SUBTITLE
Space Based Radar-System Architecture Design and Optimization for a Space Based Replacement to AWACS

5. FUNDING NUMBERS

6. AUTHOR(S)
Douglas P. Wickert

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Massachusetts Institute Of Technology

8. PERFORMING ORGANIZATION REPORT NUMBER

97-020

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
THE DEPARTMENT OF THE AIR FORCE
AFIT/CIA
2950 P STREET
WPAFB OH 45433

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

14. SUBJECT TERMS

15. NUMBER OF PAGES
167

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

DTIC QUALITY INSPECTED 2

Space Based Radar – System Architecture Design and Optimization for a Space Based Replacement to AWACS

Douglas P Wickert

May 1997

Space Based Radar – System Architecture Design and Optimization for a Space Based Replacement to AWACS

by

Douglas P Wickert

B.S. Astronautical Engineering
B.S. Mathematical Sciences
US Air Force Academy, 1995

Submitted to the Department of Aeronautics and Astronautics
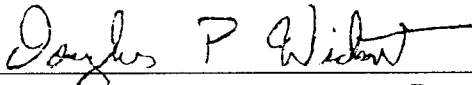in partial fulfillment of the requirements for the degree of

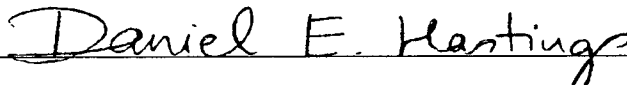Master of Science in Aeronautics and Astronautics
at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1997

© 1997 Douglas P Wickert. All rights reserved.

The author hereby grants to MIT permission to reproduce and distribute publicly
paper and electronic copies of this thesis document in whole or in part.

Signature of Author : _____
Department of Aeronautics and Astronautics
9 May 1997

Certified by : _____
Daniel E Hastings
Professor of Aeronautics and Astronautics
Thesis Supervisor

Accepted by: _____
Professor Jaime Peraire
Professor of Aeronautics and Astronautics
Chair, Graduate Office

Space Based Radar – System Architecture Design and Optimization for a Space Based Replacement to AWACS

by

Douglas P Wickert

ABSTRACT

Through a process of system architecture design, system cost modeling, and system architecture optimization, we assess the feasibility of performing the next generation Airborne Warning and Control System (AWACS) mission from a Space Based Radar platform. Initial studies by the Air Force called for systems that push the limits of currently available technology and are prohibitively expensive. We introduce a distributed operations concept for SBR that reduces the size of the satellites required, increases system reliability, improves system performance, and reduces system cost. Coupled with the system architecture optimization process, the resulting distributed SBR concept is feasible with currently available technology and is estimated to cost less than half of other, similar concepts.

The system architecture design process minimizes a system cost function with respect to system architecture independent variables. The system cost model consists of subsystem, reliability, operations, and constellation components. The architecture optimization process is a powerful analysis and concept development tool and could be adapted to other systems. A reliability model for distributed systems which quantifies the cost for reliability is also developed and should be useful in the analysis of other distributed systems.

Table of Contents

Table of Figures

Acknowledgments

I thank Chuck McQuiddy at Lockheed Martin for initially suggesting the research, reviewing our work, sharing their ideas for SBR, and ultimately for funding our research. Some of the most useful feedback for improving the accuracy and credibility of our work come from dialogues between Chuck McQuiddy and Daniel Hastings.

I also thank the Air Force for allowing me the opportunity to spend two years in pursuing this degree. I hope that this research in some way compensates their investment. I thank MIT for the best possible educational experience and I thank the National Science Foundation for the fellowship that made all of this possible.

Finally, I thank my parents for the years of love, nurture, support, and encouragement that allowed me to develop into who I am today. And I thank my beautiful bride, Jodee, for proofing the drafts, for relentless encouragement, for her ceaseless tolerance of the long hours, and above all else, for her love throughout it all.

Biographical Notes

Douglas P Wickert graduated from Beavercreek High School in Dayton, OH, in May, 1991. He entered the US Air Force Academy in June of 1991 where he received military distinction as Group Outstanding Fourthclass Cadet of the year, was a Soaring instructor pilot, was active in intramural athletics, and held various squadron leadership positions. Wickert graduated from the Air Force Academy first in his class academically in May, 1995, with a double major in Astronautical Engineering and Mathematics. He received honors as the top graduate in Astronautical Engineering, the top Engineering graduate, and the top academic graduate. Wickert entered the Massachusetts Institute of Technology in the September of 1995 on a National Science Foundation Fellowship. He is a student member of the American Institute of Aeronautics and Astronautics (AIAA), a student member of the American Association for the Advancement of Science (AAAS), a member of the Pi Mu Epsilon honor society for mathematics, and a member the Sigma Xi honor society for pure and applied science and engineering research. Following completion of an S.M. degree from MIT in June of 1997, Wickert is being reassigned to the 8th Flying Training Squadron, Vance AFB, OK, for pilot training.

10

Victory smiles upon those who anticipate the changes in the character of war,
not upon those who wait to adapt themselves after the changes occur.

Italian Air Marshall Giulio Douhet, 1928

1. Introduction

The whole of military tactics and strategy can almost be completely reduced to a single doctrine: *take and hold the high ground.* In warfare up to the twentieth century, the high ground gave soldiers the ability to survey what the enemy was doing, was easier to defend, and gave the advantage of attacking and counter-attacking downhill. With the advent of airpower the realm of high 'ground' moved into the air, but the role and importance of maintaining the high ground only grew: airpower gave improved reconnaissance capability, the ability to strike deeply - both tactically and strategically - into enemy territory, and the freedom to move on the ground at will. The lessons of the latter days of World War II and the War in the Persian Gulf are clear: air superiority is the high ground and virtually nothing can be done without it.

Space is the ultimate high ground. From space, one can project global power, do global surveillance, and have instantaneous global infrastructure. Surveillance from space has long been a tool for military planners and the update/revisit rate has steadily increased. Eventually, space surveillance will have to be real time and include the capabilities of air surveillance, ground surveillance, naval surveillance, tactical imaging, electronic intelligence, and ballistic and cruise missile warning. Thus, many of the tasks currently done by other assets will move into space as a course of technological evolution.

The time is right to seriously consider when and how some of these tasks should be moved into space. The existing Airborne Warning and Control System (AWACS) has a remaining lifetime of only 20 to 25 years. Given the current acquisition cycle, it would take about 15 years before a replacement system could be operational. Thus, in the next five years, a decision must be made regarding the platform for the next AWACS system. In 1995, the Air Force Chief of Staff, General Ronald R. Fogleman, suggested that the next generation AWACS mission should be done from space. A space based radar (SBR) for AWACS is attractive for many reasons, but it is not without technological and funding challenges.

The current AWACS platform is the E-3 Sentry [1]. The E-3 is a modified Boeing 707 airframe with a 30 foot diameter radar dome positioned above the fuselage. The E-3 provides mobile, all-weather surveillance, command, control and communications (C^3) to both battlefield command centers and pilots. The E-3 can detect and track enemy aircraft and ships, update the location and status of friendly aircraft and ships, support air to ground operations, direct interception of hostile aircraft, and perform the identification friend or foe (IFF) function. The drawbacks of the current E-3 AWACS platform include the time required to respond to a new threat associated with moving assets to the theatre, the immense support structure that must also be deployed, technology that is beginning to become outdated, and high year-to-year operating

11

and maintenance costs. It is in each of these areas that a space based surveillance platform could improve the capability of AWACS. A space based platform could respond to new threats almost instantaneously, requires relatively little support structure, would incorporate the newest technology, and is dominated by acquisition and deployment costs rather than by operation and maintenance costs. The challenges to a space based AWACS platform are a moderate to high level of technological risk and a high cost to initial operating capability.

When considering the next generation of military surveillance from space, it makes sense to initially examine the AWACS mission. The AWACS mission has the most stringent update rate requirements of all the surveillance missions. Except for differences in signal processing, an SBR system capable of performing the AWACS mission could also perform the ground surveillance mission (JSTARS), the naval surveillance mission, moderate resolution tactical imaging, provide a ballistic and cruise missile surveillance fence, and perhaps even perform the electronic intelligence (ELINT) mission. Because a replacement to the current E-3 platform is needed and because it could perform adjunct missions, we have focused on the AWACS segment of an SBR platform.

A space based system for air surveillance may have important non-military applications as well. Much effort is currently going into redesigning the Air Traffic Control system (ATC) and a space based radar for air surveillance could enable the implementation of some concepts, e.g., free-flight. Critics of relying on a primarily military system for ATC during wartime need not worry since radars would be unused when not in view of the theatre. (Airline operations in the battlefield region are unlikely, afterall.) Civilian mission adjuncts such as ATC may help justify the high cost of an SBR platform for air surveillance.

This is a detailed systems analysis of a space based replacement to AWACS. We are specifically interested in the feasibility of such a system as measured by cost. Initial, "quick-look" studies by the Air Force Space and Missile Center (SMC) generated SBR concepts with price tags of $20 to $50 billion! Funding approval for such an expensive system is highly unlikely in the political reality of the post Cold War Department of Defense. As a first step, we wanted to determine the optimality relations that may result in a reduced system cost: power, aperture, constellation, coverage, technology, sensing, and operations. We also wanted to look for new system concepts that would perform the same mission but at a reduced cost. Another goal was to identify major technology drivers that the Air Force and industry should push to study and develop that would result in a more affordable system. Section 1.1 gives more detail of the research objectives and Section 1.2 outlines the system architecture design and optimization process for SBR.

1.1 Research Objectives

Although the primary research objective is to assess the feasibility of a space based radar (SBR) system, there are several secondary objectives that are also satisfied by the system architecture design and optimization process developed here. The secondary objectives support

the primary objective but are also worthy questions by their own right. The general analysis framework of the system architecture design process is powerful and flexible enough to answer both the primary and secondary research objectives:

1. Determine the feasibility of performing the next generation AWACS surveillance mission from a space based radar (SBR) platform. Feasibility is measured by the size of the satellites and constellation required to perform the mission and the IOC cost.

2. Develop cost and performance metrics that quantify the SBR system cost as a function of system architecture.

3. Optimize the SBR system architecture with respect to the cost metric for a given level of performance.

4. Examine and quantify the effects of performance/function distribution.

5. Develop new operations concepts that may improve the feasibility of SBR.

6. Identify key technologies for SBR.

Each of the secondary research objectives in some way answers an aspect of the primary objective. In a similar manner, some secondary objectives support other secondary objectives. For example, to determine the feasibility of an SBR system as the next generation AWACS, we first wanted to optimize the system architecture. This required some method of identifying and quantifying the optimality relations in system architecture options which in turn necessitated the development of a cost function that reflects system architecture in the independent variables. New concepts of operations might also improve the feasibility of SBR. We develop a distributed operations concept that benefits from many of the advantages of distributed systems outlined in Section 1.3. Finally, identifying key technologies for SBR now might lead to developments in these areas that improve the feasibility of SBR. Thus, each of the objectives listed above are neatly addressed by the process of system architecture design and optimization.

1.2 System Architecture Design and Optimization

The system architecture design and optimization process is the analysis tool that satisfies each of the research objectives. A system cost model is developed down to the subsystem level and models for other systems engineering concerns are incorporated into the cost model. For example, a reliability model, a constellation coverage model, and an operations model that reflects system performance are developed and integrated. Additionally, the cost model is designed so that the most significant architecture options are independent variables of the cost metric. Then by analytic, semi-analytic, and numerical solutions, the cost function is minimized as a function of the desired architecture variables. This process gives the optimal system architecture that meets given system requirements and performance specifications.

The system cost model includes many submodels that are based on a fairly detailed system

analysis. These submodels may be included, omitted, or altered as desired to study the effects of different architectures, operations, performance requirements, or systems technology. It is this flexible aspect of the cost model that makes it particularly powerful. The model components are:

1. Satellite subsystems cost model. (developed in Chapter 3) Subsystems and effects included in the model:
   - Power subsystem cost model
   - Aperture subsystem cost model
   - Bus subsystem cost model, including radiation hardening model
   - Launch subsystem cost model
   - Production learning curve model

2. Reliability model for distributed systems. Cost for reliability model. (developed in Chapter 4)

3. Distributed operations concept for detection mission. (developed in Chapter 5) This is a noncoherent concept of operations which means that it is not multistatic radar. The mean time to detection performance metric is also developed and introduced as an improved metric for measuring SBR performance.

4. Constellation design for SBR systems. (developed in Chapter 6) The constellation design and analysis process gives several important results for sizing the SBR system:
   - Constellations for continuous global coverage
   - Constellation size as a function of orbital altitude
   - Mean and minimum expected satellite coverage; coverage as a function of distribution
   - Mean and maximum expected slant range to target
   - Satellite duty cycle

5. Clutter performance for distributed and non-distributed system architectures. (developed in Chapter 7) Although clutter is not directly used in the cost model, the clutter performance analysis is important in checking the performance of the optimal system design.

All of the model components are integrated into a single cost function that reflects the cost to initial operating capability (IOC cost). Operating costs are not included in the cost model for several reasons. First, the range of the system architecture options considered have similar operating costs. Second, increasing the cost function by including operating costs might obscure cost gradients due to architecture variables desired for the architecture optimization process. Third, newer space systems are dominated more by their acquisition and deployment cost than by operating costs. Our research objective is to assess the feasibility of a space based AWACS platform, not to recommend whether the next generation AWACS mission should be performed from a space-based or atmosphere-based platform. Whereas a space platform will tend to be dominated by acquisition and deployment costs, an atmosphere platform will tend to be

14

dominated by the year-to-year operation and maintenance costs. Thus a fair comparison between these two options can only be done on the basis of performance and total life cycle cost. Note that the optimization of both system architectures should be done before a fair comparison can be made.

The system architecture optimization process minimizes the IOC cost function with respect to system architecture variables for a given level of performance. The system architecture optimization process gives optimum values for four architecture variables:

1. Power
2. Aperture
3. Distribution
4. Orbital altitude

These are the most important architecture variables as they define the concept and all other architecture and system values may be derived from them. Optimal power and aperture are analytically determined from an optimality relation derived in Section 3.6. Optimum distribution is semi-analytically derived - both analytic and numerical components to the solution - as a function of system performance requirements, constellation coverage, and reliability issues. The optimum orbital altitude is numerically determined.

The system cost function is parametric based which makes it a powerful tool for several additional types of analysis besides identifying the optimal system architectures. First, system cost can be examined as a function of performance specifications. Although most military systems will have minimum requirements, cost as a function of performance may significantly influence the performance specifications. Second, different concept proposals can be compared on an equal basis. Third, cost function sensitivity to several key subsystem parameters may indicate where to direct research and development resources to help reduce the system cost. Finally, the effect of technological breakthroughs on system performance or cost can be quantified by the change in system cost.

1.3 System Distribution for SBR

The feasibility of performing the next generation surveillance missions from a space based platform is significantly improved by distributing the function of detection. A distributed operations concept for radar is developed in Chapter 5 which reduces the power and aperture required on a satellite. The constellation of satellites increases in size which has several secondary benefits in coverage, design efficiency, and detection resolution. In most cases, the reliability and survivability of the distributed constellation is also improved.

System distribution for space systems is not a new idea [2,3]. In many ways, the course of events is analogous to the changes experienced in computer systems almost two decades ago. Improvements in space qualified computers, manufacturing techniques, bus and power

subsystems, and improvements and miniaturization of sensors, and improvements in communication are making distribution more affordable and realizable. In addition, several new commercial ventures are employing large constellations and extending the envelope of what was previously deemed possible. To date, the Global Positioning System (GPS) was the largest constellation consisting of 24 satellites. The Iridium system is a 66 satellite constellation that will provide world-wide cellular phone access. The Teledesic concept is a bold proposal for a broadband communication system that initially called for a 840 satellite constellation and is currently based on a 288 satellite constellation. It is very likely that the limitations to constellation size due to satellite operations will soon be solved. This will probably usher in a new era of large constellation systems and open the possibilities for distributed space systems.

There have been several previous concepts for distribution of radar, though most of them have been multistatic in nature. NASA, the Air Force Office of Scientific Research (AFOSR), and DARPA all funded investigations of Distributed Array Radar (DAR) concepts in the early 1980s [4,5,6]. Most of these concepts used multistatic sparse aperture synthesis from separate platforms to achieve high resolution with limited aperture. More recently in 1996, the Air Force Space and Missile Center (SMC) initiated a bistatic radar sensor study that examined using uninhabited aerial vehicles (UAV) as a adjunct receiver to an SBR transmitter [7]. The primary challenge of these and other multistatic radar concepts is the requirement to maintain phase coherence between radar systems. The number of required connections between satellites - and hence complexity - increases geometrically with distribution.

The distributed radar operations concept developed here is not multistatic and thus there is no requirement for phase coherence between systems. All processing between distributed satellites is post-detection processing. Distribution is achieved by distributing the function of aircraft detection from one radar to several radars. This is done by decreasing the probability of detection for a single radar. The overall probability of detection for all the distributed satellites however remains high. Decreasing the probability of detection on a single satellite results in a decrease in the power and aperture required on that satellite. If some level of performance degradation is allowed for the distributed system, then the reliability of the distributed system is higher than a system of only one radar. The level of performance degradation and the reliability cost savings can be quantified by the distributed reliability model developed in Chapter 4. Distribution also improves constellation coverage and decreases the variability of target slant range which results in a more efficient design. The distributed radar system also has the advantage of an increased number of perspectives of the target area which results in improved target detection and clutter rejection capability. Different perspectives of the same target also improves the spatial and frequency resolution of the target.

Distribution of the space-based AWACS mission results in a decrease in cost to initial operating capability of over one third compared to a nondistributed system.

16

2. Radar Fundamentals

In this chapter, we consider the components of a general radar system and the most common ways for measuring the performance of radar. The goal is to establish the fundamentals of radar as they will be used extensively in analyzing a Space Based Radar system. Section 2.1 gives the elements of a typical radar system and discusses several different types of radar. Radar is ultimately an application of signal processing - the extraction of information from signals - so the radar transmitted and received signals are characterized in Section 2.2. Resolution is an important performance metric and the resolution due to waveform is derived in Section 2.3. The radar range equation is derived in Section 2.4. Important performance limits for a space based application - multipath, scintillation, doppler spreading, and range spreading - are detailed in Section 2.5.

2.1 Radar System

There are several types of radar systems for many different applications: air traffic control, aircraft navigation, remote sensing, weather, ship safety, space rendezvous and docking, military, and law enforcement. Each of these systems consists of the same basic elements: transmitter, receiver, and signal processor. A detailed explanation of the most common radar system elements is given in Section 2.1.1. The most significant difference between radar for different applications is in the signal processing. Several types of radar and their associated form of signal processing are discussed in Section 2.1.2.

2.1.1 Block Diagram

A general radar system block diagram is given in *Figure 2.1*. The signal generator generates the transmitted waveform which may be a pulse, pulse train, chirped pulse, Gaussian pulse, triangle, or any other of an infinite number of potential waveforms. The waveform is designed so that the signal processing of the received signal gives the desired information about the target or medium.

The transmitted waveform is modulated onto a carrier frequency before being transmitted. Modulation is necessary to reduce the size of antennas, achieve better propagation properties, obtain better angular resolution, and to avoid interference with other signals in the spectrum. The power amplifier boosts the modulated signal to give the transmitted signal sufficient energy to be detectable when received.

The antenna is often used for both transmitting and receiving, though it need not be. If the antenna is used for both functions, the radar is monostatic. A radar with different transmit and receive antennas is multistatic. Many multistatic variations exist: a single transmit antenna with a single separated receiver; a dual transmit/receive antenna with a second separate receive antenna;
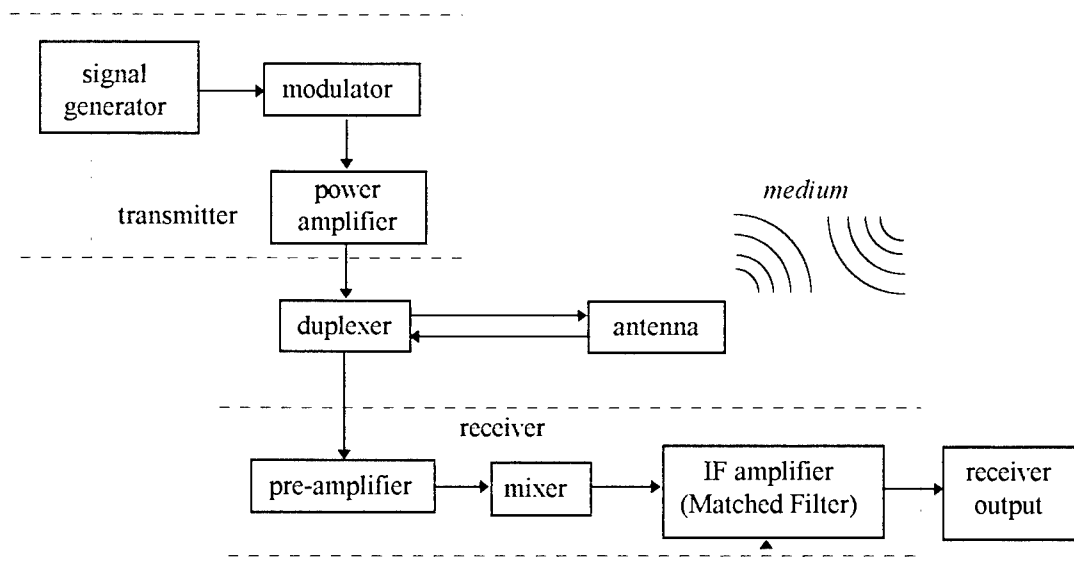
Figure 2.1: Radar system block diagram.

or many separated transmit and receive antennas. Multistatic variations with two antennas are known as bistatic radars. Multistatic radars can give extremely good angular resolution between antennas but at the cost of grating lobes and added complexity due to the need for phase coherence between systems.

Since the receiver is designed to detect signals that are often many orders of magnitude less power than the transmitter signal, the receiver must be protected from burnout when the transmitter is operating. This is the function of the duplexer. The receiver is shunted or turned off to isolate its delicate electronics during transmit.

The unprocessed received signal enters the pre-amplifier which is a low noise RF amplifier to boost the signal to a level that is within the operating range of the receiver electronics. The mixer demodulates the received signal to remove the carrier signal. Often, the signal is only taken down to an intermediate frequency where it is more easily manipulated.

The IF amplifier is a filter with a pass band at the intermediate frequency. Most of the pre-detection signal processing - e.g., coherent integration, - is done here, though post-detection processing can also be done on the final output signal. The IF amplifier is typically a matched filter – a filter whose output response signal to noise ratio is a maximum. The matched filter is covered in Section 2.2.3.

2.1.2 Types of Radar

A radar emits electromagnetic (EM) radiation and extracts information from the reflected energy. Sometimes, the absence of reflection is used to characterize the medium through which the electromagnetic wave propagated, but the vast majority of radars analyze the reflected signal.

18

Table 2.1: Radar band designations. (Compiled from [15,8])

Band	frequency	wavelength
VHF	30-300 MHz	10-1 m
UHF	300-1000 Mhz	1-.3 m
L	1-2 GHz	30-15 cm
S	2-4 GHz	15-7.5 cm
C	4-8 GHz	7.5-3.75 cm
X	8-12 GHz	3.75-2.5 cm
K_u	12-18 GHz	2.5-1.67 cm
K	18-27 GHz	17-11 mm
K_a	27-40 GHz	11-7.5 mm
mm	40-300 GHz	7.5-1 mm

There are four aspects to the information in the reflected, received signal: presence of a reflector, the EM wave's travel time to the reflector, frequency content of the received signal, and polarization of the reflection. Polarization is seldom used since antennas are typically designed for only one polarization. The presence of a reflected signal is used to detect the presence of targets. The time delay between the transmitted and received signal gives the range to the target if the speed of the EM wave is known for the medium. Finally, the spectrum of the received signal can be used to indicate the velocity of the target relative to the radar by the phenomenon of Doppler shift.

Every radar uses these aspects of information in the received signal in different ways, depending upon the function and role of the radar. For example, an air traffic control radar relies primarily upon the detection and ranging aspects to find and locate air traffic. A moving target indicator (MTI) radar relies heavily on the signal spectrum to separate moving targets from stationary clutter. MTI radar also uses range information to locate target position. A synthetic aperture radar (SAR) for imaging uses signal time delay to resolve objects in range and frequency information to resolve objects in cross-range. A range/cross-range image is constructed based on the presence and strength of reflection.

For a space based replacement to the AWACS mission, all three information aspects are important. Target detection is done by the presence of a return. The target is located by range within the radar beam footprint and the target velocity is determined by the signal spectrum. The spectrum is also used to detect moving targets in a strong clutter background.

Although radar can operate at almost any electromagnetic frequency from low frequency (LF), 30 kHz, to millimeter, 300 GHz, the majority of radar systems are designed in the ultra high frequency (UHF) and microwave bands. The microwave frequencies are divided into letter designated subbands for convenience. The letters denoting the subbands form no particular

pattern since they were originally conceived as a code during World War II. Nevertheless, the letter band designation has persisted to this day. Though the SI system of units has defined the New Band designations for frequencies from VHF through microwave, radar engineers still exclusively use the old microwave subbands. Some confusion is possible since some of the band designations are redundant but refer to different bands. For example, L-band ranges from 1 GHz to 2 GHz, but under the new band designation, L band denotes 40-60 GHz [8]. We use the traditional microwave subband designations given in *Table 2.1*.

2.2 Radar Signal

The radar emits an electromagnetic wave that propagates through some medium, reflects off an object, and then propagates back through the medium to the radar. If the transmitter signal is $s_{tx}(t)$, then the receiver signal is

$$s_{rx}(t) = \alpha s_{tx}(t - t_d)e^{j2\pi f_d t} + s_{clutter}(t) + n_{thermal}(t) \qquad eqn\ 2.1$$

where α is the signal attenuation due to propagation, medium absorption or scattering, and reflectivity of the target. The propagation time delay is

$$t_d = \frac{2r_{tgt}}{c} \qquad eqn\ 2.2$$

where r_{tgt} is the range to the target and c is the speed of light in the propagation medium. The received signal experiences a doppler shift relative to the transmitted signal given by

$$f_d = \frac{2\dot{r}_{tgt}}{\lambda} \qquad eqn\ 2.3$$

where $\dot{r}_{tgt}$ is the relative velocity between the radar and target and λ is the transmitted signal wavelength.

The purpose of signal processing is to extract information from the received electromagnetic wave. Careful design of the transmitted waveform can improve the quality of information contained in the received waveform. The complex envelope of the waveform is created in the signal generator and is then modulated on the carrier frequency. It is convenient to characterize the transmitted waveform in terms of the complex envelope of the signal since the modulation rarely encodes any additional information on the transmitted signal. The complex envelope can be found by quadrature demodulation of the transmitted signal [11].

2.2.1 Correlation Receiver

For target detection, the receiver processing must distinguish between two hypotheses. The null hypothesis is that the signal in the receiver is due to noise alone. The alternative hypothesis is that the signal is a received signal plus noise. In the notation of hypothesis testing,

20
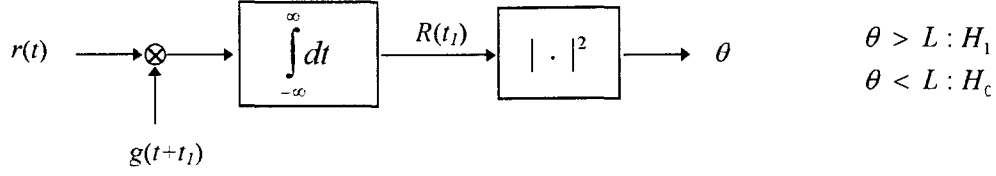
Figure 2.2: Correlation receiver block diagram and hypothesis decision criteria.

$$H_0: r(t) = n(t)$$
$$H_1: r(t) = s_r(t) + n(t)$$

eqn 2.4

where $r(t)$ is the receiver signal; $s_r(t)$ is the received signal (the attenuated, delayed, and doppler shifted transmitted signal); and $n(t)$ is the noise in the receiver. We will model the receiver noise as white noise and assume that the received signal and noise are uncorrelated.

The correlation receiver [9] decides between the null and alternative hypotheses on the basis of the value of correlation between the received signal and some internal receiver signal, $g(t)$. The process is illustrated in *Figure 2.2*. We will solve for the best $g(t)$. The correlation between the receiver signal and $g(t)$ is

$$R(t_1) = \int_{-\infty}^{\infty} r(t)g(t + t_1)dt$$

eqn 2.5

which is a random variable since the receiver signal is a function of random noise.

The correlation receiver output is θ, the squared magnitude of the correlation R. If θ is greater than some threshold value, L, the signals are correlated and the receiver signal is probably not due to noise alone. Thus, the alternative hypothesis is concluded. Conversely, if θ is less than L, the null hypothesis is concluded.

The receiver signal to noise ratio is equal to the ratio of the variance of the correlation random variables under the two hypotheses. That is,

$$\frac{S}{N} = \frac{\sigma_{R_1}^2}{\sigma_{R_0}^2}$$

eqn 2.6

The variance of the correlation under the null hypothesis is

$$E\left[R^2 \mid H_0\right] = \sigma_{R_0}^2 = E\left[\left(\int_{-\infty}^{\infty} n(t)g(t)dt\right)^2\right] = E\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} n(t_1)n(t_2)g(t_1)g(t_2)dt_1dt_2\right]$$

eqn 2.7

which under the assumption of white noise is

$$\sigma_{P_0}^2 = \sigma_n^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta(t_1 - t_2)g(t_1)g(t_2)dt_1 dt_2 = \sigma_n^2 \int_{-\infty}^{\infty} g^2(t)dt \qquad\qquad eqn\ 2.8$$

where σ_n^2 is the variance of the white noise. Note that the existence of a finite value for the noise variance implies that the white noise is band limited white noise with a power spectrum that eventually rolls off at a sufficiently high frequency. Within the bandwidth of interest, however, the noise spectrum is flat.

Similarly, the variance of the correlation under the alternative hypothesis is

$$\sigma_{P_1}^2 = E\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(s_r(t_1)+n(t_1)\right)\left(s_r(t_2)+n(t_2)\right)g(t_1)g(t_2)dt_1 dt_2\right] \qquad\qquad eqn\ 2.9$$

which, under the uncorrelated signal and noise assumption is

$$\sigma_{P_1}^2 = \left(\int_{-\infty}^{\infty} s_r(t)g(t)dt\right)^2 + \sigma_n^2 \int_{-\infty}^{\infty} g^2(t)dt \qquad\qquad eqn\ 2.10$$

The signal to noise ratio is then

$$\frac{S}{N} = \frac{\sigma_{P_1}^2}{\sigma_{P_0}^2} = \frac{\left(\int_{-\infty}^{\infty} s_r(t)g(t)dt\right)^2}{\sigma_n^2 \int_{-\infty}^{\infty} g^2(t)dt} + 1 \qquad\qquad eqn\ 2.11$$

By the Schwarz inequality[1] [10]

$$\frac{S}{N} = \frac{\sigma_{P_1}^2}{\sigma_{P_0}^2} \leq \frac{\int_{-\infty}^{\infty} s_r^2(t)dt \int_{-\infty}^{\infty} g^2(t)dt}{\sigma_n^2 \int_{-\infty}^{\infty} g^2(t)dt} + 1 = 1 + \frac{\int_{-\infty}^{\infty} s_r^2(t)dt}{\sigma_n^2} \qquad\qquad eqn\ 2.12$$

By comparing *eqn 2.11* and *eqn 2.12*, we get the maximum signal to noise ratio - the equal part of the less than or equal - when $g(t)$ equals the complex conjugate of the received signal. Since the received signal is a delayed replica of the transmitted signal (*eqn 2.1*), the maximum signal to noise ratio is achieved by correlating an appropriately delayed and frequency shifted replica of the transmitted signal with the received signal. Thus

$$g(t) = s_{tx}\left(t - \hat{t}_d\right)\exp(-j2\pi\hat{f}_d t) \qquad\qquad eqn\ 2.13$$

where $\hat{t}_d$ is the estimated time delay and $\hat{f}_d$ is the estimated doppler shift. The values of $\hat{t}_d$ and

---

[1] The Schwarz inequality for integrals: $\left|\int_a^b f(x)g(x)dx\right| \leq \left[\int_a^b f^2(x)dx\right]^{1/2}\left[\int_a^b g^2(x)dx\right]^{1/2}$
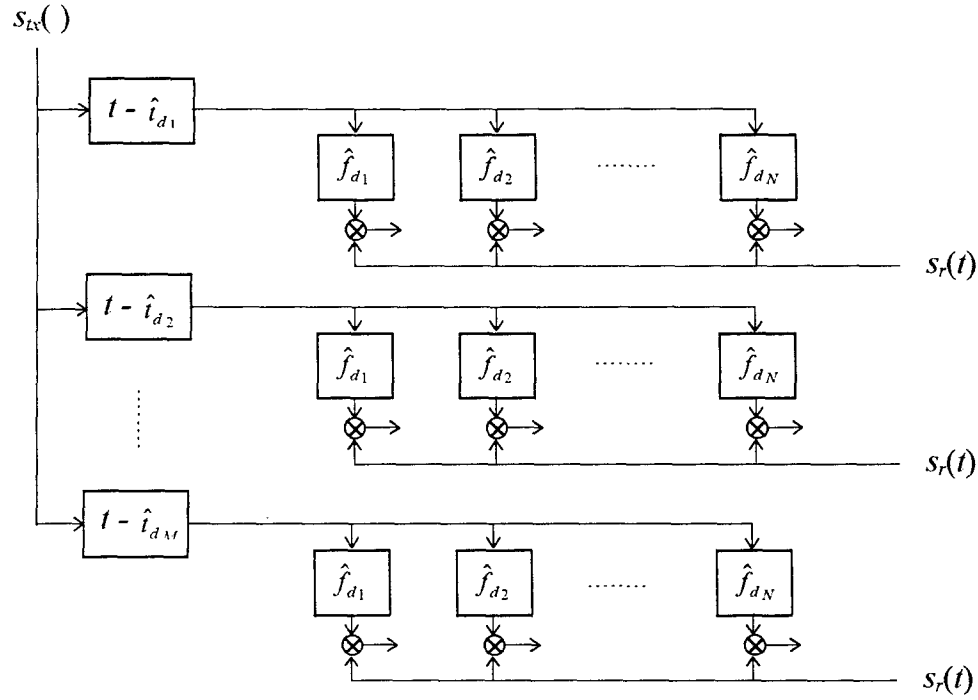
Figure 2.3: Correlation processing block diagram.

$\hat{f}_d$ which give the maximum value of correlation give the time delay and doppler shift of the target. This then implies the receiver processing algorithm depicted in *Figure 2.3*.

2.2.2 Ambiguity Plane

Each point of output in *Figure 2.3* - that is, the correlation for every time delay and doppler shift - becomes a point in the range-doppler plane. The correlation magnitude squared is the ambiguity function and the corresponding graphical representation is the ambiguity plane. It is called ambiguity because of the ever-existent tradeoff between range and frequency resolution. Because range (time) and frequency are Fourier conjugates, it is impossible to achieve better resolution in one dimension without sacrificing resolution in the other dimension. Plots of the ambiguity function are useful for visualizing the resolution inherent in different types of waveforms. The ambiguity function is the square of the correlation function magnitude. That is

$$\theta\left(t_d, f_d\right) = \left|\phi\left(t_d, f_d\right)\right|^2 \qquad eqn\ 2.14$$

where $\phi(t_d, f_d)$ is the correlation function,

$$\phi\left(t_d, f_d\right) = \int_{-\infty}^{\infty} \tilde{f}\left(t - t_d\right)\tilde{f}^*\left(t - \hat{t}_d\right)e^{-j2\pi\left(f_d - \hat{f}_d\right)t}\,dt \qquad eqn\ 2.15$$

In the definition of the correlation function, we have replaced the received signal and

transmitted signal with the complex envelope of the waveform, $\tilde{f}(t)$, since it is the waveform envelope that will be used to analyze waveform properties. When implemented in a correlation receiver, however, the received and delayed transmitted envelopes are used as implied in Section 2.2.1. The * notation on the complex envelope indicates complex conjugate.

For the true range and doppler shifts, the correlation function is an autocorrelation along the time axis and a Fourier transform along the frequency axis.

Figure 2.3 shows how large the processing load can be. For a correlation receiver with M range processing bins and N frequency processing bins, the receiver must correlate M times and perform MN Fourier transforms. This can be a significant processing load. A receiver processing technique that is mathematically equivalent to the correlation method but with fewer processing channels is the matched filter.

2.2.3 Matched Filter

The matched filter is the most widely used IF amplifier of all radar systems. By definition, the matched filter has a frequency response that maximizes the output peak signal to mean noise ratio. The frequency response of a filter gives the magnitude and phase of the filter output signal relative to the signal input. If the filter passband is wide compared to the signal bandwidth, noise energy from the extra bandwidth will degrade the output signal and decrease the signal to noise ratio. If the filter passband is narrower than the signal bandwidth, some of the signal is lost and the signal to noise ratio is again decreased. The matched filter optimizes the filter bandwidth to give the maximum signal to noise ratio.

The matched filter frequency response function is [15,11]

$$H(f) = G_a \tilde{S}^*(f) \exp(-j2\pi f t_1) \qquad eqn\ 2.16$$

where G_a is a constant filter gain, $\tilde{S}^*(f)$ is the complex conjugate of the Fourier transform of the input signal, and t_1 is the fixed time for which the input signal is a maximum. The matched filter is also known as the North filter, the conjugate filter, or the Fourier transform criterion.

The impulse response of a filter is the inverse Fourier transform of the frequency-response function. The matched filter impulse response is

$$h(t) = G_a \int_{-\infty}^{\infty} \tilde{S}^*(f) e^{-j2\pi f(t_1-t)} df = G_a \int_{-\infty}^{\infty} \tilde{S}(f) e^{j2\pi f(t_1-t)} df = G_a \tilde{s}(t_1 - t) \qquad eqn\ 2.17$$

where the relationship $\tilde{S}^*(f) = \tilde{S}(-f)$ has been used and $\tilde{s}(t)$ is the complex envelope of the input signal. Thus, the matched filter response to an impulse is the time reflected expected input signal – that is the input signal played backwards in time. This illustrates that the matched filter has "built in" the expected waveform. To construct the matched filter, the reflected time image of the transmitted waveform envelope is used to create the desired impulse or frequency response,

$$\tilde{f}(t) + n(t) \longrightarrow \boxed{h(t)} \longrightarrow x(t)$$

$$\tilde{f}(t_1 - t)$$
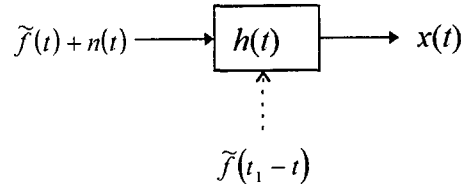
Figure 2.4: Matched Filter implementation.

Figure 2.4. Then, the received waveform - a delayed replica of the transmitted waveform corrupted by noise (*eqn 2.1*) - after matched filtering, has the peak signal to mean noise ratio. Note that the actual shape of the received signal is not preserved by the filtering process. For many applications, however, the shape of the received signal is not important as long as the frequency information and correlation peak are preserved.

The matched filter implementation described in *Figure 2.4* is mathematically equivalent to the correlation receiver derived in Section 2.2.1. The temporal output of the matched filter, $x(t)$, is the convolution of the input single and the impulse response function.

$$x(t) = \int_{-\infty}^{\infty} \left(\tilde{f}(t) + n(t) \right) h(t - \xi) d\xi \qquad \qquad eqn\ 2.18$$

where $h(t) = \tilde{f}(t_1 - t)$ by *eqn 2.17*. Therefore

$$x(t) = \int_{-\infty}^{\infty} \tilde{f}(t)\tilde{f}(t_1 - t - \xi)d\xi + \int_{-\infty}^{\infty} n(t)\tilde{f}(t_1 - t - \xi)d\xi = R(t - t_d) \qquad eqn\ 2.19$$

where $R(t\text{-}t_d)$ is the cross-correlation between the transmitted and received waveforms delayed by a time increment equal to the wave travel time. Note that the cross-correlation between the signal and noise is approximately zero. This is exactly what the correlation receiver does. With the correlation receiver, however, the signal cross-correlation had to be computed for a series of time delays, *Figure 2.3*. For the matched filter, this cross-correlation is built into a linear, time invariant filter through the filter frequency response. The matched filter implementation equivalent to *Figure 2.3* is given in *Figure 2.5*.

Whereas the correlation receiver required MN processing steps, the matched filter only needs N processing steps. The block of matched filters for different doppler shifts essentially rasters along the frequency axis in the range-doppler plane. Peaks of the ambiguity function correspond to potential targets. The position of the peak gives the signal travel time, t_d, and the target doppler shift, f_d. Range and velocity may be determined from *eqn 2.2* and *eqn 2.3*.

Although *Figure 2.5* depicts the matched filter implementation with the ambiguity function, the ambiguity function is seldom used in the actual detection process. A detection is declared when the matched filter output exceeds some detection threshold. The actual ambiguity plane need not be computed. However, the ambiguity function is useful for visualizing the
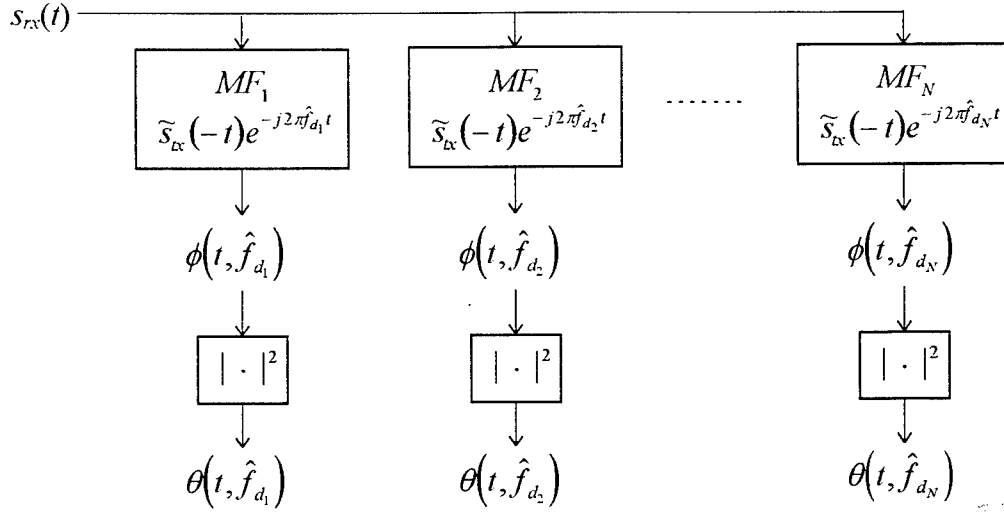
Figure 2.5: Matched Filter processing block diagram.

process and analyzing and comparing waveforms. This is done in the following section.

2.2.4 Ambiguity function properties and expressions for common signals

The ambiguity function, *eqn 2.14*, has several properties that are useful in analyzing waveforms. For the purposes of waveform analysis, we use the normalized ambiguity function so that the ambiguity integrated over the range doppler plane is one.

Ambiguity function properties: [9]

1. Volume Invariance (time-frequency uncertainty principle)

   The ambiguity function always integrates to one. Therefore, any energy taken from the mainlobe must appear somewhere else in the ambiguity plane:

   $$\iint \theta(t_d, f_d) dt_d df_d = 1 \qquad\qquad eqn\ 2.20$$

2. Symmetry

   $$\phi(t_d, f_d) = \phi^*(-t_d, -f_d) \qquad\qquad eqn\ 2.21$$

   $$\theta(t_d, f_d) = \theta(-t_d, -f_d) \qquad\qquad eqn\ 2.22$$

3. Scaling

   $$\tilde{f}(t) \Leftrightarrow \phi(t_d, f_d)$$

   $$\sqrt{\alpha}\tilde{f}(\alpha t) \Leftrightarrow \phi\left(\alpha t_d, \frac{f_d}{\alpha}\right) \qquad\qquad eqn\ 2.23$$
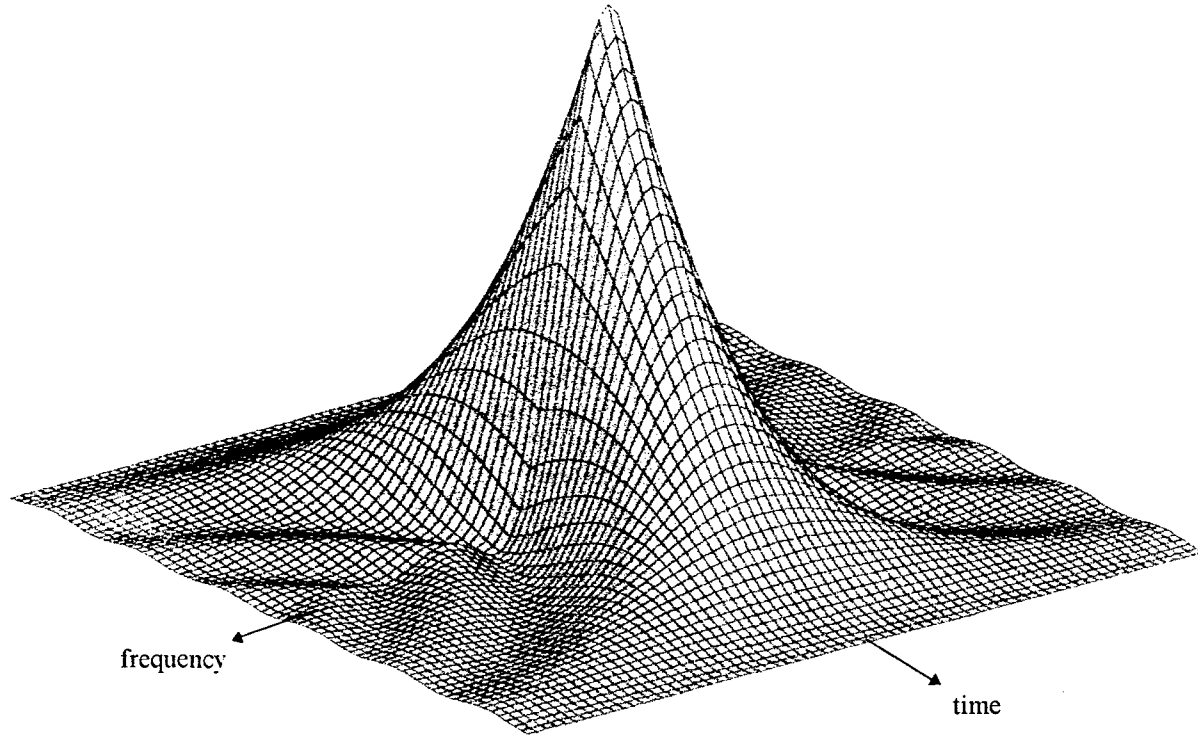
Figure 2.6: Rectangular pulse ambiguity function.

4. Multiplication

$$\tilde{f}_1(t) \cdot \tilde{f}_2(t) \Leftrightarrow \int_{-\infty}^{\infty} \phi_1(t_d, x)\phi_2(t_d, f_d - x)dx$$

$$\tilde{F}_1(t) \cdot \tilde{F}_2(t) \Leftrightarrow \int_{-\infty}^{\infty} \phi_1(x, f_d)\phi_2(t_d - x, f_d)dx$$

eqn 2.24

The waveform envelopes and ambiguity functions for several common waveforms are listed below. These are compiled from [12,9]:

1. Single rectangular pulse of width τ_p:

$$\tilde{f}(t_d) = \begin{cases} \dfrac{1}{\sqrt{\tau_p}} & -\dfrac{\tau_p}{2} < t_d < \dfrac{\tau_p}{2} \\ 0 & otherwise \end{cases}$$

eqn 2.25

$$\theta(t_d, f_d) = \begin{cases} \left(1 - \dfrac{|t_d|}{\tau_p}\right)^2 sinc^2\left(2\pi f_d \dfrac{\tau_p}{2}\left(1 - |t_d|/\tau_p\right)\right) & |t_d| \le \tau_p \\ 0 & otherwise \end{cases}$$

eqn 2.26

2. Single Gaussian pulse:

$$\tilde{f}(t_d) = \left(\frac{1}{\pi\tau_p^2}\right)^{1/4} \exp\left(-\frac{t_d}{2\tau_p^2}\right) \qquad\qquad eqn\ 2.27$$

$$\theta(t_d, f_d) = \exp\left[-\frac{1}{2}\left(\frac{t_d^2}{\tau_p^2} + \left(2\pi f_d \tau_p\right)^2\right)\right] \qquad\qquad eqn\ 2.28$$

3. Rectangular pulse train:

$$\tilde{f}(t_d) = \frac{1}{(2n+1)T_p^{1/2}} \sum_{k=-n}^{k=n} \tilde{u}(t_d - kT_p) \qquad\qquad eqn\ 2.29$$

$$\theta(t_d, f_d) = \left|\frac{1}{2n+1}\left(\frac{\sin\left(2\pi f_d (n+1/2)T_p\right)}{\sin\left(2\pi f_d T_p/2\right)}\right)\left[\left(1-\frac{|t_d|}{\tau_p}\right)sinc\left(2\pi f_d \frac{\tau_p}{2}\left(1-|t_d|/\tau_p\right)\right)\right]\right|^2 \quad eqn\ 2.30$$

for $|t_d| \leq \tau_p$ and 0 otherwise.

2.3 Waveform Resolution

Resolution is the minimum amount by which two distinct objects have to be separated before their signal returns are separable. For radar, there are two important types of resolution: range resolution and frequency resolution. Both are governed primarily by the waveform and the signal to noise ratio. The waveform determines how the uncertainty between time and frequency is allocated. The signal to noise ratio determines how accurately the signal can be characterized in the presence of noise. Note the difference between processing resolution and waveform resolution. Processing resolution depends on the number of discrete processing bins, e.g., Doppler shift bins in *Figure 2.5*. Waveform resolution depends on the waveform uncertainty in time and frequency.

Since time and frequency are Fourier conjugates, they cannot both be simultaneously measured without uncertainty. This is the radar uncertainty principle (analogous to Heisenberg's uncertainty principle) and can be stated as [13]

$$\Delta t \Delta f \approx 1 \qquad\qquad eqn\ 2.31$$

where Δt is the uncertainty in time and Δf is the uncertainty in frequency. The uncertainty diagram is a level curve of the main lobe of the waveform ambiguity function. For example, the rectangular pulse uncertainty diagram given in *Figure 2.7* is a horizontal slice through the ambiguity diagram in *Figure 2.6*.

There are many different definitions for waveform resolution. One convention is that range resolution is given by the width of the main peak of the autocorrelation function and
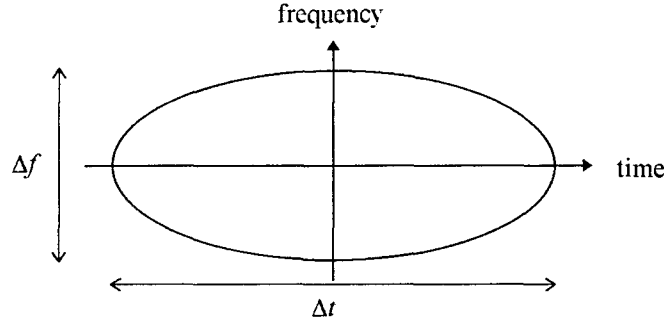
Figure 2.7: Uncertainty diagram for rectangular pulse.

frequency resolution is defined as the width of the main lobe of the power spectrum. Since the ambiguity function is an autocorrelation and Fourier transform of the waveform along two axes, waveform resolution is proportional to uncertainty.

The error in measuring time delay or Doppler shift from the received signal depends on the waveform uncertainty (waveform resolution) and the signal to noise ratio. Increasing the signal to noise ratio increases the accuracy of the measurement. For a matched filter process, Skolnik [15] gives the range error as

$$\delta R = \frac{c}{2} \frac{1}{\beta \sqrt{2E N_0}}$$ eqn 2.32

where β is the effective bandwidth of the waveform and E/N_0 is the signal energy to noise per unit bandwidth ratio. The reciprocal of β is analogous to the waveform uncertainty described above. The effective bandwidth is defined as

$$\beta^2 = \frac{\int\limits_{-\infty}^{\infty}(2\pi f)^2 \tilde{S}(f)\tilde{S}^*(f)df}{\int\limits_{-\infty}^{\infty}\tilde{S}(f)\tilde{S}^*(f)df} = \frac{1}{E}\int\limits_{-\infty}^{\infty}(2\pi f)^2 \tilde{S}(f)\tilde{S}^*(f)df$$ eqn 2.33

where E is the energy in the signal.

Similarly, the error in measuring frequency is [15]

$$\delta f = \frac{1}{\alpha \sqrt{2E N_0}}$$ eqn 2.34

where α is the effective time duration of the signal defined as

29
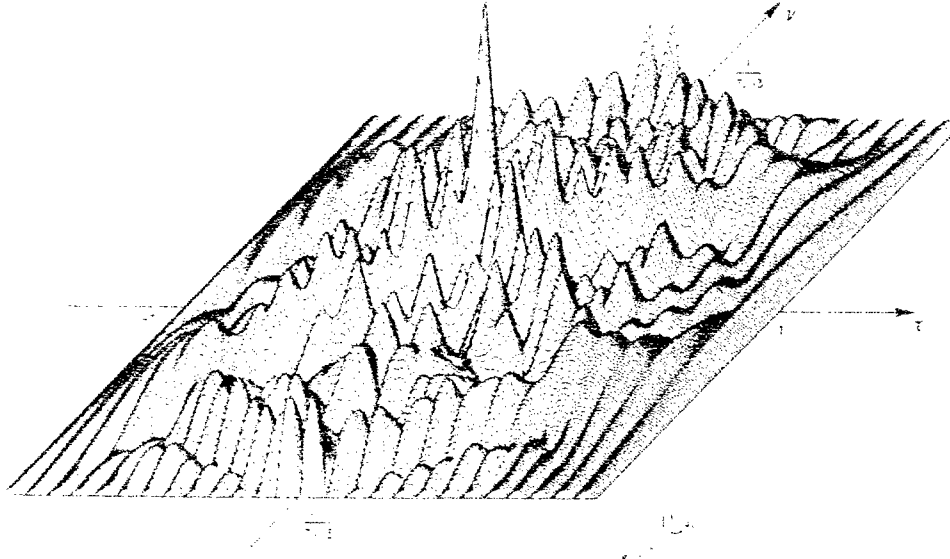
Figure 2.8: Ambiguity diagram for the 13 bit Barker code. The uncertainty in the mainlobe has decreased, but at the expense of increased ambiguity in the sidelobes. Figure from [13].

$$\alpha^2 = \frac{\int\limits_{-\infty}^{\infty}(2\pi t)^2 \tilde{s}(t)\tilde{s}^*(t)dt}{\int\limits_{-\infty}^{\infty}\tilde{s}(t)\tilde{s}^*(t)dt}$$

eqn 2.35

Another version of the uncertainty relation often used by radar engineers is [15]

$$\beta\alpha \geq \pi$$

eqn 2.36

At first glance, there is contradiction in the interpretation of the two versions of the uncertainty relation, *eqn 2.31* and *eqn 2.36*. This is due to differences in the definitions -- *eqn 2.31* governs global uncertainty and *eqn 2.36* local uncertainty. Skolnik's [15] version of uncertainty relation, *eqn 2.36*, implies that there is no upper bound to the accuracy achievable by the waveform. Large effective bandwidth-effective time duration products give simultaneous improvements in range and frequency resolution. This requires waveforms that are long in duration and of wide bandwidth. Coded sequences such as Barker codes and pseudo-random noise (PRN) codes are two ways of achieving long duration waveforms with wide bandwidth. Pulse chirping is another technique. Note that although the main lobe resolution can be increased with these techniques, the ambiguity is conserved and appears in the sidelobes of the ambiguity functions (compare the ambiguity diagram for the 13 bit Barker code with rectangular pulse ambiguity diagram, *Figure 2.6*). This is the essence of the uncertainty relation as defined by *eqn 2.31*. Without a sufficiently high signal to noise ratio, the transition of uncertainty from the mainlobe to the sidelobes may not be advantageous. The simultaneous range and frequency error measurements are governed by

30

$$\delta t_d \delta f_d = \frac{1}{\beta \alpha (2E \ N_0)}$$

2.3.1 Pulse Compression

Pulse compression is a technique for encoding more information on the transmitted waveform so as to improve the resolution of the received waveform. Pulse compression increases the product of the waveform effective bandwidth and time duration. The most effective pulse compression techniques employ phase coded sequences, though they also require more complex transmitters, receivers, and processing. The most widely used pulse compression technique in radar is linear frequency modulation or chirping.

Coded pulses can be thought of as a sequence of subpulses strung together to form a much longer pulse. The effective pulse duration is the length of the entire pulse and the effective bandwidth is the reciprocal of the subpulse duration. Thus, the coded waveform has a large effective bandwidth-duration product compared to the subpulse. Phase between the subpulses is a common method for coding, though it is conceivable - albeit less practical - to code the subpulses by amplitude or frequency. The sequence of subpulse phases, either 0 or π, defines the code. Examples of coded sequences include Barker codes, pseudo-random noise (PRN) sequences, linear recursive sequences, and binary-shift-register sequences. The Barker code has the optimal autocorrelation function with the lowest and most uniform sidelobes, however the longest Barker code is only 13 bits long. This is too short to be practical for many applications. PRN codes have the advantage of being resistant to spoofing.

The most widely used form of pulse compression used in radar is linear frequency modulation, also known as chirping. Linear frequency modulation is a constant increase in the frequency of modulation. Note that this modulation is distinct from the modulation onto the
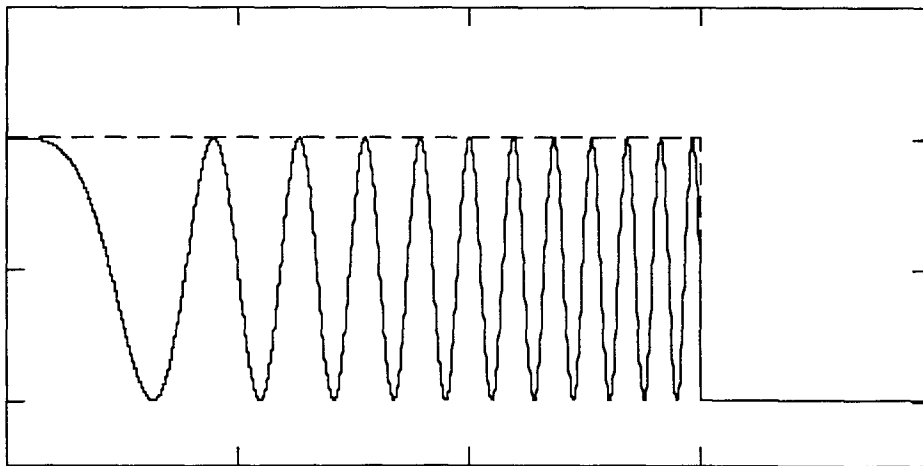
Figure 2.9: Chirped rectangular pulse complex envelope.
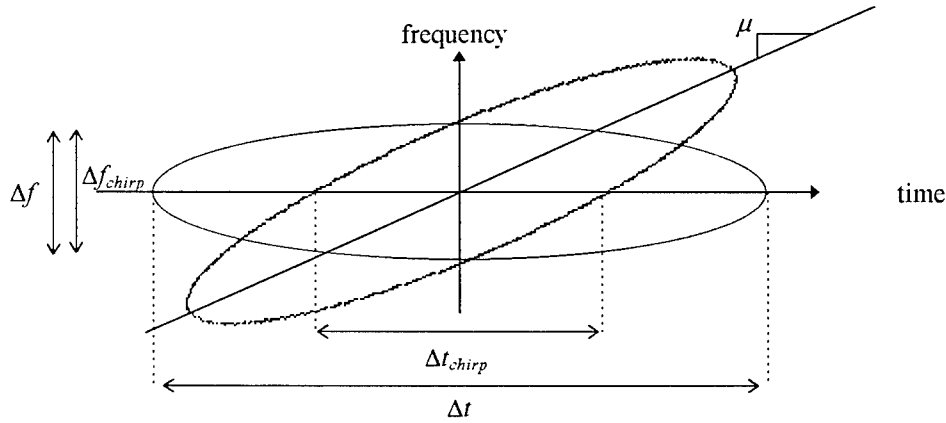
Figure 2.10: Chirped waveform uncertainty.

carrier frequency. The waveform is first chirped and then modulated on the carrier frequency. The chirped waveform is a frequency modulated version of the original waveform, *Figure 2.9*. That is,

$$\tilde{s}_{chirp}(t) = \tilde{s}(t)e^{j2\pi\frac{\mu t^2}{2}}$$

eqn 2.38

where μ is the chirp rate, e.g. in Hz per sec. Chirping improves the range resolution by increasing the effective bandwidth of the signal. The range resolution of the unchirped signal is proportional to α, but the range resolution of the chirped signal is proportional to $1/\mu\alpha$. The frequency resolution is often decreased only negligibly. Linear frequency modulation results in a rotation of the waveform uncertainty as depicted in *Figure 2.10*.

2.4 Radar Range Equation

The radar range equation is one of the most well known techniques for analyzing radar performance. It relates the maximum range at which targets can be detected to the transmitter power, antenna gain and area, signal to noise ratio, signal integration, system losses, thermal noise, and target radar cross section. Different versions of the radar range equation exist for search radar, track radar, radar under jamming, radar with clutter backgrounds, and many other scenarios. We derive the general radar range equation and then adapt it for a space based search radar.

An electromagnetic source emits radiation that expands in free space as a spherical wave. The amplitude of the wave decreases as the inverse distance that the wave has traveled from the source [14]. Since energy is proportional to the square of the wave amplitude, energy decreases with the inverse of the range squared. Signal power, as the temporal derivative of the energy, also decreases with the inverse of the squared range from the signal source. This is equivalent to thinking of the signal power being distributed across the surface of an ever increasing sphere. The

power density is the signal power divided by the surface area of the sphere. The power density of an isotropic source is then

$$power\ density\ at\ R = \frac{P_t}{4\pi R^2}$$

where P_t is the transmitted power and R is the range from the source. Note that power density has the units of power per area, e.g. W/m^2.

Radar systems almost always employ an antenna to direct the transmitted energy in a preferred direction. The directive gain or directivity is the ratio of the maximum radiation intensity per unit solid angle to the average radiation intensity of the beam pattern. Equivalently,

$$G_D = \frac{maximum\ power\ radiated\ unit\ solid\ angle}{total\ power\ radiated\ 4\pi}$$

The maximum power radiated per unit solid angle must be determined from the beam pattern. Solid angles are measured in steradians and relate surface area to the solid angle in the same manner that radians relate circumferential length to angles. That is, a solid angle of 1 steradian intersects R^2 surface area of a sphere of radius R. 4π steradians encompass the entire surface of a sphere in the same way that 2π radians cover the entire circumference of a circle. If $P(\theta, \phi)$ is the radiation intensity pattern as a function of azimuth and elevation with respect to the antenna, the directive gain is

$$G_D = \frac{4\pi P(\theta,\phi)_{max}}{\iint P(\theta,\phi)d\theta d\phi}$$

Closely related to the directive gain is the power gain. Directivity does not account for dissipative losses in the antenna. Power gain is always less than directive gain and is calculated by

$$G = \frac{maximum\ power\ radiated\ unit\ solid\ angle}{net\ power\ accepted\ by\ antenna\ 4\pi}$$

Gain is related to effective aperture. An antenna with effective aperture A_e (units of area) has a gain given by

$$G = \frac{4\pi}{\lambda^2} A_e$$

where λ is the transmitted signal wavelength. Thus, gain is a function of the antenna area and the radar frequency. Higher frequencies diffract less over a given range than lower frequencies. This implies a tighter radiation pattern and hence greater power gain. The effective aperture area is related to the true aperture area by the aperture efficiency, η_A.

$$A_e = \eta_A A$$

The signal power density of an antenna with gain G is

Table 2.2: Typical radar cross sections at microwave frequencies. Compiled from [15,16].

target	RCS (m^2)
conventional, unmanned, winged missile	.5
small, single engine aircraft	1
fighter	2-6
bomber / jet airliner	20-40
jumbo jet	100
pickup truck	200
automobile	100
bicycle	2
man	1
bird	.01
insect	10^{-5}

$$\textit{power density at } R = \frac{P_t G}{4\pi R^2} \qquad\qquad eqn\ 2.45$$

This is the power density incident upon a target at range R. The total power incident on the target is the product of the cross-sectional area of the target and the power density. The radar cross-section (RCS), σ, of a target is the fictional area that - reflecting in all directions - gives the power density at the radar receiver. That is, the target is an isotropic source that emits an amount of power equal to the product of the power density at the target and the RCS. Typical values for the RCS of common objects are given in *Table 2.2*. These values represent an average RCS. Actual RCS of a target depends strongly on the azimuth and elevation perspective, *Figure 2.11*. The power density of the received signal at the radar due to a target at range R with RCS σ is then

$$\textit{power density at radar } = \textit{ power density at } R\ \frac{\sigma}{4\pi R^2} = \frac{P_t G \sigma}{(4\pi)^2 R^4} \qquad eqn\ 2.46$$

This power density is incident upon the radar antenna with effective aperture area A_e. The power received by the radar is the product of the power density and the aperture area. This assumes that the aperture is normal to the incident wave. If the antenna is pointed in a different direction, the area is the apparent cross-sectional area and there is a corresponding squint loss or scan loss in the power received. The power of the lossless signal is then

$$\textit{power received by radar } = \textit{ signal power } = \frac{P_t G \sigma}{(4\pi)^2 R^4} A_e \qquad eqn\ 2.47$$

Every radar system experiences some signal power loss. There are three classes of losses: system loss, propagation medium loss, and ground plane loss. System losses include plumbing loss, polarization loss, antenna pattern loss, pulse width loss, squint loss, limiting loss, collapsing loss, operator loss, and non-ideal equipment loss. Propagation medium loss accounts for
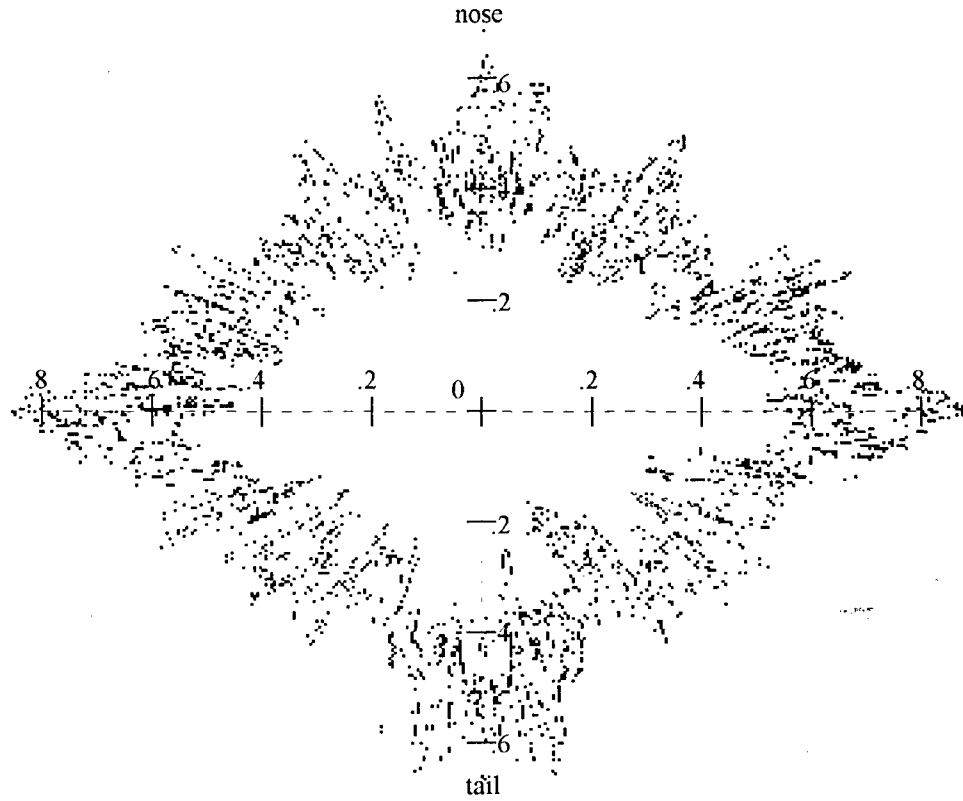
Figure 2.11: F-15 radar cross section in relative magnitude plotted in azimuth. [17]

scattering and absorption of the signal by gasses and particles in the atmosphere. A good example is rain loss. Ground plane loss comes from interference of the received signal with multipath returns. Losses are always greater than or equal to unity and give the fraction of the signal power that is received. For example, 3 dB losses indicate that only half of the possible received power actually becomes signal -- thus L equals 2. For a system with total loss given by L, the signal power is

$$S = \frac{1}{L} \frac{P_t G A_e \sigma}{(4\pi)^2 R^4}$$

eqn 2.48

Signal to noise ratio is a crucial performance driver for signal processing. The signal power is given by *eqn 2.48*. Even in the absence of external noise sources, there is noise due to the thermal motion of electrons in the resistive portions of the receiver electronics. The thermal noise power is

$$N = kTB_n$$

eqn 2.49

where k is Boltzman's constant ($1.38 \ 10^{-23}$ J/deg K), T is the temperature, and B_n is the receiver noise bandwidth. If H(f) is the matched filter transfer function, the noise bandwidth is

$$B_n = \frac{\int_{-\infty}^{\infty} |H(f)|^2 \, df}{|H(f_{max})|^2} \qquad\qquad eqn\ 2.50$$

where f_{max} is the frequency of the peak response of the matched filter. The noise bandwidth is the bandwidth of the rectangular filter with the same noise power output as the matched filter. There is a characteristic relationship given below between the pulse length and noise bandwidth for the matched filter that will allow us to avoid calculating the bandwidth.

Combining *eqn 2.48* and *eqn 2.49* gives the signal to noise ratio of the received signal.

$$\frac{S}{N} = \frac{P_t G A_e \sigma}{(4\pi)^2 R^4 k T B_n L} \qquad\qquad eqn\ 2.51$$

Signal integration is used to increase the signal to noise ratio. Since we are dealing with discrete radar pulses, integration is actually summation of the received pulses. The motivation behind integration is that the signal tends to reinforce itself and the noise tends to cancel itself. The integration process is most efficient in reinforcing the signal if the signal is coherent between integration steps. Coherence means that the signal phase is preserved. In the block diagram in *Figure 2.1*, there are two obvious places to do signal integration: before the IF amplifier and after the IF amplifier. The IF amplifier destroys the phase information in the signal, so that integration after the IF filter is noncoherent. Signal integration after matched filtering is also known as post-detection integration since the integration occurs after detection processing. The coherent integration before the IF filter is pre-detection processing. The n-pulse integrated signal to noise ratio is

$$\left(\frac{S}{N}\right)_n = n \left(\frac{S}{N}\right)_1 \eta_{int}(n) \qquad\qquad eqn\ 2.52$$

where $\eta_{int}(n)$ is the efficiency of the integration process. For coherent integration, this efficiency is exactly unity. For post-detection integration, the efficiency is less than one. The integrated signal to noise ratio is then

$$\left(\frac{S}{N}\right)_n = \frac{P_t G A_e \sigma n \eta_{int}}{(4\pi)^2 R^4 k T B_n L} \qquad\qquad eqn\ 2.53$$

Instantaneous transmitter power is a term in *eqn 2.51*. Average transmitter power is a better term to use in the radar range equation since it gives a better measure of the average signal to noise ratio and the average detection range. Average power is computed from instantaneous power by

$$P_{av} = \frac{1}{T_p} \int_0^{T_p} P_t(t) \, dt = \frac{P_t \tau}{T_p} = P_t \, \tau prf \qquad\qquad eqn\ 2.54$$

The pulse width, τ, times the pulse repetition frequency, *prf*, is the radar duty cycle. In terms of

average power, the signal to noise ratio is

$$\frac{S}{N} = \frac{P_{av}GA_e\sigma}{(4\pi)^2 R^4 kTB_n\tau prf L} \qquad eqn\ 2.55$$

For a matched filter the pulse width bandwidth product, $B_n\tau$, is approximately unity (see Section 5.2.1). Thus, detailed knowledge of the matched filter response or waveform is not necessary to use the radar range equation. Solving for range, the radar range equation becomes

$$R^4 = \frac{P_{av}GA_e\sigma n\eta_{int}}{(4\pi)^2 kTB_n\tau prf L(S\ N)_1} \qquad eqn\ 2.56$$

where $(S/N)_1$ is the signal to noise ratio for a single pulse.

Substituting *eqn 2.43* for the power gain, the radar range equation is

$$R^4 = \frac{P_{av}A_e^2\sigma n\eta_{int}}{4\pi\lambda^2 kTB_n\tau prf L(S\ N)_1} \qquad eqn\ 2.57$$

2.4.1 Radar Range Equation for Search Radar

The radar range equation, *eqn 2.57*, is valid for the purpose of detecting a target signal in receiver noise. For other roles, the radar range equation should be changed to more closely reflect the performance of the mission. For example, maximum detectable range increases with the square root of frequency in *eqn 2.57*. This is because the narrower beam concentrates more energy on the target and sees a smaller noise background. For search radar, a narrower beam would increase the signal to noise ratio but would also increase the time to search a given area - decreasing the area search rate. Thus the radar range equation for search must scale differently with frequency. In this section, we derive the radar range equation for a search radar with a circular aperture. Though a search radar need not have a circular aperture, the analysis is simpler and the same scaling relationships will result.

The footprint area illuminated by a circular aperture with a diffraction limited beam is

$$A_I = \left(\frac{R}{2}\frac{\lambda}{D}\right)^2 \pi k_A \qquad eqn\ 2.58$$

where R is the range to the footprint, D is the diameter of the aperture, λ is the wavelength of the transmitted signal, and k_A is the ellipse factor. The ellipse factor is the elliptical footprint's major to minor axis ratio (*Figure 2.12*):

$$k_A = \frac{R_E\left[\lambda(\varepsilon_{toe}) - \lambda(\varepsilon_{heel})\right]}{R_{tgt}\Delta\theta} \qquad eqn\ 2.59$$

where R_E is the radius of the Earth, λ is the central angle (*eqn 6.1*), R_{tgt} is the slant range to the target, and $\Delta\theta$ is the radar beamwidth.

$$k_A = \frac{A_{ellipse}}{A_{circle}} = \frac{a}{b}$$
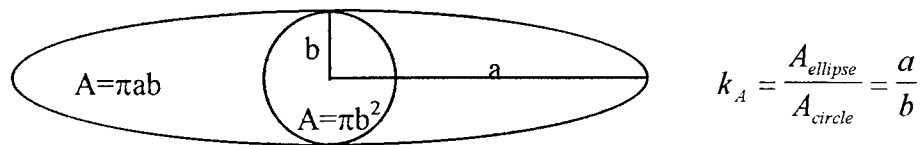
Figure 2.12: Ellipse factor definition.

The ellipse factor is a function of both orbital radius and frequency (*Figure 2.13*). As the radar altitude increases, the elliptical footprint becomes more eccentric - k_A increases - for a fixed target elevation because of the increased apparent curvature of the Earth's surface. As frequency increases, the beamwidth decreases and the footprint elongation due to Earth curvature is less significant which results in a decrease in k_A. This is only true at elevations for which the toe of the beam footprint does not exceed the maximum beam nadir offset coverage requirements. When part of the beam extends beyond the region of coverage, that portion of the beam is lost which decreases the ellipse factor. This is edge coverage degradation. Edge coverage degradation is more significant at lower frequencies since lower frequencies have larger beamwidths which encounter the edge of the viewing region before higher frequencies would. The average ellipse factor over orbital radius is plotted in *Figure 2.14*. The edge coverage degradation loss explains why lower frequencies have smaller ellipse factors in *Figure 2.14* but higher ellipse factors in *Figure 2.13*. At 20 deg target elevation (beam midpoint), none of the footprint extends beyond the region of coverage and there is no edge coverage degradation.
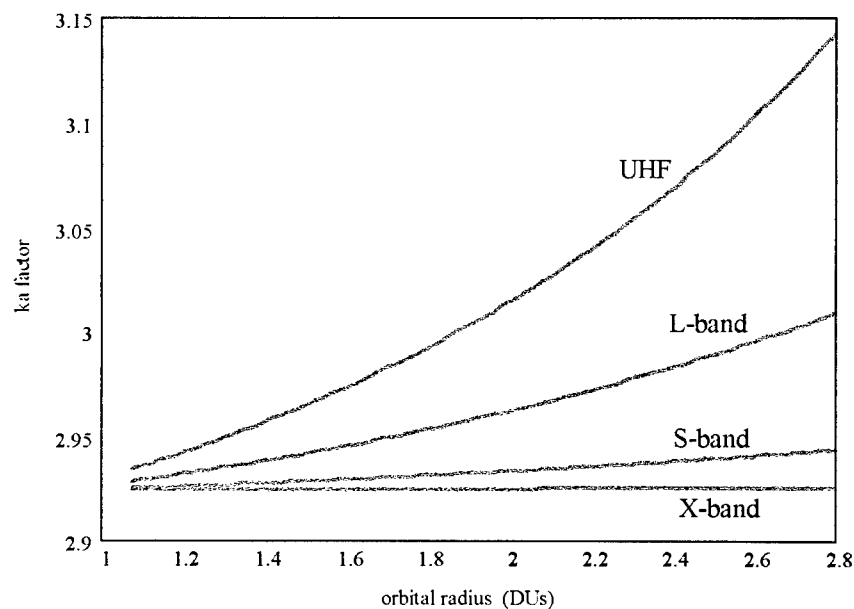
Figure 2.13: Ellipse factor, k_A, as a function of orbital radius and transmitter frequency for a fixed target (beam midpoint) elevation of 20 deg. A distance unit (DU) is equal to one Earth radii, or 6378 km.
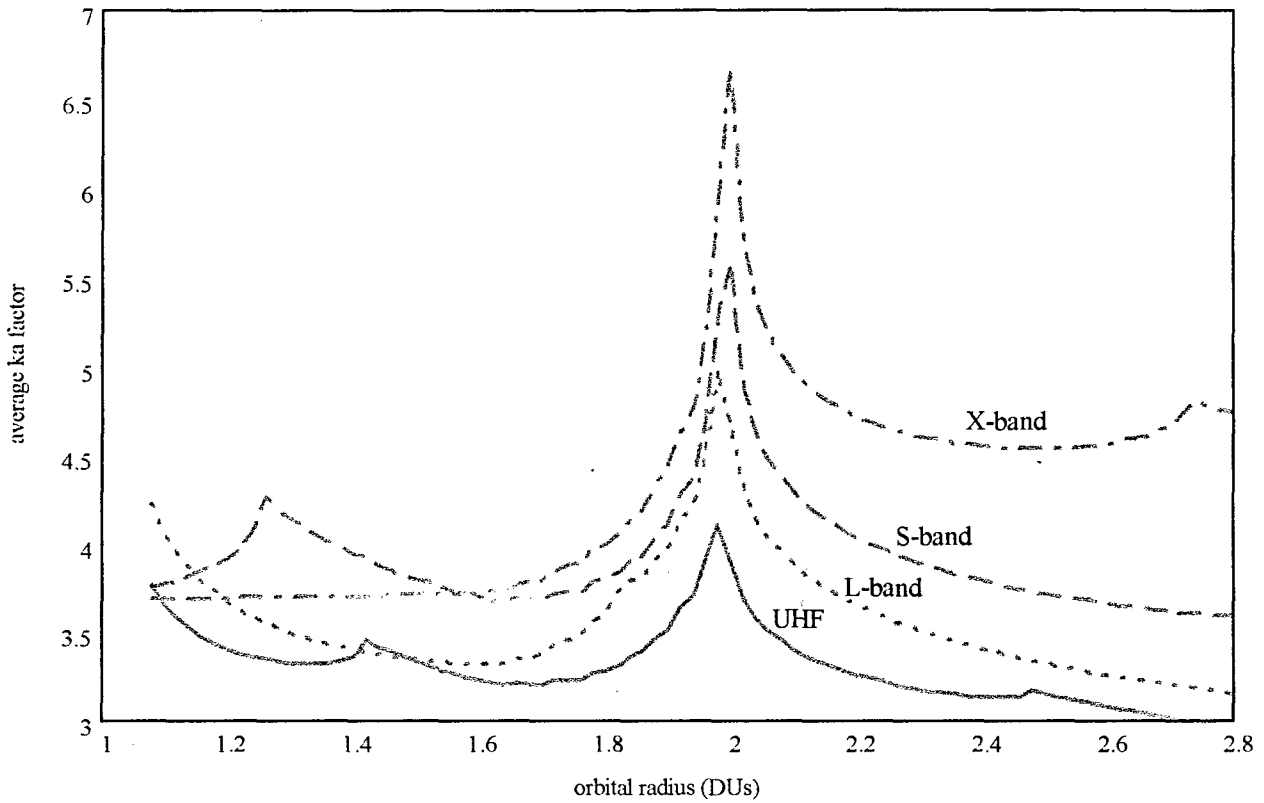
Figure 2.14: Ellipse factor weighted average over all coverage elevations as a function of orbital radius.

Dwell time is the time that it takes a point to completely transit the footprint area as the footprint is being scanned. Dwell time equals the product of the number of pulses incident upon the target and the time between pulses. That is

$$\Delta = nT_p = \frac{n}{prf}$$

eqn 2.60

The area search rate is the illuminated area divided by the dwell time.

$$ASR = \frac{A_I}{\Delta} = \frac{R^2 \lambda^2 k_A prf \pi^2}{4^2 n A_e}$$

eqn 2.61

The search radar range equation can be rewritten with area search rate as

$$\frac{S}{N} = \frac{P_{av} A_e \sigma \eta_{int} k_A \pi}{4^3 R^2 kTB_n \tau prf L} \frac{1}{ASR}$$

eqn 2.62

or in terms of range

$$R^2 = \frac{P_{av} A_e \sigma \eta_{int} k_A \pi}{4^3 kTB_n \tau prf L (S\ N)_1} \frac{1}{ASR}$$

eqn 2.63

The search radar range equation depends on range squared instead of range to the fourth

and on the power aperture product instead of the power-aperture-aperture product. The search radar range equation predicts performance only in terms of detection and area search rate. Other important factors, e.g. resolution, must be accounted for in other ways when doing system comparison. The search radar range equation can be expressed more compactly as

$$\frac{S}{N} ASR = P_{av} A_e K_s$$

eqn 2.64

where

$$K_s = \frac{\sigma \eta_{int} k_A \pi}{4^3 R^2 k T B_n \tau L}$$

eqn 2.65

2.5 Multipath and Channel Effects

Multipath is the result of receiving the desired signal through more than one path between the source and receiver. In most radar applications, only the signal due to the direct transmission path between the source and receiver is desired and the presence of multipath can severely limit the performance capability. The ground plane loss discussed in the derivation of the radar range equation is an example. Multipath is generally modeled as a random scattering function that results in reverberation. Since space based radar typically has a high target viewing aspect, multipath is typically only a problem for low grazing angles. More important for a space based radar is target scintillation due to channel effects. Since the transmission channel for a space based radar includes the ionosphere, irregularities in the ionization structure can cause severe signal scintillation [18]. Scintillation is a channel effect that can be analyzed in the same manner as multipath. The effects of multipath and channel effects are signal dependant and particularly troublesome since increasing signal gain simply increases the power of the undesired reverberation.

If the transmitted complex envelope is

$$\widetilde{s}_{tx}(t) = \sqrt{E_{tx}}\, \widetilde{f}(t)$$

eqn 2.66

where E_{tx} is the energy in the transmitted waveform and $\widetilde{f}(t)$ is the complex envelope of the transmitted waveform, then received complex envelope is

$$\widetilde{s}_{rx}(t) = \sqrt{E_{tx}}\, \widetilde{b}\, \widetilde{f}(t - t_d) e^{j2\pi f_d t}$$

eqn 2.67

where $\widetilde{b}$ is a random variable that accounts for the signal attenuation through the medium. $\widetilde{b}$ is a non-stationary, time varying process that can vary in both magnitude and phase. The variance of $\widetilde{b}$ is $\sigma_b{}^2$.

$$\widetilde{b} = |b| e^{j\phi_b}$$

eqn 2.68

Two important effects of random attenuation are doppler spreading, and range spreading.

40

Doppler spreading results in selective time fading of the signal and range spreading causes signal frequency fading. Both effects can occur simultaneously and result in range/doppler spreading.

Channel doppler spreading causes selective time fading of the signal and is characterized by variations of the signal amplitude in time. It is manifested by a loss of correlation of the signal when delayed by more than the reciprocal of the bandwidth of the channel spreading process. Doppler spreading most commonly results from atmospheric waves (internal waves), and relative target motion. Analogously, range spreading causes frequency fading and results from multipath and extended targets.. Range spreading is also only significant if the time bandwidth product is greater than one. Recall from Section 2.3 that large time bandwidth products were characteristics of waveforms with good resolution.

Change Change

42

3. Cost Model

The subsystem cost model lies at the heart of the space based radar system architecture design and optimization process. The cost metric is used to quantify the system cost of different proposed architectures for a given level of performance. The *optimal architecture* is the system architecture for which, at a given level of performance, the system cost to initial operating capability (IOC) is a minimum. For the case in which an architecture variable can be analytically defined throughout the valid range of the variable, the "optimization process" gives a global, optimal solution for that variable. An example of an architecture parameter that can be continuously and completely defined is orbital altitude. Many candidate architectures, however, are often discrete sets of different options. The architecture solution space is then the set of all unique combinations of architecture variables. The optimization process can only give the minimum cost system within this solution space. Thus, the architecture design and optimization process developed here is best used as an analysis tool for comparing proposed architectures. It is not a substitute for the process of generating a set of candidate architectures. Nevertheless, as a tool for *quantitatively* comparing the differences between systems, it is extremely powerful.

The cost metric can also be used to demonstrate how cost scales with performance. It is common for many commercial or scientific endeavors to demand the maximum *cost per performance* ratio. Cost per performance is very typically architecture dependant. Performance may be quantified as the expected revenue for commercial systems and as the volume, integrity, and/or resolution for a science mission. Cost, of course, is quantified by the metric. In contrast, most military systems are driven by requirements, not cost; though this mentality is beginning to change as budgets shrink. All systems, though, have minimum performance requirements. A significant change in the cost as a function of performance relationship warrants some interest if only to justify the need for requirements in the marginal return of performance region. For the Space Based Radar system, such a transition exists in the region of interest when search mission performance is measured by mean time to detection. Requirements justification is an important step in the system design process and quantifying cost as a function of performance is an invaluable tool.

The cost model is a parametric model – built around subsystem and performance parameters. This is one of its most powerful aspects as it allows a wide range of options and scenarios in both subsystems and performance to be studied. Because of the parametric nature, the model is also very flexible. The results presented in Chapter 8 have been continuously recomputed for adjusted parameter values following the review and advice of Air Force Space and Missile Center (SMC) [39], Air Force Phillips Lab [50], Lockheed Martin [27], and Lincoln Labs [49]. Through a process of briefing and review, we have settled on a set of system parameters that are acceptable to both the Air Force and industry. In some instances, we treat the parameters as free variables and determine the range of performance as a function of the expected or anticipated range of potential values. For some parameters, we derive a theoretical basis for

the value of the parameter. Ultimately, though, the model is flexible and the architecture optimization process can be repeated if knowledge of certain key parameters improves.

The parametric nature of the cost model can also be used to examine the cost sensitivity to various subsystem capabilities. The relative changes in subsystem parameters reflect technological advances. For example, power mass density is a function of battery storage capacity, depth of discharge, and conversion efficiency. Technological advances in batteries are reflected by a quantifiable change in the power mass density. The sensitivity of the overall system cost to advances in subsystems has important implications for funding research and development. Thus, the cost model may also be used to identify where likely improvements in system cost or cost per performance may be realized. Cost sensitivity analysis is presented in Section 8.4.

Distribution is introduced as an architecture variable in order to quantify the effects on system cost and performance and determine the optimum level of distribution. The benefits of satellite distribution have been attractive to mission planners for some time, but only recently is analysis being accomplished to seriously quantify the drivers and issues of space system distribution. In order to introduce distribution as an architecture variable in the cost model, the operations concept for using distribution had to be developed. (Chapter 5)

There are primarily two options for distributed radar operations: coherent and noncoherent distribution. Coherent distribution includes multistatic radar and the synthesis of sparse apertures from distributed elements. Coherent distribution has been the subject of several previous space based radar studies in the early 1980s [4,5,6]. From some preliminary analysis, we concluded that coherent operations were too operationally complicated. The number of coherent processing links between satellites increases geometrically with distribution and probably exceeds the envelope of feasibility for the time frame of a replacement to AWACS. Concepts of operations for noncoherent distribution did not exist so we developed a concept which distributes the function of aircraft detection (Chapter 5). We also developed a model for reliability and cost-of-reliability for a distributed radar system designed for detection (Chapter 4). Both the distributed operations concept and reliability model are important elements of the system cost model developed in this chapter. Much of the analysis we develop for analyzing a distributed space based radar should be generalizable to other *active* distributed systems.

For radar missions, there are typically two system functions: search and track. Although they often use the same system components, search and track missions are very different in function, operation, and performance metrics. An immediate example is the radar range equation which is directly proportional to the power-aperture-aperture product for track radar and only the power-aperture product for search radar. To consider an overall cost per performance metric for a system that functions in distinct ways, the system cost should be evaluated by performance in each of the function dimensions. Thus, it is necessary to formulate a cost-per-search-mission metric and a cost-per-track-mission metric and then combine them for an overall system cost/performance metric. One method for combination is to weight the two mission metrics depending on the relative time that the system functions in that mission. For the follow-on AWACS mission, the mission functions traditionally done by track radar may be accomplished

44

from a space based search radar. This is discussed in Section 5.1. Additionally, track function can be achieved from several different platforms including fighters and missiles. The search function can only be performed by the AWACS system. A cost metric that quantifies cost as a function of search mission performance captures the most significant attributes for a space based AWACS replacement. Thus, we concentrate on the search mission.

The cost metric is developed in this chapter. Section 3.1 is an overview of the system cost model which consists of several subsystem cost models derived from the USAF Unmanned Spacecraft Cost Model (USCM) and theoretical approximations. The power subcost model is developed in Section 3.2; the aperture subcost model is developed in Section 3.3; the bus subcost model is developed in Section 3.4; and the launch cost model is developed in Section 3.5. The system performance constraints developed in Chapter 5 are incorporated into the cost model with the power aperture optimality relationship derived in Section 3.6. The production learning curve model is developed in Section 3.7. The total integrated model including subsystem cost, reliability, distributed operations performance, and constellation coverage is summarized in Chapter 8, Section 8.2.

3.1 IOC Cost Metric

The cost metric is cost to initial operating capability (IOC). IOC costs include all development and acquisition costs up to and including the initial deployment costs. Thus operations costs, support costs, ground facility costs, etc., are not included in the metric. Although operating costs are important for life cycle system analysis, the baseline missions considered are assumed to have very similar operating costs. Operations costs in the metric would decrease the effect that different system parameters have on the overall cost and disguise the effects of system architecture. Of course, if different architectures have different operational requirements and cost, they must be considered during the comparison and modeling. For the SBR mission, cost of operations are assumed constant across the range of architectures.

The IOC cost is the cost to develop, acquire, and deploy the constellation of satellites. It is convenient to model the cost of each satellite as a function of the most significant subsystem cost drivers which are in turn functions of the system architecture and operations. The metric is then the product of the cost per satellite and the number of satellites in the constellation, N.

$$IOC\ cost = \phi(N)\ cost_{sat} = \phi(N)\left(cost_{power} + cost_{aperture} + cost_{bus} + cost_{launch}\right) \quad eqn\ 3.1$$

where $\phi(N)$ is the constellation size cost multiplier that includes learning curve effects. The four components of satellite cost are power, aperture, bus, and launch. These are the significant drivers in development, production, and deployment for the SBR system. Power and aperture costs are very strong functions of architecture. In the SMC Space Sensors Study [52], power and aperture were the constraints that pushed the limits of practicality. Launch is also a significant function of architecture through constellation size (number of launches) and altitude (size of the booster). Bus costs are included to capture the cost of all remaining subsystems.

45

Each subsystem cost is modeled as a power law ($y = c\,x^{\alpha}$) based on the USAF Unmanned Spacecraft Cost Model (USCM) [19,20,21] and theoretical approximations. Power laws are ubiquitous both in nature and human practices. Power laws are convenient in cost analysis since they scale without respect to dimension. This is important for two reasons. First, changing the units of the independent variable (x) will not change the shape of the power law. For example, expressing power in watts or ergs only changes the scaling factor (c) and not the power law exponent (α). Second, the scaling is independent of the absolute dimension – the power law scaling is valid regardless of the order of magnitude of the independent variable. For most real world power law realizations, there is a characteristic length for which the power law is valid. This characteristic length typically spans several order of magnitudes, however.

The USCM actually consists of two models: the research, development, test, and evaluation cost model (RDT&E), and the theoretical first unit cost model (TFU). RDT&E costs are also known as nonrecurring costs and quantify the development phase of the product life cycle. TFU costs are recurring costs and quantify the production phase of the life cycle. A rule of thumb is that development costs typically run 2 to 3 times the first unit production cost (TFU). This is generally true of the RDT&E and TFU models (*Figure 3.1*, *Figure 3.4,* and *Figure 3.6*), though the factor is usually better approximated as 3 to 4 times the first unit production cost. (The shape of this relation does not hold for the power subsystem TFU and RDT&E models – this is discussed in Section 3.2.) Both the RDT&E and TFU models are a set of cost estimating relationships (CERs) which are regressed from historical data (DMSP, DSCS3, FLTSAT-COM, GPS, I-IV, I-V, TDRSS, TACSAT, and others). Because the models are regressed from past systems, the CERs may not necessarily reflect modern trends or practices. Nevertheless, they have been widely used and reflect a standard unit that serves as a baseline. Additionally, we are primarily interested in the way that the system cost scales with respect to its key system parameters and mission drivers and not to absolute costs. It is very unlikely that the relative nature or shape of the CERs has changed significantly and it is the power law aspect of the models that primarily affects the scaling between architectures. For the search mission cost model, the TFU CERs are used to quantify subsystem costs while developments costs are modeled as a constant function of first unit production cost. This constant multiple for development costs represents another parametric degree of freedom in the cost model. The use of a constant development cost multiplier is done for model simplicity and is justified by the relative relationship between the RDT&E and TFU models. Thus, the USCM RDT&E model in not used.

As a parametric model, the cost metric is a function of several parameters; e.g. aperture area density, power density, battery capacity, solar array efficiency, T/R module mass, etc. Nominal values were selected to reflect current technology, technology that is expected in the near future, and technology that may be possible in several decades. The nominal values are based on theoretical approximations and frequent dialogues with the Air Force and industry [39,27,49,50]. The technique of a range of potential parameter values is used in sensitivity analysis to predict where significant cost reductions are likely. Subsequently, this provides the basis for recommendations for where the Air Force and industry should push the hardest in

developing new technology to make a space based replacement to AWACS both feasible and affordable.

The cost model is also a function of performance or capability. Thus, it is necessary to establish a baseline mission -- a set of mission requirements based on probable mission scenarios. The system sizing is very dependant on the scope of the mission and thus the scale of the system. The difficulty of selecting a mission baseline is that we must extrapolate current political, technological, and military trends at least 25 years into the future. For example, what is the likelihood of autonomous aircraft or missiles to perform the mission done currently by fighters? This, in turn drives the nature of target illumination which changes the priority of the SBR continuous tracking mission. The mission baseline selected is the same being used by the USAF Space and Missile Center development branch in their analysis. The mission baseline is summarized in Section 8.1.

The cost metric also incorporates mass production models that have not yet been proven in the satellite industry but that have been realized in several other industries including automobiles and computers. The uncertainty of realizing production learning in the early aircraft industry provides a historical analog to the beginnings of mass production in the satellite industry [22]. Until recently, satellites have been built in small numbers and with little standardization between satellites for different missions. The commercial satellite industry, most notably Hughes and Lockheed Martin, are changing this historical trend. Several current concepts and trends foretell testing of these assumptions in the next couple years -- notably, the number of large constellations telecommunications satellites, including Iridium and Teledesic. Lockheed Martin is reportedly realizing 15% learning curves in the production of Iridium buses [27]. Learning curves are developed in Section 3.7.

The number of satellites in the constellation, N, is a function of the constellation altitude; the level of distribution, n_s; and the number of spare satellites required for reliability/survivability. The altitude coverage function gives the constellation size, N, as a function of distribution and orbital altitude and is developed in the constellation design process (Chapter 6). The constellation design process also gives models for system duty cycle and maximum expected target slant range which are required to size the system.

3.2 Power Cost

The power subcost model is derived from the USCM and architecture effects. The power cost is

$$cost_{power} = \phi_P(P_t, \{C_i\}, \{R_i\})\left(\left[P_t(\rho_P P_t)\right]^{\alpha_P} + \beta_P\right) \qquad eqn\ 3.2$$

where $\phi_P(P_t, \{C_i\}, \{R_i\})$ is the power cost function which captures the effects of architecture. The other model component in parentheses is the USCM which has units of dollars. The USCM model is plotted in *Figure 3.1*. The power mass density, ρ_P, has units of kg/W and converts
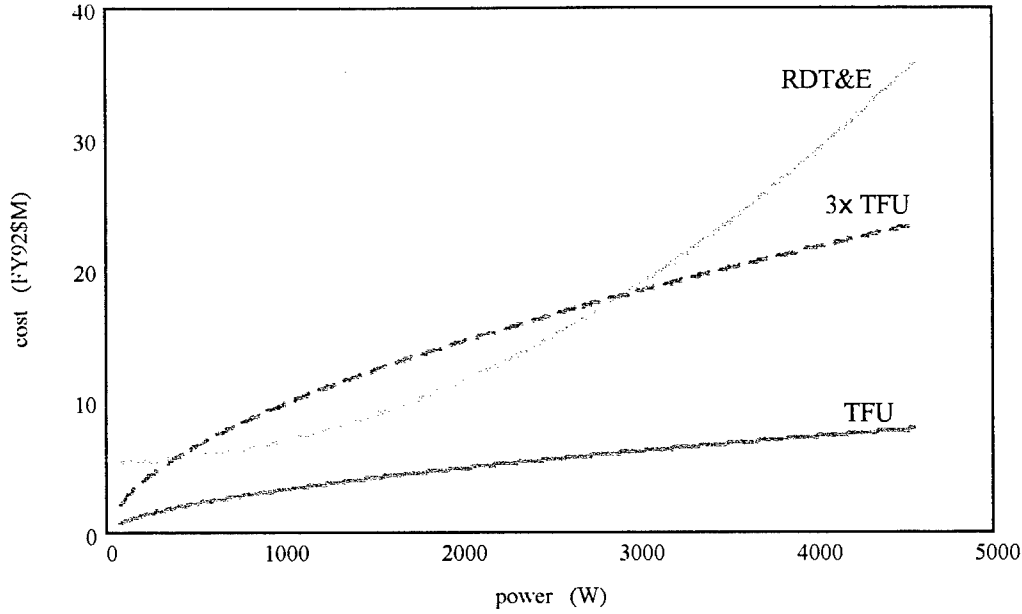
Figure 3.1: Power cost from USCM5. RDT&E and TFU CERs are shown. The ordinate spans the model's valid range. The dashed curve is 3 times the TFU.

power to mass. Thus, the quantity $(\rho_P P)$ is the total mass of the power system.

The power cost multiplier, ϕ_P, is a function of transmitted power, P_t; constellation parameters, $\{C_i\}$, (e.g. the number of satellites, N, and mean target distance); and reliability parameters, $\{R_i\}$. The power cost multiplier has the form

$$\phi_P = \phi_R\big(n_s,\{R_i\}\big)\phi_{P_0} + \phi_d(a) \qquad\qquad eqn\ 3.3$$

where ϕ_R is the reliability cost factor developed in Section 4.4, n_s is the level of distribution, and ϕ_{P0} is the TFU model parameter given in *Table 3.1*. The function $\phi_d(a)$ is the subsystem development factor which is a function of orbital altitude. a is the constellation semi-major axis.

A rule of thumb is for development costs to typically run two to three times the cost to produce the first unit. The USCM RDT&E CERs generally fall into a range of three to four times the TFU CER. The RDT&E power law, however, is significantly different from the TFU power law and the 3 times rule of thumb does not accurately predict RDT&E behavior. This is most likely due to the fact the historical systems forming the basis for USCM 5 were developed at a time when high power subsystems were uncommon and more costly to develop. Satellite power has been steadily increasing and 4.5 kW no longer represents a difficult power level to achieve on a satellite. Thus, the USCM 5 RDT&E probably does not accurately reflect the current development cost for power subsystems. This is another reason for using a constant first unit cost multiplier to characterize development cost. The development cost multiplier may be a function of orbital altitude because of the increased cost to develop hardened components for systems operating within the radiation belts.

Table 3.1: Power subcost model parameters [19]

source	α	φ₀ (FY92$M)	β
USCM5 RDT&E	.97	$1.08 \ 10^{-4}$	$4.9 \ 10^4$
USCM5 TFU	.29	.183	0

The power mass density is an important parameter in determining the subsystem cost and is a function of the power subsystem architecture. The subsystem architecture options are summarized in *Figure 3.2*. Power storage is required if the primary power source is solar. Currently, only batteries and fly-wheels are viable storage devices. Their capacities are summarized in *Table 3.1*. Because of storage and conversion losses, it may be desirable to directly power the transmitter from the primary source when the solar arrays are illuminated. This requires a separate power conditioning unit and steerable arrays which may significantly increase the power system mass. For the SBR mission, the solar arrays constitute less area and mass than the aperture – operationally, then, the satellite (and aperture) should be pointed at the target and the arrays steered normal to the sun. This will not, in general, be possible at all points throughout the orbit. Additionally, because the peak transmitted power often exceeds the instantaneous power that can be generated by the solar arrays, the transmitter power must be supplemented by the stored capacity. For these reasons and because it simplifies the analysis, we do not allow the primary power source to power the transmitter. This will result in a slight loss in subsystem efficiency due to storage conversion loss. This loss is typically on the order of only 5-10% which is roughly comparable to the loss in the power conditioning unit. If the primary power source is nuclear, no storage capacity is required and this problem is obviated. A nuclear power source also gives the possibility for high power at low specific cost relative to solar photovoltaic (*Table 3.2*). However, a nuclear power source is probably not politically viable.

Under the assumption that the stored power capacity must provide all the power to the
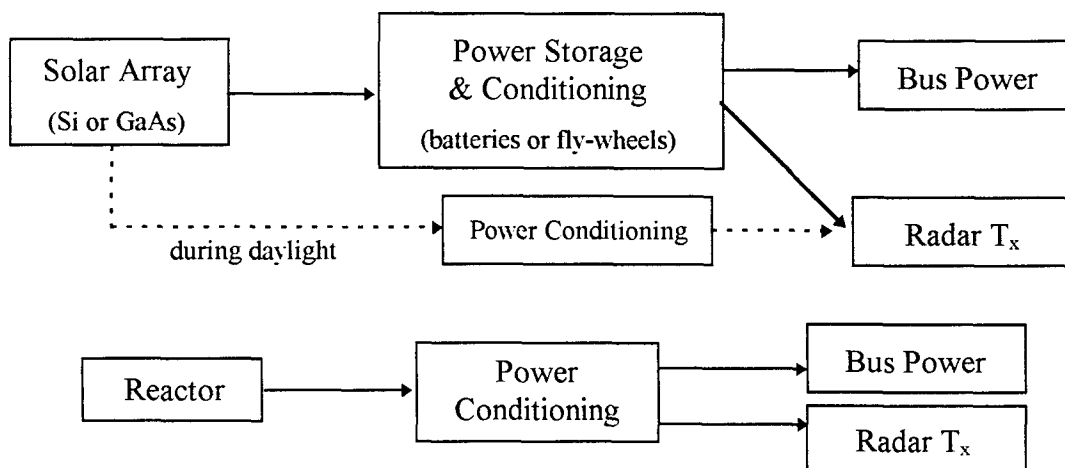
Figure 3.2: Power subsystem architecture options.

source	type	specific ratio	η	DOD	remarks
Si solar array	primary	26-100 W/kg	18%	N/A	.2-25 kW range; significant experience; high drag, easy target, 2500-3000 $/W
GaAs solar array	primary	26-100 W/kg	23%	N/A	more expensive than Si
nuclear reactor	primary	15-22 W/kg		N/A	25-100 kW range; very hardened but high IR signature; 400-700 $/W
NiCd battery	storage	25-30 Whr/kg	.9-.95	20-70%	wide experience, low DOD
NiH₂ battery	storage	25-60 Whr/kg	.9-.95	40-90%	common pressure vessel gives better specific capacity
NaS battery	storage	140-210 Whr/kg	.9-.95		under development
fly wheel	storage	60 Whr/kg	.9-.95	very high	storage only, better power density than currently available batteries, unproven

bus and payload, the required capacity is

$$C_{req} = \frac{P_{bus}Per + P_t\mu_{ms}Per}{\eta_{store}DOD}$$

eqn 3.4

where P_{bus} is the bus power, P_t is the average transmitted power, Per is the orbital period, μ_{ms} is the mission cycle fraction (the percentage of the orbital period for which the radar must operate), η_{store} is the storage conversion efficiency, and DOD is the depth of discharge. Many of these parameters are in turn dependant on other architecture parameters. Depth of discharge is a function of the type of storage device and the lifetime number of charge-discharge cycles. Mission cycle is a function of orbital altitude and operation concept. Transmitter power depends on the mission requirement for mean time to detect a target.

Because the power replenishment required from the solar arrays during one period is equal to the power discharge in a period, the required solar array power is

$$P_{SA} = \frac{DOD\, C_{req}}{(1-\mu_e)Per\,\eta_{SA}}$$

eqn 3.5

where the eclipse fraction, μ_e, must not only include the percentage of the orbital period spent in eclipse conditions, but the percentage of the orbit when the solar arrays are not illuminated because of operation coverage constraints. The eclipse fraction is a random variable that depends on the orbit, operation theatre, time of year, time of day, and operational requirements. For simplicity, we treat it as a constant that may be interpreted as the mean eclipse fraction.

The total power subsystem mass is the sum of the mass from the primary and storage components (where power conditioning and cabling are included with storage mass). The power mass density, ρ_P, is defined as the mass of the power system divided by the average transmitted power so that
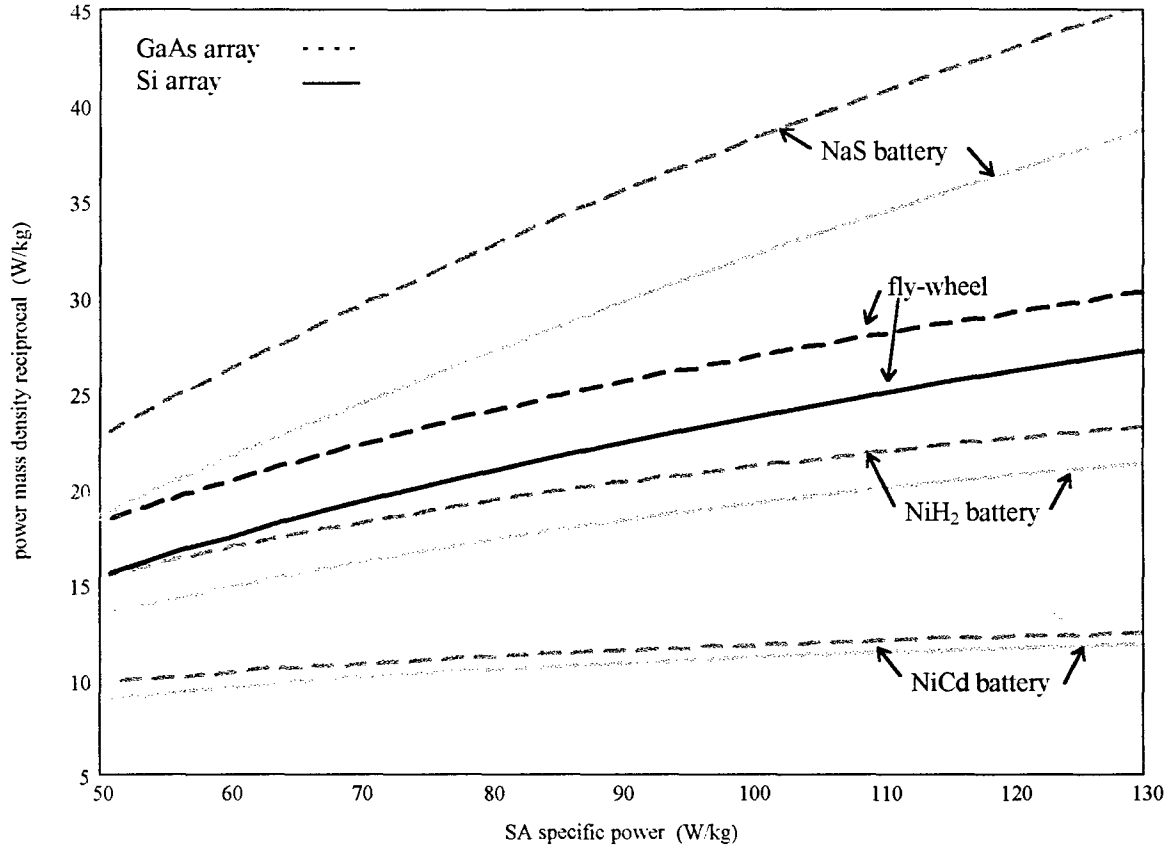
Figure 3.3: Reciprocal of power mass density as a function of solar array specific power and power system component options.

$$m_P = \rho_P P_t = \rho_{SA} P_{SA} + \rho_c C_{req} \qquad\qquad eqn\ 3.6$$

If the bus power is taken as 10% of the average transmitted power (on the order of 500 W to 1 kW), the power mass density parameter is

$$\rho_P = \frac{\rho_{SA}\left(1+\mu_{ms}\right)}{\left(1-\mu_e\right)\eta_{SA}\eta_{store}} + \frac{\rho_c\left(1+\mu_{ms}\right)Per}{\eta_{store}DOD} \qquad\qquad eqn\ 3.7$$

The power mass density parameter is plotted in *Figure 3.3* for the range of power system options given in *Table 3.2*.

3.3 Aperture Cost

As with the power subcost model, the aperture cost model is derived from the USCM and architecture factors. The aperture cost is

$$cost_{aperture} = \phi_A\left(A,\{C_i\},\{R_i\}\right)\left(\left[\rho_A A\right]^{\alpha_A} + \beta_A\right) \qquad\qquad eqn\ 3.8$$

source	α	φ₀ (FY92$M)	β
USCM5 RDT&E	.59	1.015	0
USCM5 TFU	.59	.230	$8.7\ 10^{-2}$

where the aperture cost multiplier, $\phi_A(A, \{C_i\}, \{R_i\})$, reflects constellation and reliability parameters. A is the effective antenna aperture[1] and ρ_A is the aperture mass density. The aperture cost multiplier has the form

$$\phi_A = \phi_R\left(n_s, \{R_i\}\right)\phi_{A_0} + \phi_d(a) \qquad\qquad eqn\ 3.9$$

where ϕ_R is again the reliability cost factor and ϕ_{A0} is given in Table 3.3.

The aperture mass density, ρ_A, has units of kg/m² and converts effective aperture area to mass. Thus, the quantity $(\rho_A A)$ is the total array mass. For a phased array antenna, the array mass depends on the mass of the individual transmit/receive or transceiver (T/R) modules and the structure that supports the modules. There are essentially two aperture options for the mission being considered: a parabolic antenna or a phased array. A parabolic antenna is in the shape of a parabola with a signal source at the focus. The antenna must be physically pointed to direct the transmitted signal towards the area of interest. Because of the aperture areas required for the

Figure 3.4: Aperture cost from USCM5. RDT&E and TFU CERs are shown. The dashed curve is 4 times the TFU. The ordinate spans the model's valid range.

[1] filled aperture, subarrays are not modeled

52

SBR AWACS mission and launch shroud constraints, the parabolic antenna would have to be deployable which results in surface accuracy errors. In contrast, a phased array can be made to conform to almost any shape. Because the beam is electronically steered by phase shifts in the radiating elements, the beam can be directed through a wide angle without moving the array and surface accuracy errors can be electronically processed. Since they are most commonly flat, a phased array can be unfolded for deployment. Additionally, a phased array is capable of generating as many simultaneous beams as it has elements. For these reasons, a phased array is more attractive than the parabolic antenna for the SBR AWACS mission. We assume that the array elements are separate T/R modules.

Table 3.4: T/R module mass examples [24,27]

year	manufacturer	freq (GHz)	T/R mass
70-74	TI	10	163 g
67-70	Westinghouse	9.5	450 g
72-73	RCA	1.2	31 g
80-83	Raytheon	1-12	2 g
82	MIT Lincoln Lab	29-31	.03 g
83	Raytheon	3.1	55 g
83-87	TI	9-11	50 g
97	Lockheed Martin	10	10 g

The antenna beam pattern is the Fourier transform of the aperture illumination function which spatially describes the signal across the aperture surface. To avoid grating lobes, which are the result of spatial aliasing, the radiating elements in a phased array cannot be greater than $\lambda/2$ apart where λ is the wavelength of the carrier signal. If the antenna has dimensions of l by w, the array must be filled with $2l/\lambda$ and $2w/\lambda$ elements along the sides. If the mass of each transceiver element, $m_{T/R}$, includes the structure mass distributed among the elements, the array mass is

$$m_A = m_{T/R} \frac{4f^2 A}{c^2}$$

eqn 3.10

where f is the transmitter carrier frequency and c is the speed of light. The aperture mass density parameter is defined as

$$\rho_A = \frac{4m_{T/R}f^2}{c^2}$$

eqn 3.11

The aperture mass density parameter is plotted for a range of frequencies in *Figure 3.5*. *Table 3.4* includes examples of typical T/R module mass. Radiation hardening is an important factor in the module mass. Note that the values given in *Table 3.4* do not include the support structure mass that must be included when using *eqn 3.11*.
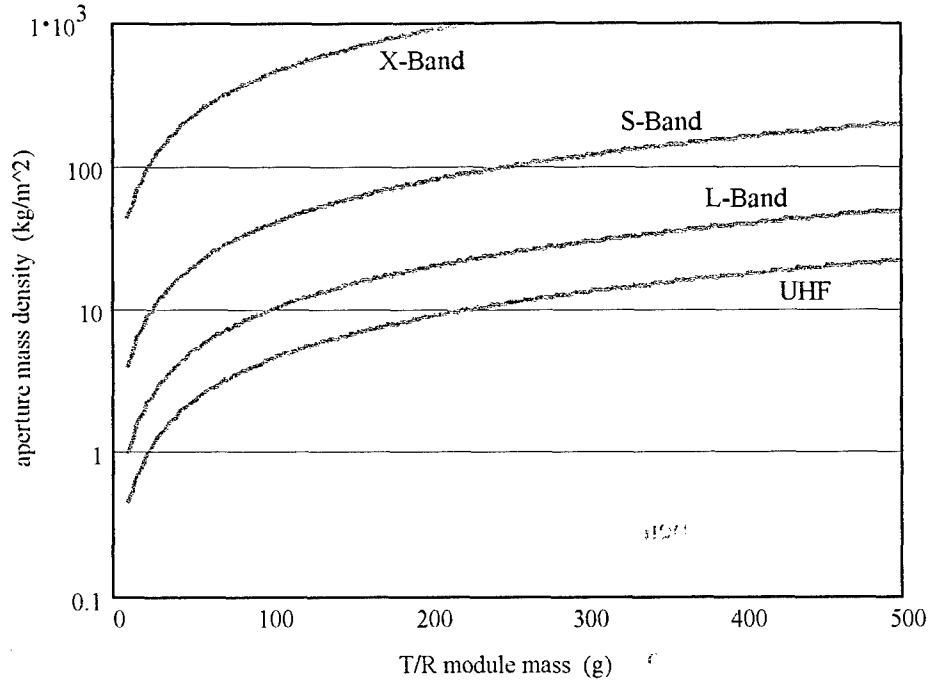
Figure 3.5: Aperture mass density parameter as a function of T/R module mass and frequency.

3.4 Bus Cost

The most significant differences in cost due to architecture result from power, aperture, launch, and processing. Therefore, the cost of all other satellite subsystems and components (structure, thermal, attitude control, telemetry and communication) are combined and reflected in the bus subcost. The bus subcost model is derived from the USCM model and architecture factors. The subcost model is

$$cost_{bus} = \phi_{bus}\left(\{C_i\},\{R_i\}\right)\left(\left[m_{bus}\left(P_t, A, \{C_i\}\right)\left(1 + \mu_{hard}\phi_{hard}(a)\right)\right]^{\alpha_{bus}} + \beta_{bus}\right) \qquad eqn\ 3.12$$

where the bus cost multiplier, $\phi_{bus}(\{C_i\}, \{R_i\})$, reflects constellation and reliability parameters. P_t is the transmitter power and A is the effective antenna aperture. The bus mass, m_{bus}, is a function of transmitter power, aperture, and constellation parameters that drive required spacecraft hardening. The aperture cost multiplier has the form

$$\phi_{bus} = \phi_R\left(n_s, \{R_i\}\right)\phi_{bus_0} + \phi_d(a) \qquad eqn\ 3.13$$

where ϕ_R is again the reliability cost factor and ϕ_{bus0} is given in *Table 3.3*.

The bus mass function is a simple parametric relationship that depends on the power and aperture mass and an assumed payload mass fraction, μ_{PA}. The power-aperture mass fraction is

$$\mu_{PA} = \frac{\rho_P P_t + \rho_A A}{m_{S/C}} \qquad eqn\ 3.14$$

54

Table 3.5: Bus subcost parameters [19]

source	α	ϕ_0 (FY92$M)	β
USCM5 RDT&E	1	.110	148
USCM5 TFU	.77	.185	0

Since the bus includes everything that is not payload, the mass of the bus is

$$m_{bus} = \frac{1 - \mu_{PA}}{\mu_{PA}} \left(\rho_P P_t + \rho_A A \right)$$

eqn 3.15

The higher orbital altitudes considered in the range constellation architecture options lie within the Van Allen radiation belts. Spacecraft operating within these belts require radiation hardening. The radiation belts consist of electrons and ions trapped by the Earth's magnetosphere and vary in strength (measured by the number and energy of trapped particles) with altitude. The highest particle flux occurs in two toroidal regions centered at approximately 1.3 and 5 Earth radii. The National Space Science Data Center (NSSDC) maintains several models of the Van Allen belts which can be accessed via the Internet [25]. Component protection can be achieved in several ways -- one of the simplest is to place vulnerable components behind a protective metal shield that blocks particles. The NSSDC maintains a model called SHIELDOSE which gives the radiation dose absorbed for different shield materials and thicknesses [26]. SHIELDOSE was used to determine the aluminum (Al) shield thickness required to give a dose rate of 10^4 rads/yr as a function of orbital altitude. A dose of 10^4 rads/yr is considered high since many unhardened
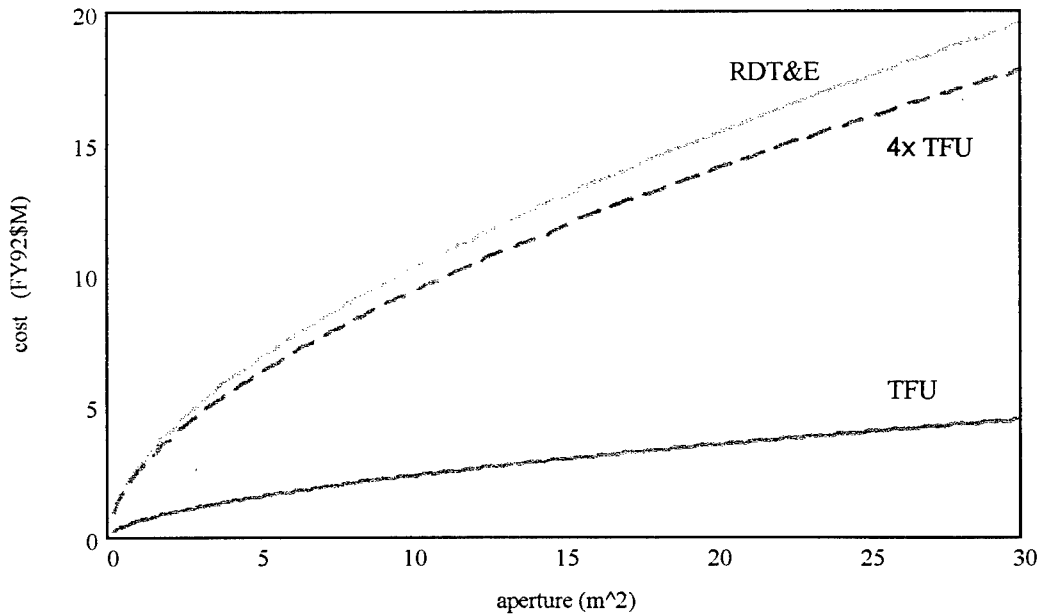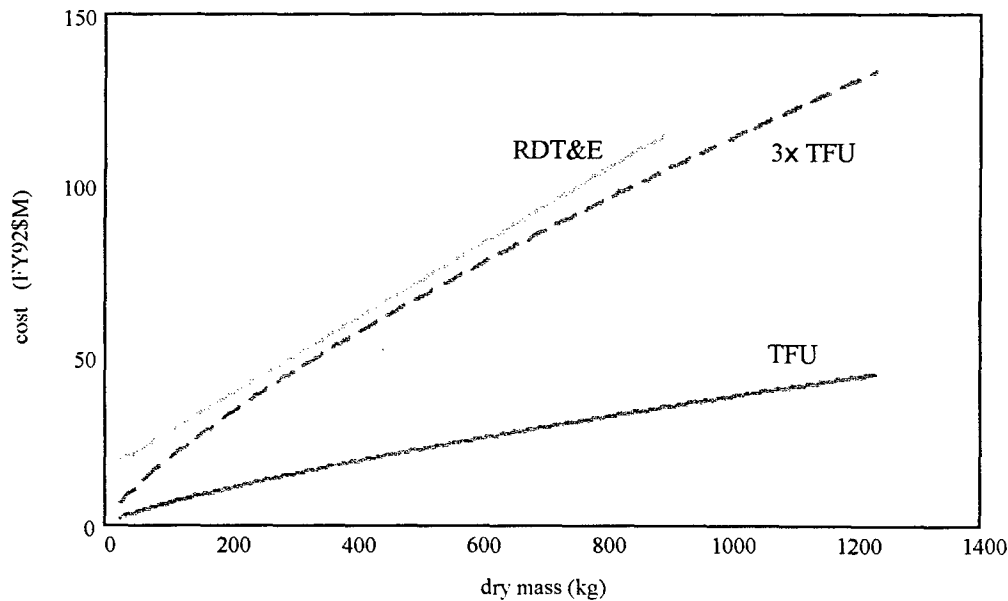

Figure 3.6: Bus cost from USCM5. RDT&E and TFU CERs are shown. The dashed curve is 3 times the TFU. The ordinate spans the model's valid range.
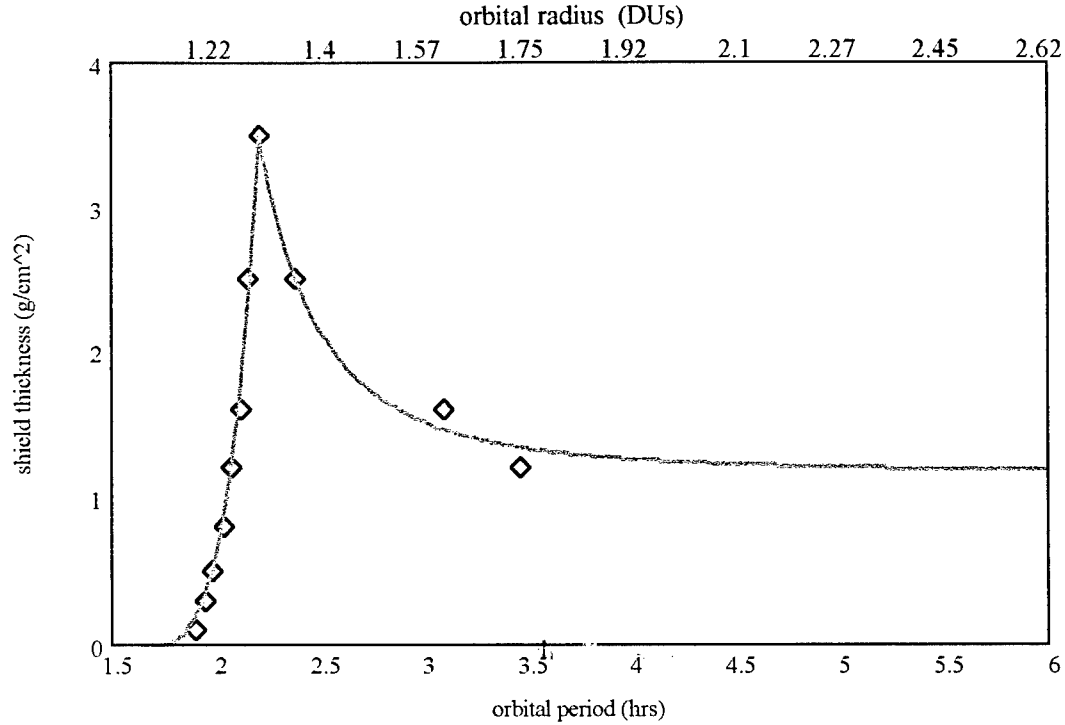
Figure 3.7: Hardening required (AL shield thickness) to give a dose rate of 10^4 rads/yr as a function of orbital period.

components can survive total doses of only 10^4 rads. However, because of the parametric nature in which hardening/shielding is modeled below, the total dose is not important in the model. What is important is the shape of the shield thickness results plotted in *Figure 3.7*. Decreasing the total dose requirement will shift the entire shield thickness curve up, but will not significantly alter the shape of the curve. Since hardening mass is modeled as a percentage of bus mass and peak hardening, changing the dose rate does not affect the nature of the results. This is an example of one of many ways in which parametric modeling allows analysis that would not otherwise be possible without extensive complex modeling and simulation.

The best fit power law regressed to the data in *Figure 3.7* is

$$thick_{Al}(a) = \begin{cases} -.035 + 1949(a-1)^6 & 1.1 < a \le 1.349\ DU \\ 1.155 + .1(a-1)^{-3} & 1.349 < a < 2.6\ DU \end{cases} \qquad eqn\ 3.16$$

where a is the semi-major axis expressed in DUs (1 DU = 6378 km) and $thick_{Al}$ is the aluminum shield thickness in g/cm^2.

Required shielding may be as much as 20% of the spacecraft mass [27]. Using this as a maximum gives a simple method to bootstrap shielding mass into bus mass. If the shielding to spacecraft mass ratio, μ_{hard}, peaks at 20% at the inner radiation belt maximum (1.3 DUs), the hardened bus mass is then

$$m_{bus}^{(hard)} = \left(1 + \mu_{hard}\phi_{hard}(a)\right)m_{bus} \qquad eqn\ 3.17$$

where the hardness multiplier, ϕ_{hard}, is

$$\phi_{hard}(a) = \frac{thick_{Al}(a)}{thick_{Al}(1.3)} \qquad eqn\ 3.18$$

Spacecraft hardening may be desirable for reasons other than the radiation belts. Likely future threats to military space systems include directed energy weapons, directed particle weapons, kinetic kill devices, and other anti-satellite means. It should be possible to develop shielding that gives radiation hardening as well as protection against hostile enemy actions.

The total spacecraft dry mass (without station keeping and attitude control propellant) is then

$$m_{S/C}^{(dry)} = m_{bus}^{(hard)} + \rho_P P_t + \rho_A A \qquad eqn\ 3.19$$

Some of the low earth orbits considered in the range of constellation altitudes experience atmospheric drag. The satellites must periodically thrust to compensate for the orbital energy lost to atmospheric drag. The propellant required for a lifetime of drag delta V's increases the mass of the spacecraft. Since only the propellant required to compensate for drag is important in system architecture comparison - attitude control propellant is required regardless of orbital altitude - we define the spacecraft wet mass as

$$m_{S/C}^{(wet)} = \exp\left(\frac{\Delta V_{life}}{gI_{sp}}\right)m_{S/C}^{(dry)} \qquad eqn\ 3.20$$

eqn 3.20 is an expression of the rocket equation (eqn 3.26). I_{sp} is the specific impulse of the thrusters used for drag compensation, and ΔV_{life} is the total delta V for the life of the satellite. Since the delta V to compensate for the drag of one orbit is

$$\Delta V_{rev} = \pi \frac{\rho_{atm}}{bc} a\sqrt{\frac{\mu}{a}} \qquad eqn\ 3.21$$

where bc is the satellite ballistic coefficient, a is the orbit semimajor axis, ρ_{atm} is the atmospheric density, and μ is the gravitational constant (3.986 10^5 km^3/sec^2). The lifetime delta V is then

$$\Delta V_{life} = \Delta V_{rev} \frac{T_{life}}{2\pi}\sqrt{\frac{\mu}{a^3}} \qquad eqn\ 3.22$$

A convenient expression for the atmospheric density during solar maximum for the altitudes of interest is

$$\rho_{atm} = \exp\left(-38.69 + 22.27a^{-8}\right)\ \frac{kg}{m^3} \qquad eqn\ 3.23$$

where a must be expressed in DUs. eqn 3.23 is regressed from solar maximum atmospheric density data given in [53] and is valid for altitudes from 500 to 1500 km. Above 1500 km

altitude, the delta V required to overcome drag is negligible.

3.5 Launch Cost

Launch costs can frequently require from one quarter to as much as one half of the total cost of a space system. The launch cost gradients due to system architecture can be significant and are thus an important aspect of architecture design and optimization. Launch costs depend on spacecraft mass, distribution (shared launch, separate launch), reliability, and constellation (altitude, number of satellites, inclination). The launch subcost model is the product of architecture factors and the best fit power law to the current launch capabilities plotted in *Figure 3.8*.

$$cost_{launch} = \phi_l\left(m_{S/C}^{(wet)}, \{C_i\}, \{R_i\}\right)\left(\left[\phi_{boost}(a)\left(m_{S/C}^{(wet)}\right)\right]^{\gamma_{a_l}} + \beta_l\right) \qquad eqn\ 3.24$$

where the launch cost multiplier ϕ_l is a function of vehicle mass, constellation and reliability parameters. The launch cost multiplier has the form

$$\phi_l = \phi_R\left(n_s, \{R_i\}\right)\phi_{l_0} \qquad eqn\ 3.25$$

where ϕ_R is again the reliability cost factor and ϕ_{l0} is given in Table 3.6.

The high cost of access to space always receives a lot of attention from those seeking ways to make space more affordable. Significant savings in system cost would result from even moderate reductions in launch costs. Current launch system costs range from $7000-20,000 per
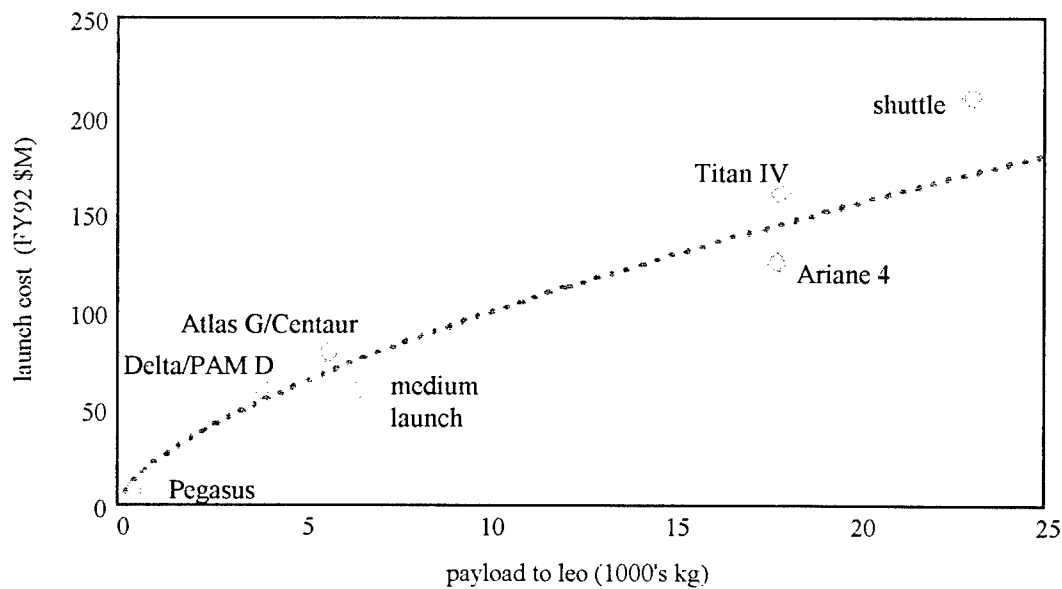
Figure 3.8: Launch cost data as a function of leo payload and best fit power law relationship. The shuttle data point was not included for purposes of regression. [21]

Table 3.6: Launch subcost parameters

source	α	ϕ_0 (FY92$M)	β
regression	.6	.415	-13.76

kg. The stated goal of the Reusable Launch Vehicle (RLV) is a launch cost of $1000 per pound, or approximately $2200 per kg.

eqn 3.24 treats booster capability as a continuum. In reality, launch capabilities are very discrete and a lifter may not always exist in the desired payload mass range. However, it is the margin of return nature of the power law in *Figure 3.8* that is important and not the actual launch vehicle. It is more efficient to launch heavy payloads than to launch the same total mass on several smaller boosters. This will most likely continue to be true for the next generation of launch vehicles. The reason for modeling launch cost as a continuum is to account for payload options that are currently voids but may come into existence with future launch systems.

The power law in *Figure 3.8* gives the cost to launch a payload into low earth orbit (LEO). The booster required to lift the spacecraft from LEO into its final orbit increases the mass that must be launched into LEO. The rocket equation relates the change in mass and the change in velocity.

$$\Delta V = g I_{sp} \ln\left(\frac{m_0}{m_f}\right)$$

eqn 3.26

where g is the gravitational constant, I_{sp} is the specific impulse, m_0 is the mass before the burn,
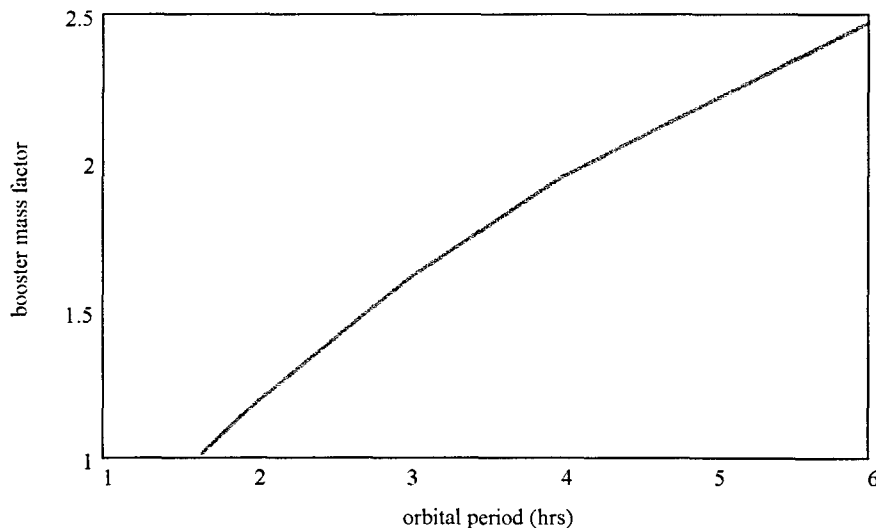
Figure 3.9: Booster mass factor to go from LEO to final orbital altitude expressed as a function of orbital period. LEO is taken to be 1.1 DU. [eqn 3.28]

59

and m_f is the mass following the burn. For a Hohmann transfer between circular orbits with semi-major axes a_1 and a_2, the required ΔV is

$$\Delta V = \sqrt{\frac{\mu}{a_2}} - \sqrt{\frac{\mu}{a_1}} + \sqrt{\frac{2\mu}{a_1} - \frac{2\mu}{a_1 + a_2}} - \sqrt{\frac{2\mu}{a_2} - \frac{2\mu}{a_1 + a_2}} \qquad eqn\ 3.27$$

where μ is the gravitational constant for Earth.

The booster mass factor as a function of the final orbit semi-major axis is then

$$\phi_{boost}(a) = \exp\left[\left(\sqrt{\frac{\mu}{a}} - \sqrt{\frac{\mu}{a_{leo}}} + \sqrt{\frac{2\mu}{a_{leo}} - \frac{2\mu}{a_{leo} + a}} - \sqrt{\frac{2\mu}{a} - \frac{2\mu}{a_{leo} + a}}\right)\frac{1}{gI_{sp}}\right] \qquad eqn\ 3.28$$

For the cost metric analysis below, a_{leo} is taken as 1.1 DU and the booster I_{sp} is 300 sec.

There are two strategies for sizing launch costs when modeling distributed system architectures. The conservative approach is to require a separate launch for each satellite. The more reasonable approach and one which slightly pushes distribution is to allow multiple satellites per launch. The difference in total system cost only becomes significant ($\geq 5\%$) for configurations of more than five satellites (*Figure 3.10*). Since shared launches require similar orbits, a good model is to allow the entire distributed configuration to be launched on a single launch vehicle. For large configurations, this is probably not practical, but shared launches of subsets of the
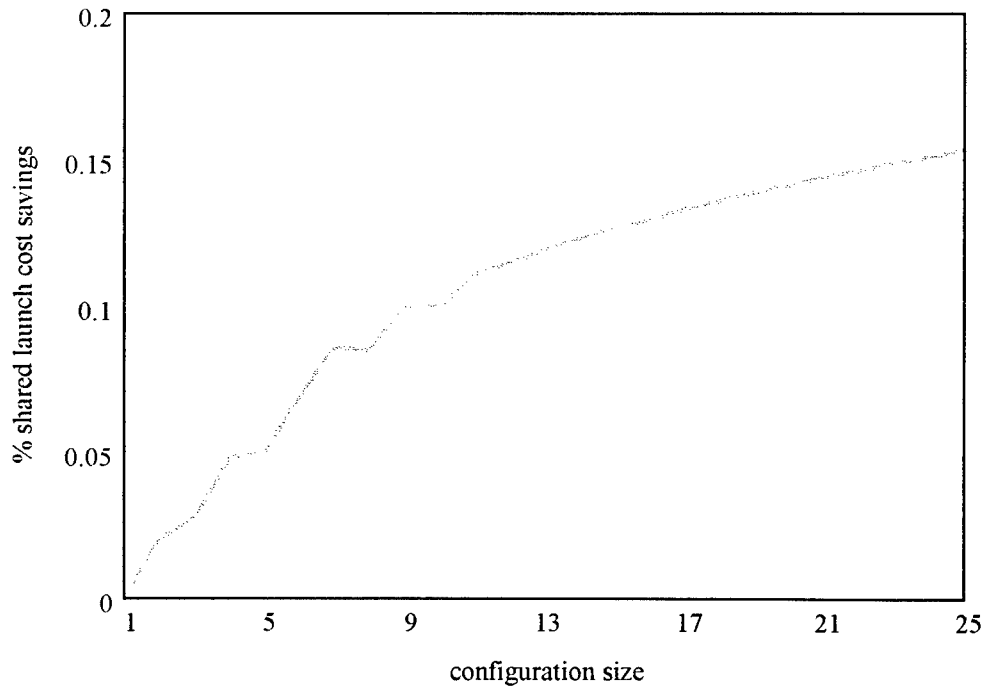
Figure 3.10: Shared launch cost savings as a function of configuration size (distribution). The entire configuration must be launched on the same booster which may not be practical for very large configurations. (This plot was created from results that are presented in Chapter 8.)

60

configuration will still result in some cost reduction. The realizable number of spacecraft per booster depends on several other factors as well: booster capability, orbit, and spacecraft mass.

3.6 Power Aperture Constraints and Optimality Condition

The search mission is constrained by several minimum performance requirements. These include detection probability, mean time to detection, area search rate, and minimum detectable velocity. Detectable velocity is constrained by the clutter return, waveform, and signal processing and is treated in Chapter 7. The other mission requirements are all functions of the signal to thermal noise ratio. Probability of detection, mean time to detection, and area search rate only represent two degrees of freedom (Section 5.3). The requirements are embedded in the cost function through area search rate, ASR, and signal to noise ratio, S/N. The radar range equation ($eqn\ 2.63$) further couples the mission requirements to a single constraint given by the product of area search rate and signal to noise ratio:

$$ASR\ S/N = K_s A P_t \qquad\qquad eqn\ 3.29$$

where K_s is a proportionality factor given by ($eqn\ 2.65$)

$$K_s = \frac{\sigma \pi k_A}{4^3 R_t^2 L k T_{rx}} \qquad\qquad eqn\ 3.30$$

and P_t is the average transmitted power.

$eqn\ 3.29$ gives hyperbolic solutions in power or aperture. Because the mission requirements represent only one constraint in an equation with two degrees of freedom (power and aperture), we are free to optimize the power aperture relationship. Performance and cost are two obvious metrics by which to optimize the power aperture relationship. In turn, performance may be quantified by two metrics: mean time to detection and minimum detectable velocity. Mean time to detection is an established requirement and embedded in the area search rate constraint. Using minimum detectable velocity as a performance metric for optimization leads to the ridiculous solution of infinite aperture and no power. Any other solution is suboptimal when measured by minimum detectable velocity. Cost is a better metric for optimizing the power aperture relationship, particularly since cost is the metric for optimization of the system architecture.

Define the payload cost as

$$cost = cost_P + cost_A = \phi_P\left(\left[\rho_P P^2\right]^{\alpha_P} + \beta_P\right) + \phi_A\left(\left[\rho_A A\right]^{\alpha_A} + \beta_A\right) \qquad\qquad eqn\ 3.31$$

The power-aperture constraint, $eqn\ 3.29$, may be expressed simply as $PA = K$. Substituting this constraint into the payload cost model gives

$$cost = \phi_P \rho_P^{\alpha_P} P^{\alpha_P} + \phi_A \rho_A^{\alpha_A} \frac{K^{\alpha_A}}{P^{\alpha_A}} + \phi_P \beta_P + \phi_A \beta_A \qquad\qquad eqn\ 3.32$$

which is now only a function of power, P. The minimum cost occurs where the first derivative is zero. That is

$$\frac{\partial}{\partial P} cost = 2\alpha_P \phi_P \rho_P^{\alpha_P} P^{2\alpha_P - 1} - \alpha_A \phi_A \rho_A^{\alpha_A} K^{\alpha_A} \frac{1}{P^{\alpha_A + 1}} = 0 \qquad eqn\ 3.33$$

which can be solved for the optimal power, P^*

$$P^* = \left(\frac{\alpha_A \phi_A \rho_A^{\alpha_A}}{2\alpha_P \phi_P \rho_P^{\alpha_P}}\right)^{\frac{1}{2\alpha_P + \alpha_A}} \left(\frac{ASR\ S\ N}{K_s}\right)^{\frac{\alpha_A}{2\alpha_P + \alpha_A}} \qquad eqn\ 3.34$$

From the power-aperture constraint, the optimal aperture, A^* is then

$$A^* = \left(\frac{2\alpha_P \phi_P \rho_P^{\alpha_P}}{\alpha_A \phi_A \rho_A^{\alpha_A}}\right)^{\frac{1}{2\alpha_P + \alpha_A}} \left(\frac{ASR\ S\ N}{K_s}\right)^{\frac{2\alpha_P}{2\alpha_P + \alpha_A}} \qquad e_{q}\ 3.35$$

The power to aperture cost ratio is then

$$\frac{cost_P}{cost_A} = \frac{\phi_P \rho_P^{\alpha_P} P^{*2\alpha_P} + \phi_P \beta_P}{\phi_A \rho_A^{\alpha_A} P^{*2\alpha_A} + \phi_A \beta_A} \qquad eqn\ 3.36$$

If the cost bias parameters, β, are small (as they are for the TFU model), the cost function is then in the form of a homogenous power law and the cost ratio simplifies to

$$\frac{cost_P}{cost_A} = \frac{\alpha_A}{2\alpha_P} \qquad eqn\ 3.37$$

For the TFU model, the ratio $\alpha_A/2\alpha_P$ is approximately one which implies that the optimal power aperture solution to *eqn 3.29* is when the power and aperture both contribute equally to the payload cost. In general, this *optimality condition* is given by *eqn 3.36*. The optimality condition is useful as a rule of thumb to quickly compare the optimality of the power and aperture sizing of different system proposals.

3.7 Learning Curve Model

Learning curves in production have been identified in several industries including aircraft, automobile, and computer manufacturing. The learning curve represents improvements and experience in the manufacturing process that reduce the cost of late production units. Because most satellite systems have historically been unique and were only built in small production runs, learning curves are as yet not fully realized in the space industry. This is beginning to change, mostly because of the efforts of commercial practices in the booming satellites-for-telecommunications market. Part of the strategy of the proposed Teledesic system is the averaged cost savings resulting from the production of over 800 satellites. More modest systems should also realize significant learning curves including Iridium (66 satellites), Orbcomm (46 satellites),
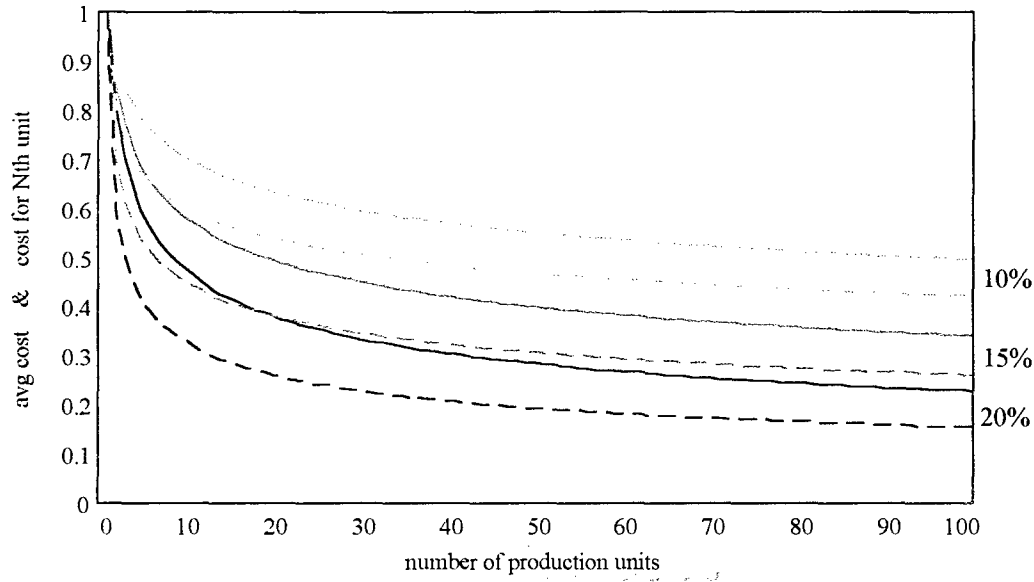
Figure 3.11: Learning curve showing average cost (solid) and production cost for the Nth unit (dashed) for discount factors: 10%, 15%, and 20%.

Spaceway (17 satellites), and Astrolink (9 satellites) [28]. Lockheed Martin has already reported realizing 15% learning curves in production of the buses for the Iridium system [27].

A learning curve is defined by the discount in the cumulative average when the number of production units are doubled. The average cost of N production units is then

$$\langle c_N \rangle = N^{\log_2 (1-dis)} c_1 \qquad eqn\ 3.38$$

where c_1 is the first unit cost, dis, is the discount rate (typically ranging from 5% to 20%), and c_N is the Nth unit cost. The operator $\langle \cdot \rangle$ gives the cumulated average; that is, the sum of the cost of all units through N divided by N. *Figure 3.11* demonstrates both average production cost and N^{th} unit cost.

Learning curves are reflected in the constellation size cost multiplier in *eqn 3.1*.

$$\phi_N = \left[acf(a, n_s) \right]^{1+\ln(1-dis)\ \ln(2)} \qquad eqn\ 3.39$$

where acf is the altitude coverage function derived in Section 6. The altitude coverage function gives the constellation size required to meet coverage requirements as a function of orbital radius and distribution.

A common strategy for learning curve implementation is to increase the discount rate as the number of production units increases. Wong [21] recommends 5% for less than 10 units, 10% for between 10 and 50 units, and 15% for over 50 units. Higher learning curves may as yet be realized for the space industry. A significant cost driver in satellite manufacturing is the man-hours required to test and evaluate the final satellite. There is currently much work being done to automate this process which would result in higher learning curves for large production runs. The

cost of developing the autonomous evaluation software would be prohibitive for small production runs unless some inter-application could be realized between different products. However, for large production runs, the cost savings in man-hours would be very significant.

4. Reliability Model for Distributed Systems

Reliability is a significant advantage of a distributed architecture. Distribution in computers and processing is a well practiced means for achieving reliability in networks. The redundant flight computers on modern aircraft use distributed algorithms to increase system reliability. Distribution can be achieved at several levels and diminishes the effect a failure at a single point or node has on the system. The tradeoff of distribution is that more elements are required in the system.

For SBR, a distributed operations concept is employed to reduce the power and aperture required on a satellite (Chapter 5). In this context, the distribution is at the level of the satellite. More specifically, mission function is distributed between several satellites so that no single satellite is solely responsible for performing the mission. One of the most significant advantages of system distribution is increased reliability. This is reflected by a decreased cost of reliability. A reliability cost model is developed in Section 4.4.

Reliability in the context of the SBR mission should include satellite and subsystem reliability, performance reliability, and system survivability. Satellite and subsystem reliability is the complement of the probability of failure of a single satellite. Equivalently, it is the cumulative probability that the satellite has not yet failed. It should be noted that a subsystem failure does not always imply satellite failure. Another advantage of distribution is the possibility that satellites with different subsystem failures may be able to cooperate and maintain some system performance. For example, if one satellite suffers a power failure that precludes operating as a radar transmitter, and another satellite with good power loses processing capability, the two satellites together may still give the functionality of one good satellite. For simplicity, subsystem failures are included in satellite failure and are not distinct in the reliability model. Performance reliability is a measure of the probability that the SBR constellation can satisfy minimum performance requirements at any given time. Performance reliability is similar to the concept of system availability. Performance failure is distinct from satellite failure – a missed link is a performance failure when the satellite continues to operate. Finally, system survivability is the probability of the system continuing to perform following hostile enemy action. Survivability includes vulnerability to both anti-satellite devices and electronic countermeasures (jamming).

This Chapter models the distributed system reliability and the cost associated with that reliability and solves for the optimum point. Section 4.1 defines the key reliability parameters and includes an overview of the reliability model. The reliability required for distribution is developed in Section 4.2 and the cost of reliability is developed in Section 4.3. The reliability cost factor is derived in Section 4.4.

4.1 Reliability Model Overview and Example

Mission reliability is the cumulative probability that the system has not failed. Since a

Table 4.1: Reliability parameters.

parameter		model value
F_{ms}	instantaneous mission probability of failure	.01
R_{ms}	instantaneous mission reliability/availability	.99
R_0	individual satellite reliability for a m-out-of-n redundant configuration	*eqn 4.4*
P_c	configuration probability of detection	*eqn 5.10; .95, .75*
P_d	individual radar probability of detection	*eqn 4.2*
P_e	minimum acceptable mission effectiveness	*eqn 4.3; .9, .85*
$1-P_e$	maximum acceptable performance degradation	*eqn 4.3; .9, .85*
R_{sc}	baseline satellite reliability	.75
V_r	reliability expenditure	*eqn 4.11*
ϕ_R	system reliability cost factor	*eqn 4.14*

space based replacement to AWACS is a critical system for achieving and maintaining air and space superiority, the required mission reliability, R_{ms}, will be high. For analysis and sizing, we assume a instantaneous 99% mission reliability requirement. Mission reliability is usually modeled as a function of time. For example, one model for reliability that has been widely used for system analysis is the exponential distribution [29]

$$R(t) = R_i \exp\left(-\int_0^t \lambda(\tau)d\tau \right) \qquad\qquad eqn\ 4.1$$

where $\lambda(t)$ is the failure rate and R_i is the initial reliability. A Weibull distribution is also used sometimes for certain subsystems or components. In systems engineering, the end of life reliability is typically a system specification. With estimated failure rates for subsystems and the overall satellite, the required beginning of life reliability can be determined and designed. The reliability model presented below employs a constant mission reliability specification which should be viewed as the instantaneous reliability value. This again illustrated the ease and flexibility of a parametric model: a reasonable reliability specification can be used without having to develop complex failure models that may not actually reflect the actual system. The parameter for mission reliability is easily changed.

The individual satellite reliability, R_0, is the probability that the satellite performs its portion of the mission and includes both performance reliability and survivability. A baseline spacecraft reliability, R_{sc}, is used to scale the relative cost between satellites with different reliabilities. The cost expenditure to increase a satellite from baseline reliability R_{sc} to the required satellite reliability R_0 is the subject of Section 4.3. The performance parameters P_d and P_c represent the satellite and configuration probability of detection respectively from the distributed operations concept (Chapter 5). The minimum acceptable performance effectiveness, P_e, is the

percent capability of the original mission that is still acceptable for mission performance. Equivalently, it represents the maximum allowable performance degradation. These parameters are summarized in *Table 4.1* and defined in Section 4.2. *Figure 4.1* illustrates the effect of each of the parameters in an example that is discussed in detail below.

It is easiest to introduce the reliability model through an example. *Figure 4.1* depicts the process and logic of reliability model. For a distribution of one satellite, the satellite must be 99% available to meet 99% mission reliability specification. The desired mission capability is 95% probability of detection. As distribution increases, the cooperative operations concept developed in Chapter 5 allows the individual satellite detection capability to decrease and still meet the detection specification. For a distribution of two, each satellite must only be capable of 78% probability of detection to achieve the desired mission capability of 95% probability of detection (that is, 95% configuration probability). However, now the two satellites are two critical mission components in series and the reliability/availability of each satellite must increase to give the desired mission reliability. For a configuration of two satellites, each satellite must be 99.5% reliable. This increase in satellite reliability/availability results in an increased reliability cost that then increases overall system cost.

The distributed operations concept then results in decreased reliability rather than increased reliability as desired. In fact, this is counter to the conventional wisdom regarding distributed systems. In general distributed systems are more reliable because a failure at a single node does not result in total system failure. The cooperative operations concept for SBR, however, places the nodes (satellites) in series which is less reliable. However, the satellites in the configuration that continue operating after a satellite failure still have some capability which should be included in the model. Thus, we define the maximum allowable mission degradation (or equivalently, the minimum acceptable mission performance) as a percentage of the original mission capability.

Figure 4.1 uses a maximum 10% mission degradation to illustrate the process. A maximum degradation of 10% is equivalent to a configuration that has at least 90% of the capability of the original configuration. For a configuration of two satellites, one satellite alone cannot achieve the 90% effectiveness. Thus, at a distribution of two there is no reliability advantage for distributing (for the assumed mission parameters). For a configuration of three satellites, however, a loss of one satellite results in less than 10% performance degradation for the remaining two satellites in the configuration. Since only two of three satellites must operate at end of life or during hostile enemy action, each satellite must only be 93.9% reliable to achieve the desired mission reliability of 99%. This represents a significant savings in satellite cost. At a distribution of five satellites, only one satellite is allowed to fail to achieved the desired mission effectiveness. Thus, four satellites must operate in series which requires the individual satellites to be 96.6% reliable. A configuration of five satellites will have a higher cost for reliability than the configuration of three satellites (for these assumed mission parameters). For a distribution of ten satellites, three satellites may fail and each satellite must only be 90.6% reliable.

baseline S/C reliability	reliability cost factor		configuration prob detection	configuration size
$R_{S/C} = .75$	$3.3 \times \$_{S/C} \longrightarrow$	$R_0 = .99$ $P_d = .95$	$P_c = .95$	$n = 1$
$R_{S/C} = .75$	$3.8 \times \$_{S/C} \longrightarrow$	$R_0 = .995$ $P_d = .78$	$P_c = .95$	$n = 2$
$R_{S/C} = .75$	$2.0 \times \$_{S/C} \longrightarrow$	$R_0 = .939$ $P_d = .63$	$P_c = .95$	$n = 3$
$R_{S/C} = .75$	$2.4 \times \$_{S/C} \longrightarrow$	$R_0 = .966$ $P_d = .45$	$P_c = .95$	$n = 5$
$R_{S/C} = .75$	$1.7 \times \$_{S/C} \longrightarrow$	$R_0 = .?'?$ $P_d = .26$	$P_c = .95$	$n = 10$

Figure 4.1: Reliability model overview. Reliability required per satellite for a 99% probability of mission success and a 10% allowable mission degradation. The reliability cost factor is the expenditure required to bring a 75% reliable spacecraft up to the required spacecraft reliability, R_0. The configuration probability of detection is P_c and the individual satellite detection capability is P_d. The X-out satellites represent the number of satellites that can be lost from the configuration and still achieve at least 90% of the desired mission capability.

The cost for reliability model is developed in Section 4.3. The cost for an increase in satellite reliability is expressed as a percentage of the satellite cost. For the process illustrated in *Figure 4.1*, a normalizing, 75% reliable satellite is a baseline to which expenditure for reliability is required to bring the satellite up to the desired reliability. The reliability cost is then a cost factor relative to some baseline.

4.2 Reliability for Distributed Systems

Reliability is the complement of failure. Thus, reliability is quantified as one less the probability of failure. In the following, "reliability" is meant to imply both reliability and survivability.

To model system reliability, the required mission reliability is given by the mission specifications. For an SBR search radar, the mission is detection of targets. The search concept of operations specifies the probability that a distributed search configuration of n_s satellites will

detect a target: P_c (*eqn 5.10*). This can be rewritten as

$$P_d = 1 - \left(1 - P_c\right)^{1/n_s}$$

<div align="right">*eqn 4.2*</div>

where P_d is the probability of detection capability for a single satellite (*eqn 5.7*).

If m out of the n_s configuration satellites function properly, the mission effectiveness as a fraction of the original mission capability is

$$P_e = \frac{1 - \left(1 - P_d\right)^m}{P_c}$$

<div align="right">*eqn 4.3*</div>

where P_e is the minimum acceptable mission performance or mission effectiveness. Mission effectiveness is plotted in *Figure 4.2* for several values of m as a function distribution (n_s) and configuration capability, P_c. Recall that the configuration probability of detection is a system specification or mission requirement.

For a given mission effectiveness, P_e, the number of satellites that can be lost from the configuration and still provide the desired mission effectiveness is n_s - m. m must be determined iteratively (for example, by a root-finding method) from *eqn 4.3*. The result is plotted as a function of n in *Figure 4.3* – note the step nature which is due to the discrete increments of satellites that can be lost The paradigm of reliability for losing a certain number of satellites is an
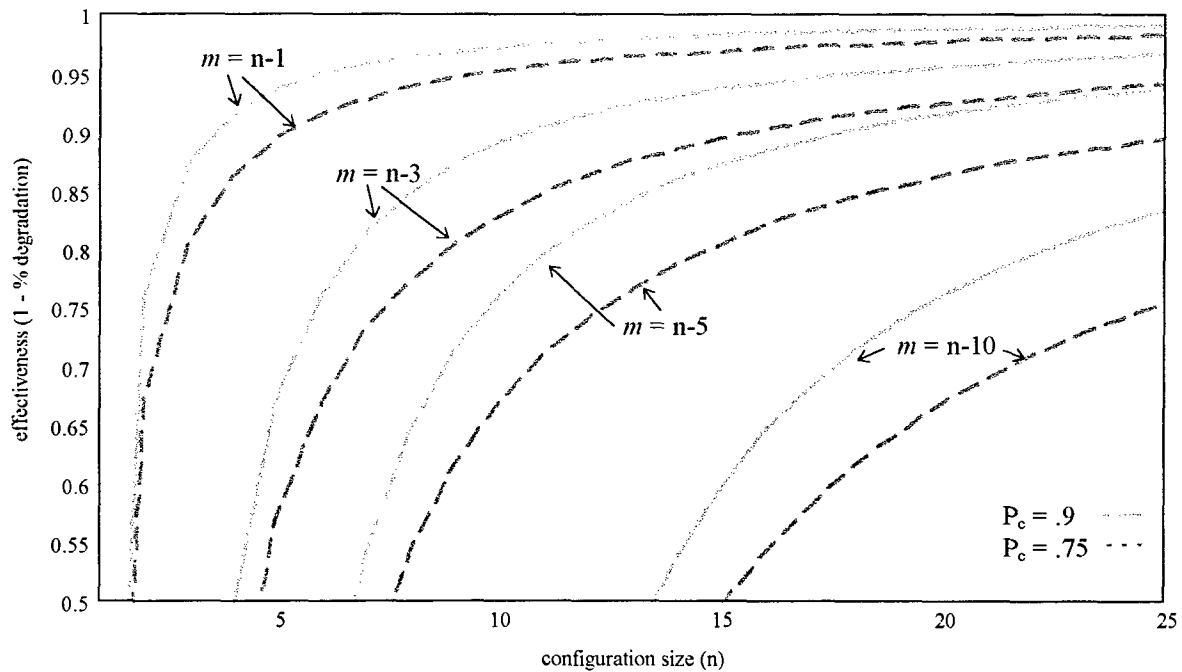
Figure 4.2: Mission effectiveness expressed as a fraction of prior mission capability as a function of the number of satellites in the operational configuration. Two configuration probability of detections are plotted: P_c = .95. and .75. The inverse function relates the number of satellites that can be lost in a configuration for a given mission effectiveness (see Figure 4.3).

m-out-of-n redundant system. For example, a car with four tires and a spare is a 4-out-of-5 redundant system: one tire can be lost and the car can still perform its mission. This illustrates that there are two perspectives for viewing the relationship between P_e and P_c: acceptable mission degradation and spare capability.

The perspective of minimum acceptable performance is given by eqn 4.3. P_e is a mission specification that quantifies the percentage of the mission that is considered essential. Equivalently, it quantifies the maximum acceptable performance degradation. For example, 90% minimum mission effectiveness corresponds to a decrease of no more than 10% in target detection capability for the SBR mission. The minimum acceptable performance parameter is also a measure of how "over-spec'ed" the system requirements are. It is typical of the defense acquisition process to pad system requirements with a little bit of margin to ensure system performance when there is uncertainty as to what the actual battlefield requirements should be. In this context, P_e is original performance requirement and P_c is what the contractor is tasked to provide.

The second perspective for the relationship between P_e and P_c is in the ℓ⋅ ∪ᵢ space capability. Another common practice for defense systems is to achieve performance robustness by using spares. For example, the 24 satellite GPS constellation includes three on orbit spares. If spare resources are put into orbit, they should be used since they enhance the overall system performance without increasing the system failure rate. Thus, P_e is a measure of the performance without spares and P_c is a measure of the system performance when the spares are used to augment performance.

For the m-out-of-n redundant system, the binomial probability distribution models the number of system elements that continue to operate. If the reliability of a single element or
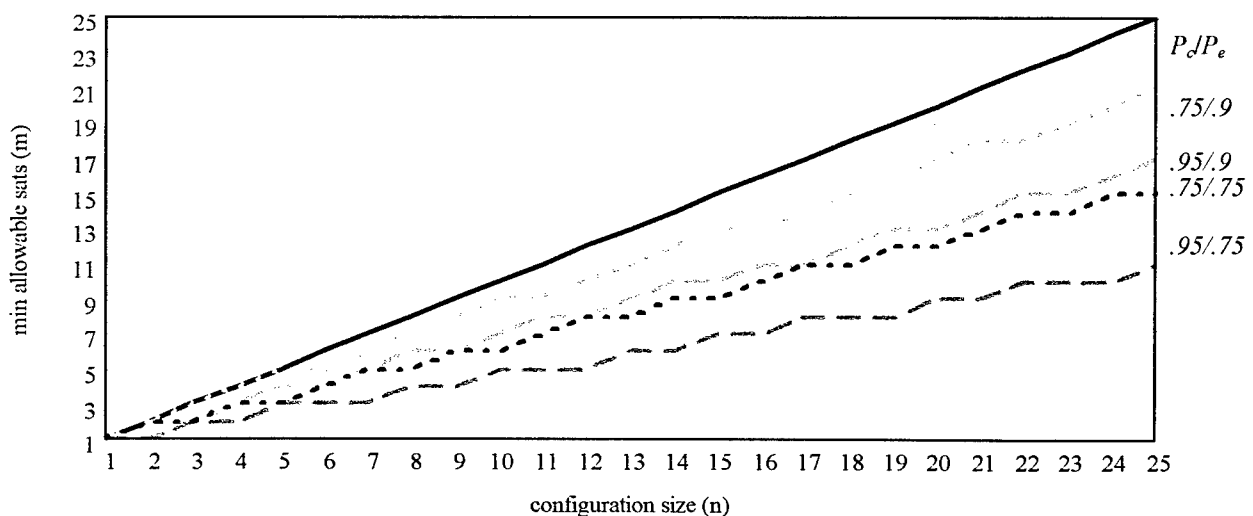
Figure 4.3: The number of satellites required for a given mission effectiveness as a function of the number of satellites in the operation configuration. The reliability parameters are given in the form P_c/P_e.

satellite is R_m, the probability that at least m satellites function is

$$P(sats \geq m) = \sum_{k=m}^{n} \frac{n!}{k!(n-k)!} R_m^{\,k} \left(1 - R_m\right)^{n-k} = 1 - F_{ms} \qquad\qquad eqn\ 4.4$$

where F_{ms} is the probability of system failure. The mission reliability, R_{ms}, is one less the system probability of failure.

eqn 4.4 can be solved iteratively for the value of R_m for which the *m*-of-n_s reliability is equal to the desired mission reliability, R_{ms}. This gives the individual satellite reliability required to achieve the specified mission reliability when only *m* out of *n* functioning satellites satisfies the minimum acceptable performance effectiveness. The result as a function of configuration size is plotted in *Figure 4.4*.

Note that the saw-toothed nature of *Figure 4.4* is the result of the incremental steps in minimal allowable satellites, *m*, in *Figure 4.3*. At times, the satellite reliability actually increases as distribution increases, because the overall configuration size has not yet reached the critical point at which another satellite can fail and the system still meet performance requirements. For example, for the 95% configuration probability of detection with 25% allowable performance degradation system (the .95/.75 system), the configuration of two distributed satellites can lose one satellite and still meet performance. When the configuration is increased to three distributed satellites, only one satellite can be lost from the configuration to meet the performance
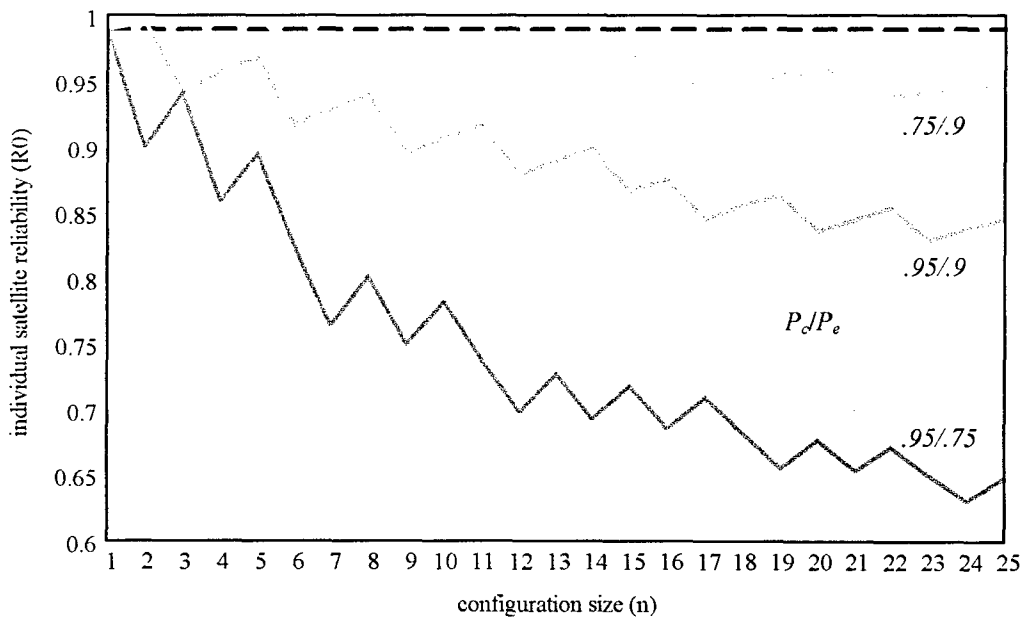
Figure 4.4: Individual satellite reliability as a function of the number of satellites in the operational configuration. Individual spacecraft reliability occasionally increases due to the discrete jumps of m required (Figure 4.3). Because of the operations concept, several satellites in series must have greater reliability until the configuration is large enough that losing a satellite does not exceed the acceptable degradation.

specifications. That is, two out of three satellites must function which requires more reliability per satellite than if only one out of two satellites must function. Thus, the individual satellite reliability increases when the distribution is increased from two to three. If the configuration is a set of four distributed satellites, two out of four function satellites meets performance requirements and the reliability per satellite can decrease. This saw-toothed feature of reliability is so strong that it manifests itself in the reliability cost factor (Section 4.4) and ultimately in the total system cost (Section 8).

4.3 Cost for Reliability

A relationship for the cost to increase reliability is desired to compare systems with different reliabilities. For example, two system architectures may have different inherent reliabilities: system A is more reliable than system B. To increase the reliability of system B to match the reliability of system A requires some resource expenditure: the reliability expenditure. If the cost for additional reliability can be quantified, the two systems can be fairly compared o the basis of system cost. This is the cost for reliability model that we develop here. The model relates the additional resource expenditure required to improve reliability to the total system cost and is modified from [30,31]. *Figure 4.5* aids in the description of the model.

For a system with probability of failure F_s, the expected cost of failure, V_F, is the probability of failure times the cost of failure, V_{comp}.
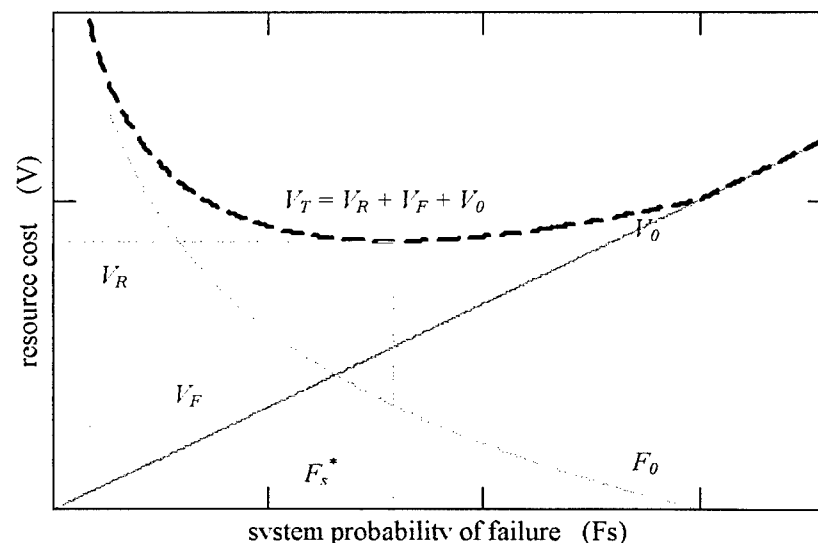
Figure 4.5: Reliability cost model. Reliability expenditure, V_R, reaches zero when the system probability of failure, F_s, equals the baseline probability of failure F_0. Total expected system cost, V_T, is then equal to cost of failure, V_F. At F_0 the cost of failure is equal to the baseline cost, V_0 ; that is the cost of another unit. There is a shallow minimum for total cost when the slopes of V_R and V_F are equal and opposite.

72

$$E[V_F] = V_F = F_s V_{comp}$$ <div align="right">*eqn 4.5*</div>

where V_{comp} is the cost expenditure required to compensate for failure. Assuming that failure compensation cost is constant, then

$$V_{comp} = \frac{V_F}{F_s} = \frac{V_0}{F_0}$$ <div align="right">*eqn 4.6*</div>

Therefore, if the reliability of a system with an initial failure probability of F_0 is increased (e.g., through redundancy or spares) to $R_s = 1 - F_s$, then the expected cost of failure after the increase in reliability is

$$V_F = \frac{V_0}{F_0} F_s$$ <div align="right">*eqn 4.7*</div>

where V_0 can be taken as the baseline system cost. The increase in reliability required some expenditure V_R, the cost of reliability. Total expected system cost is then the baseline system cost plus the reliability expenditure plus the expected cost of failure.

$$V_T = V_R + V_F + V_0$$ <div align="right">*eqn 4.8*</div>

We desire a relationship between reliability expenditure, V_R, and the subsequent system reliability. If V_R equals V_0, then it is possible to purchase a spare baseline unit. The total system now consists of two units in parallel and the probability of system failure is F_0^2. If V_R equals kV_0, it is possible to purchase $k + 1$ units, and the system reliability is $F_s = F_0^{k+1}$. This expression is valid for all non-negative integer values of k, including $k=0$: no reliability expenditure is made and the system reliability is equal to the unit reliability. If the relationship is extended to include rational values of k, then V_R is a fraction of the baseline unit cost and the system reliability becomes

$$F_s = F_0^{\frac{V_R}{V_0} + 1}$$ <div align="right">*eqn 4.9*</div>

This is equivalent to the relationship between system reliability and reliability expenditure given in [30,31]:

$$\frac{F_s}{F_0} = F_0^{V_R/V_0}$$ <div align="right">*eqn 4.10*</div>

The system reliability and reliability expenditure (eqn 4.9) can be expressed in the equivalent and more convenient form

$$V_R = \frac{\ln(F_s) - \ln(F_0)}{\ln(F_0)} V_0$$ <div align="right">*eqn 4.11*</div>

This is the reliability expenditure required to reduce the initial probability of failure, F_0, to the desired mission probability of failure, F_s.

The minimum expected total system cost expenditure occurs when the derivatives of V_R and V_F with respect to F_s cancel each other. That is, the optimum reliability expenditure occurs when the slope of V_R as a function of F_s is the negative slope of the cost-of-failure line. Thus, at the minimum total expected cost, the marginal increase in reliability expenditure cancels the marginal decrease in failure cost. The optimum (minimum expected system cost) system probability of failure is then

$$F_s^* = \frac{-F_0}{\ln(F_0)}$$

and the optimum reliability expenditure is

$$V_R^* = \frac{\ln\left(\frac{-F_0}{\ln(F_0)}\right) - \ln(F_0)}{\ln(F_0)} V_0$$

Note that the optimal reliability expenditure is optimum only in the respect that it minimizes the expected system cost (which includes the cost of failure). Mission requirements may dictate a system reliability other than the optimal reliability. Furthermore, the total system cost is typically fairly flat in the vicinity of the optimum reliability (*Figure 4.5*) so that there is not a significant penalty for systems designed at a different reliability specification. The cost does become significant as the system reliability approaches unity. For these reasons, the optimum reliability is seldom used in practice.

4.4 System Reliability Cost Factor

From *eqn 4.11*, the reliability cost factor required to increase a satellite with baseline reliability, R_{sc}, to the desired individual reliability, R_m, is then

$$\phi_R = \begin{cases} 1 + \dfrac{\ln(1-R_m) - \ln(1-R_{sc})}{\ln(1-R_{sc})} & \text{if } R_{sc} < R_m \\ 1 & \text{if } R_{sc} \geq R_m \end{cases}$$

eqn 4.14 is plotted in Figure 4.6 for two values of baseline spacecraft reliability.

The reliability cost factor is actually a function of distribution by the process developed in Section 4.2. For completeness,

$$\phi_R(n_s) = \begin{cases} 1 + \dfrac{\ln(1-R_m(n_s)) - \ln(1-R_{sc})}{\ln(1-R_{sc})} & \text{if } R_{sc} < R_m(n_s) \\ 1 & \text{if } R_{sc} \geq R_m(n_s) \end{cases}$$

where $R_m(n_s)$ is given by the solution to
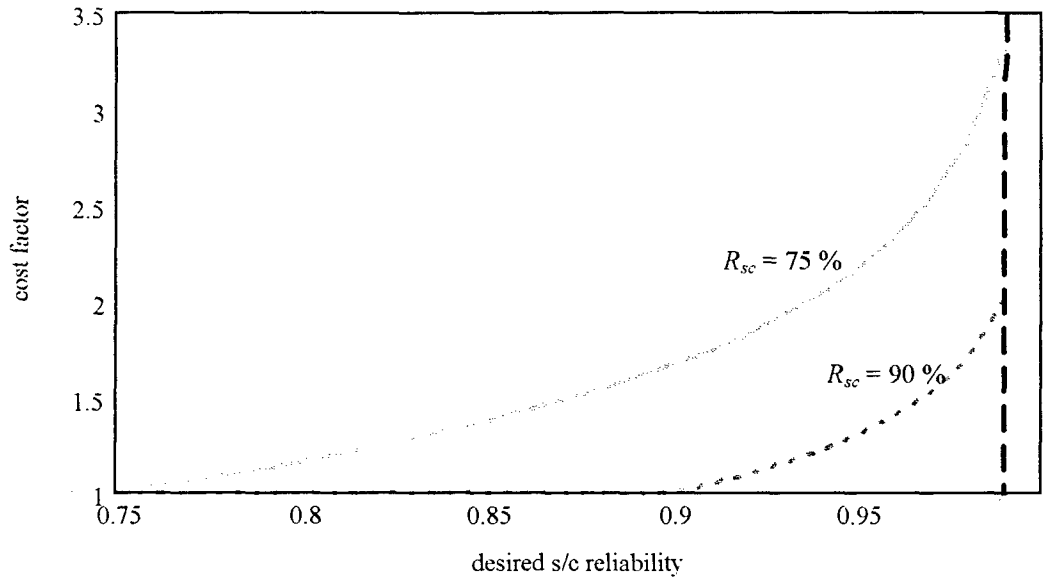
Figure 4.6: Reliability cost multiplier as a function of the desired satellite reliability. The solid line is for an initial 75% reliable baseline satellite; the dashed line is for a 90% reliable satellite. The vertical boundary is the mission reliability goal of 99%. Note that a 90% reliable spacecraft should be approximately 1.63 times as expensive as a 75% reliable spacecraft.

$$R_m(n_s) = \left\{ p \ni R_{ms} - \sum_{k=m(n_s)}^{n_s} \frac{n_s!}{k!(n_s-k)!} p^k (1-p)^{n_s-k} = 0 \right\} \qquad eqn\ 4.16$$

where the lower limit of summation, $m(n_s)$, is

$$m(n_s) = \left\{ m' \ni P_e - \frac{1-(1-P_c)^{m' n_s}}{P_c} = 0 \right\} \qquad eqn\ 4.17$$

Both *eqn 4.16* and *eqn 4.17* require an iterative root-finding process to solve the equations in brackets. Finding the desired root for *eqn 4.16* is an unstable numerical process that is sensitive to the initial seed for p. Nevertheless, we have never found a set of parameters for which the numerical process would not converge for any seed.

When *eqn 4.15* is combined with the inverse of *eqn 4.16* and *eqn 4.17*, the reliability cost factor is a function of the number of satellites in the configuration. The result is plotted in *Figure 4.7* for three different sets of system requirements P_c and P_e. The reliability cost factor is employed by each of the subsystem cost multipliers developed in Section 3.

A summary for the process of determining the reliability cost factor for a general distributed system:

1. Establish mission reliability requirement: R_{ms}. This is the cumulative probability of a correctly functioning system.

75
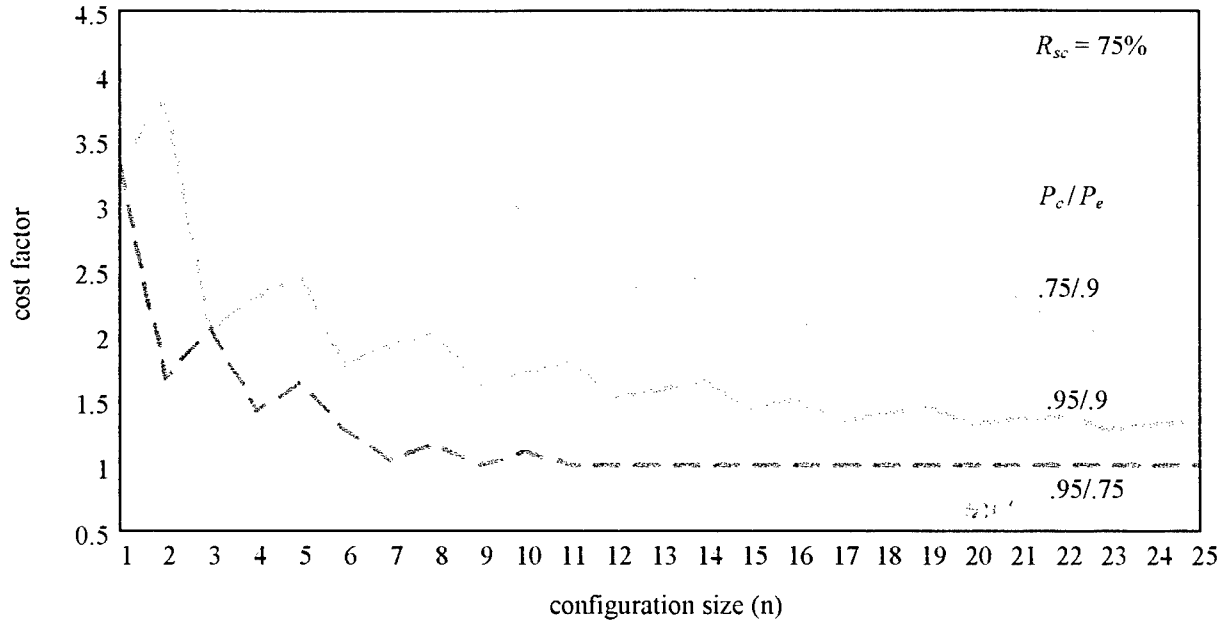
Figure 4.7: System cost factor to achieve desired mission reliability as a function of the number of satellites in the configuration. The baseline reliability is 75%.

2. Establish mission performance requirement: P_c. Mission performance must be an explicit function of distribution.

3. Establish minimum acceptable mission performance or effectiveness, P_e, as a percentage of specified mission performance, P_c.

4. Solve *eqn 4.17* which gives the number, m, of distributed elements that can be lost without exceeding the maximum allowable degradation: $1-P_e$.

5. Solve *eqn 4.16* which gives the reliability per distributed element, R_m, to achieve the desired mission reliability, R_{ms}.

6. *eqn 4.15* gives the reliability cost factor relative to some element reliability, R_{sc}. The reliability cost factor can be interpreted as the system cost normalized by the baseline system cost, R_{sc}.

5. Operations Concept

The operations concept is the manner in which the system function is implemented with the system assets. We make a distinction between operations and operations concept, though they are closely related. Operations is the day to day use and administration of the system to perform the mission. Operations concept is the more theoretical perspective of how the system assets are used to achieve certain objectives, including mission performance. Different operations concepts will in general result in different system cost and performance. Thus, operations concept is an important part of the system design process.

The Space Based Radar operations concept analysis should include the system and function distribution options that could fulfill the mission requirements. Operation variants could include co-aperture transmitter/receiver, multistatic transmit or receiver with coherent or non-coherent units, Unmanned Aerial Vehicle (UAV) receiver as either an adjunct or bistatic receiver, space based search radar with atmospheric track radar (atmospheric receiver with space illuminator or atmospheric transmit/receive platform), and an endless hybrid of these options. The operations concept given here employs a distributed set of co-aperture transmitter/receivers.

Section 5.1 discusses the mission tasks that should be considered for a next generation space based replacement to AWACS. A space based, distributed search operations concept that reduces the power and aperture required per satellite is presented in Section 5.2. The mean time to detection for a space based radar is derived in Section 5.3.

5.1 SBR Mission Tasks

For the AWACS mission, four different mission tasks must be considered. These are outlined in *Table 5.1*. The mission tasks may be divided into two classes: search and track. The purpose of the search class is primarily to detect the existence of targets. The purpose of the track class is to determine the location of the targets so that they may be neutralized.

Before any action can be taken against a target, the presence of the target must be determined. This is the purpose of the search mission task. The search process involves searching a given area (volume) with a beam footprint (solid angle) that is generally much smaller than the area to be searched. Thus, to cover the entire search area, the beam must be scanned through the area. The beam footprint defines a resolution cell with coarse granularity. Better resolution can be achieved through beam splitting and signal processing. Beam splitting improves the cross-range or azimuth resolution by determining the center of the n pulses reflected by a point as the beam is scanned past the point. Signal processing can be used to improve both range and Doppler resolution through waveform selection and waveform processing. Section 2.3 discusses these topics in detail.

Table 5.1: Space Based Radar mission task description.

Mission Task	Description	Class
search	detect and locate targets with coarse granularity	search
track-while-search	update of coarse target position with subsequent search scans while continuing search for other targets (undersampled)	search
track	target location with good resolution and sufficiently high update rate to maintain track (also includes initial track acquisition)	track
continuous track	slave of track radar to target for the purpose of target illumination for intercept – requires error signal for feedback	track

Using an updated target position with positions from previous search scans to form a "track" is the track-while-search mission task. Track-while-search is the "grease pencil" approach – targets at each scan are marked with a grease pencil and the points are connected with a curve to represent the "track". Because the "track" is not sampled at a rate sufficient to predict the target position at the next scan, track-while-search is classified as a search radar. The difference between true track and search should become clear in the track discussion that follows. Track-while-search can be very useful, particularly against targets that are low threat but for which some intelligence may reveal future intentions. For example, track-while-search could be used while searching outside a given threat radius to free other assets for search or track within the threat radius.

Once a target has been detected, it is desirable in some instances to track the target. Scenarios requiring a good granularity track include precision intercept vectoring, hand-off to another tracking radar, and target illumination. We distinguish tracking for the purpose of target illumination as continuous tracking. Another aspect of track radar is the need for a track error signal that is used in feedback to guide the beam. Methods for generating error signals include phase monopulse, wideband monopulse , sequential lobbing, and conical or off-axis scan [15,13]. Each of these methods work by inducing a known error in the beam direction and comparing the expected and actual signal return.

Error signals are particularly important for a continuous track mission since the beam must remain locked on the target. In continuous track, the radar is used to illuminate the target for some other platform, particularly for intercept by a missile with no transmitter of its own or for intercept by an aircraft that does not wish to reveal its position. Continuous track could also be used to illuminate the target for hand-off to another radar.

Track necessarily requires better position granularity than used in search. Though this need not imply a smaller beam (the methods for improving resolution discussed above for search also apply to track), a smaller beam decreases the amount of energy that must be transmitted by the radar. For typical ranges of a space based radar platform, track beamwidths at the target comparable to atmospheric platforms would require very large apertures. Track in the traditional paradigm would be inefficient from space. When considering the mission tasks that track serves,

78

however, track from a space based platform may well be anachronous. The narrow beams required for efficiency also help generate a better error signal. With the typical beam footprint (10's to 100's of km at UHF) of a moderate aperture space platform, the tasks given to track may be performed from a platform designed for search. For example, continuous track requires an error signal to keep a tight beam on the target for the purpose of target illumination. If the beam is so large that the target is unlikely to leave the footprint in the time required for intercept, no error signal or "track" update is required.

For these reasons, we concentrate on defining and optimizing the search mission.

5.2 Distributed Search Operations Concept

As described in other chapters, distribution of function results in several positive effects for reducing the cost of performing the next generation AWACS mission from a space based platform. The distributed operations concept presented here takes advantage of the physics and probability of detection to reduce the power and aperture required by a distributed space based radar platform.

The distributed search operations concept distributes the detection function among separate satellites which reduces the radar power and aperture required for each satellite. The search operations concept involves a configuration of n_s search radars that are cooperatively but independently searching the area, A. The function is distributed in that no one radar is responsible for detecting a target in any given subarea. The configuration size, n_s, is a measure of the degree of distribution. Because several radars are searching the same area, they can each have a reduced probability of detection and still maintain a given configuration probability of detection. Because of the nature of the relationship between signal to noise ratio and probability of detection, the reduction in satelitte power-aperture product is more than just dividing the power-aperture between n_s satellites. The details are given below.

5.2.1 Probability of Detection:

The purpose of the detector (*Figure 2.1*) is to detect a valid radar signal return in the presence of noise while rejecting false returns. False returns are signals that look like valid targets but are actually due to noise, multipath, clutter, reverberation, or other effects. An envelope detector sets a threshold envelope above the rms noise envelope. The envelope of a signal is the complex amplitude given by quadrature demodulation (Section 2.3). When the received signal envelope exceeds the threshold, a detection is recorded. If the threshold is set too low, noise will tend to give too many false detections; too high a threshold results in missing potential targets at moderate signal-to-noise levels. The threshold envelope can be determined analytically for a desired probability of detection and false alarm rate (or equivalently, a probability of false alarm). Generally, a detection is not noted unless M out of N pulses exceed the threshold. This is known as an M of N detection scheme.
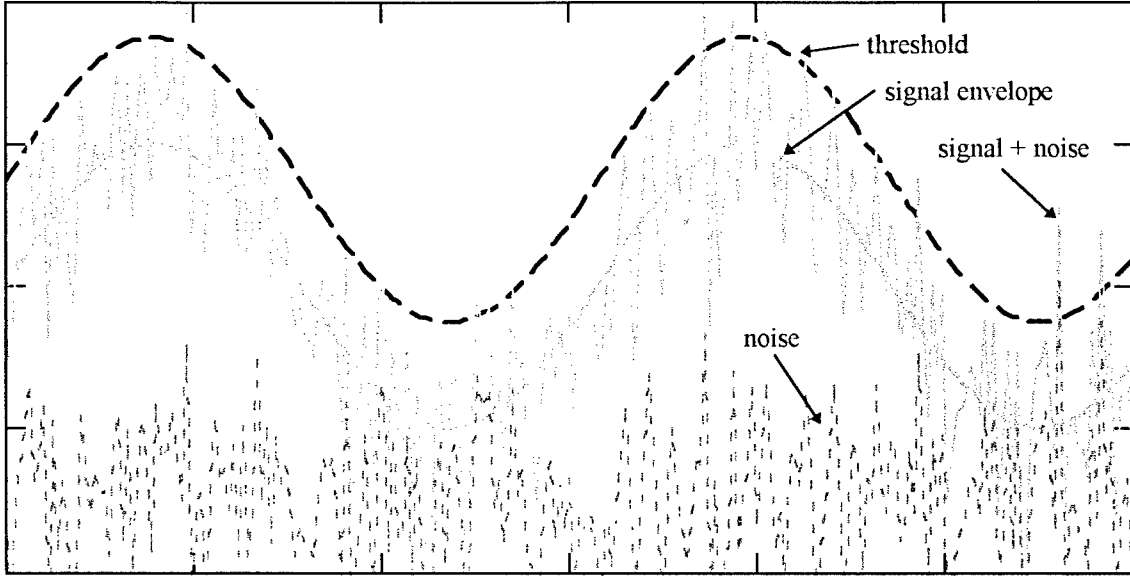
Figure 5.1: Envelope threshold with noise, des'.'ed signal, and signal plus noise envelope.

The received signal that enters the detector from the intermediate filter (IF) is the filtered signal plus noise envelope. The IF typically approximates a matched filter -- an ideal filter whose transfer function maximizes the peak signal to mean noise power ratio (Section 2.3). For a rectangular transmitted pulse of width, τ, the filter bandpass, B, for the matched filter satisfies

$$B\,\tau \cong 1 \qquad\qquad\qquad\qquad eqn\ 5.1$$

This is the same assumption used in the derivation of the radar range equation, *eqn 2.57*.

If the noise entering the intermediate filter has a Gaussian probability distribution with zero mean and RMS noise voltage of $\sqrt{\sigma_0}$, the probability density of the noise voltage output is a Rayleigh distribution [9] with

$$P(R) = \frac{R}{\sigma_0}\exp\left(-\frac{R^2}{2\sigma_0}\right) \qquad\qquad\qquad eqn\ 5.2$$

where R is the noise envelope amplitude. The probability of false alarm is then

$$P_{fa} = P\left(V_T < R < \infty\right) = \int_{V_T}^{\infty} P(R)dR = \exp\left(-\frac{V_T^2}{2\sigma_0}\right) \qquad eqn\ 5.3$$

where V_T is the threshold voltage. The false alarm time is the mean time between false alarms

$$T_{fa} = \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} T_i \qquad\qquad\qquad eqn\ 5.4$$

where T_i is the time between successive crossings of the voltage threshold by noise. The false alarm rate is then
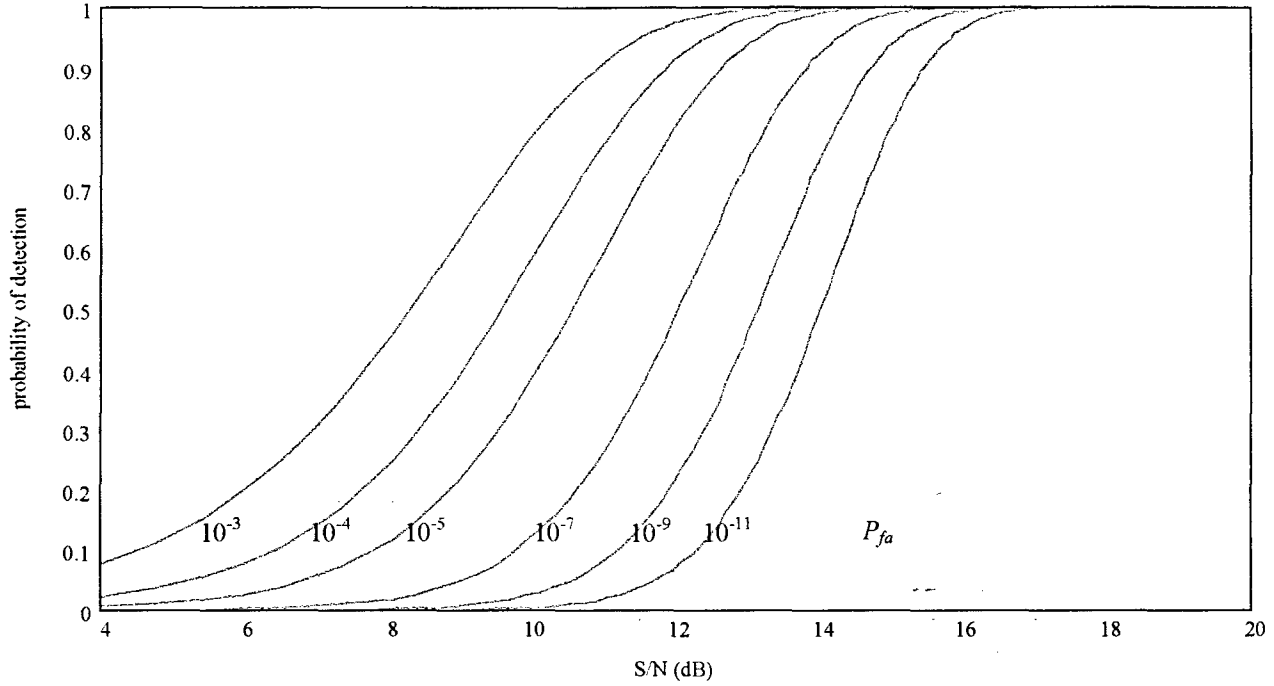
Figure 5.2: Probability of detection as a function of signal to noise ratio for a range of false alarm probabilities.

$$far \approx \frac{1}{T_{fa}} = P_{fa}B \qquad\qquad eqn\ 5.5$$

This gives a convenient method for determining reasonable values for the probability of false alarm. For a *far* of one false alarm per hour at a 1 MHz bandwidth, the probability of false alarm has to be 2.78×10^{-10} and the threshold envelope must be 4.4 times the noise RMS. In practice, the threshold voltage is set higher to avoid a flood of false alarms resulting from the false alarm time sensitivity to instabilities in the threshold level [15].

If the desired signal is a sine wave with amplitude, A, is corrupted by noise with amplitude R, the probability density of the output from the envelope detector is

$$P(R) = \frac{R}{\sigma_0}\exp\left(-\frac{R^2 + A^2}{2\sigma_0}\right)I_0\left(\frac{RA}{\sigma_0}\right) \qquad\qquad eqn\ 5.6$$

where I_0 is a modified Bessel function of zero order. eqn 5.6 is the Rice or Rician probability density function [32]. The probability of detecting the target from the received signal is

$$P_d = P_{sig}\left(V_T < R < \infty\right) = \int_{V_T}^{\infty} \frac{R}{\sigma_0}\exp\left(-\frac{R^2 + A^2}{2\sigma_0}\right)I_0\left(\frac{RA}{\sigma_0}\right)dR \qquad\qquad eqn\ 5.7$$

for which no closed form solution exists. Rice gives the series approximation

$$P_d = \tfrac{1}{2}\left(1 - erf\left(\frac{V_T - A}{\sqrt{2\sigma_0}}\right)\right) + \frac{\exp\left(\frac{-(V_T-A)^2}{2\sigma_0}\right)}{2\sqrt{2\pi}\,\frac{A}{\sigma_0}}\left[1 - \frac{V_T - A}{4A} + \frac{1+(V_T-A)^2/\sigma_0}{8A^2/\sigma_0} - \cdots\right] \qquad eqn\ 5.8$$

or equivalently

$$P_d = \tfrac{1}{2}\left(1 - erf\left(\sqrt{\ln\left(\frac{1}{P_{fa}}\right)} - \sqrt{\frac{S}{N}}\right)\right) + \frac{\exp\left(-\left(\sqrt{\ln\left(\frac{1}{P_{fa}}\right)} - \sqrt{\frac{S}{N}}\right)^2\right)}{4\pi\sqrt{\frac{S}{N}}}\left[\frac{3}{4} - \frac{\sqrt{\ln\left(\frac{1}{P_{fa}}\right)}}{4\sqrt{\frac{S}{N}}} + \frac{1+\left(\sqrt{2\ln\left(\frac{1}{P_{fa}}\right)} - \sqrt{2\frac{S}{N}}\right)^2}{16\sqrt{\frac{S}{N}}} - \cdots\right] \qquad eqn\ 5.9$$

In practice, the signal to noise ratio that should be used to calculate the probability of detection is the integrated signal to noise ratio (Chapter 2). Another factor that significantly alters the probability of detection for a given signal to noise ratio is target RCS fluctuations. In general, when the desired probability of detection is greater than approximately .3, an increased signal to noise ratio is required on a single hit basis. A slight decrease in required signal to noise is realized for a probability of detection less than .3. The literature on fluctuating targets (also known as the Swerling fluctuation cases) is extensive [15,33,34,35,36]. For the model used in Chapter 3, only the effects of integrated signal to noise were used.

5.2.2 Search Configuration Probability of Detection:

The probability of detection, P_d, for a single radar is a function of the signal to noise ratio of the received signal that is reflected from the target. The probability of detection for the distributed operations concept outlined above is a function of the number of radars, n_s, and the probability of detection for each of the radars. The probability of detection for the configuration is henceforth referred to as the probability of the configuration, P_c. For n_s radars searching the same area at a detection probability of P_d, the probability that a target is detected by at least one radar is

$$P_c = 1 - \left(1 - P_d\right)^{n_s} \qquad eqn\ 5.10$$

Similarly, the configuration probability of false alarm is

$$P_{fa}^{(c)} = 1 - \left(1 - P_{fa}\right)^{n_s} \qquad eqn\ 5.11$$

The operations concept also results in an undesirable increase in the probability of false alarm for the configuration. To counter this, an m-out-of-n detection scheme can be used. Since noise is uncorrelated and independent, the probability of m false alarms at the same processing bin is

$$P_{fa}^{(c)} = \left(P_{fa}\right)^{m} \qquad eqn\ 5.12$$
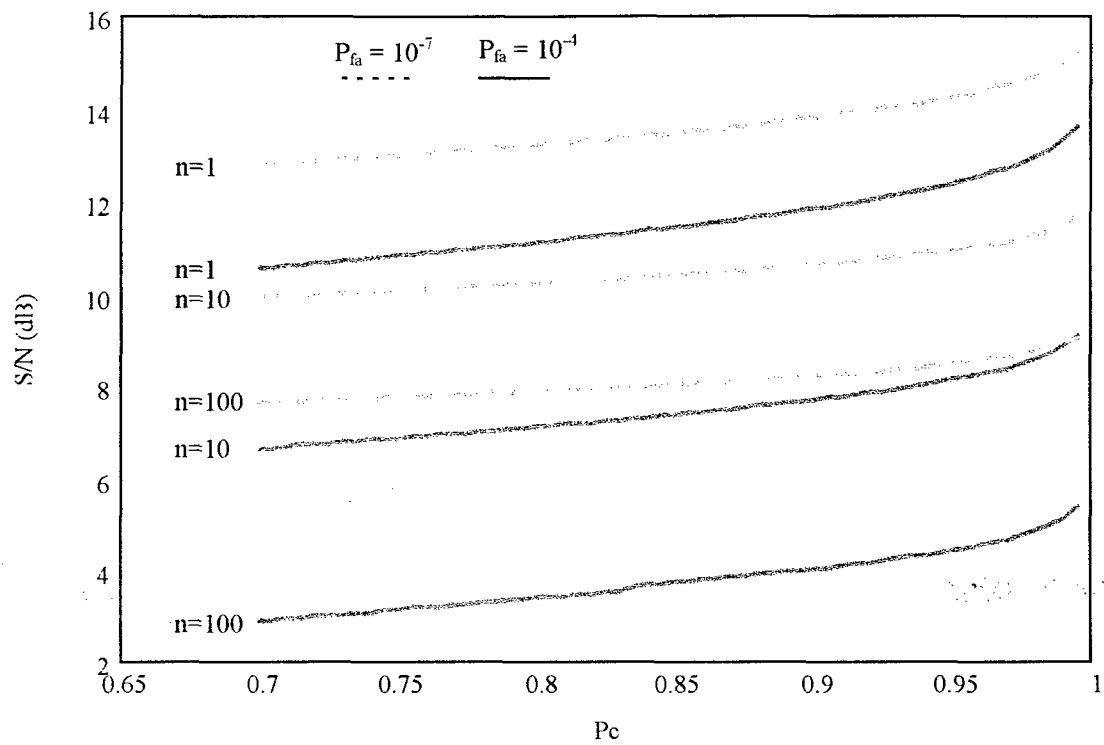
Figure 5.3: P_c *and required* S/N *for different size configurations at two false alarm probabilities.*
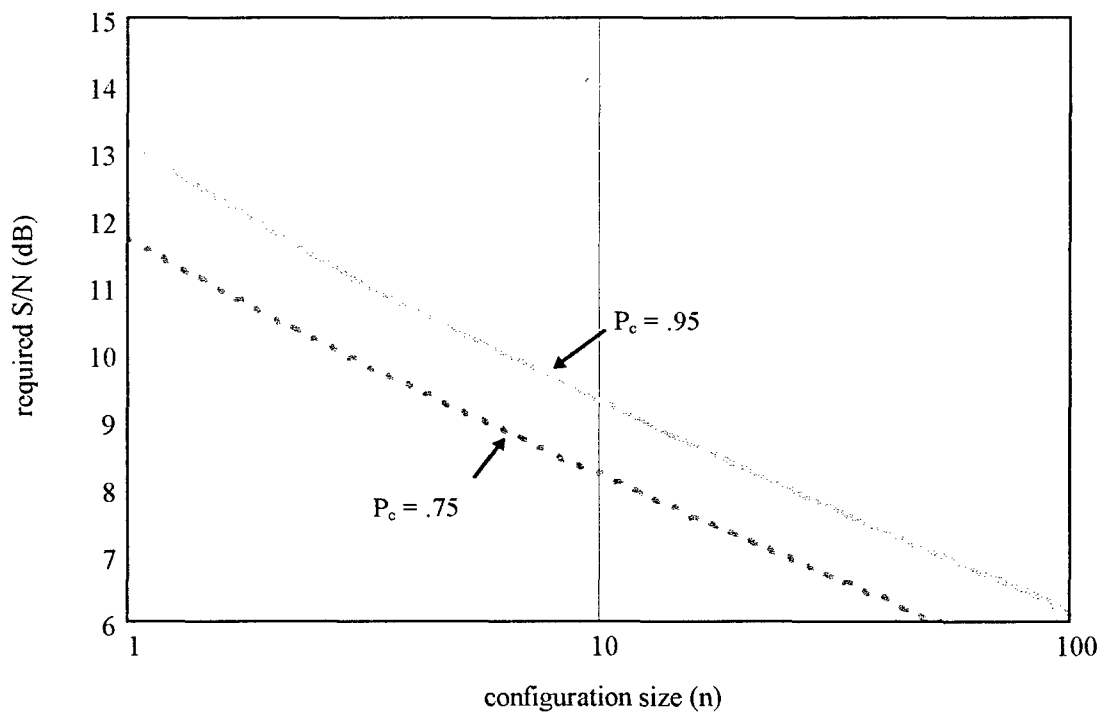
Figure 5.4: S/N as a function of search configuration size. $(P_{fa} = 10^{-5})$

83

The individual radar probability of detection, P_d, need not be great to ensure a good configuration detection probability, P_c. By using the n_s radars independently, the probability of detection, P_d, (and hence the signal to noise ratio) is reduced for a given mission – that is for a given target detection capability which is now given by P_c. *Figure 5.3* and *Figure 5.4* show the decrease in signal to noise for several configuration sizes.

Due to the asymptotic S-curve relationship between signal to noise and probability of detection (*Figure 5.2*), the increase in power-aperture for high detection probability quickly reaches a state of diminishing marginal return. In the same sense, a distributed radar network can reduce the power-aperture required at a single node and maintain a high probability of detection for the network (*Figure 5.5*). This is one the benefits of performing search from a distributed system.

Note that although power-aperture per satellite is reduced, the reduction is not sufficient to offset the total power-aperture increase that results from adding satellites. Thus, the overall system power-aperture increases, though less than it would for linear additions of satellites. The reduction in power-aperture for a distributed operations concept cannot by itself justify the distribution of the SBR AWACS mission. Rather, the reasons for distribution are other cost and performance model elements that result in lower cost – the reduction in power-aperture is a benefit that makes distribution more efficient once the decision to distribute has been made. This should be clear in the context of the metric analysis (Chapter 3) and is only mentioned here to
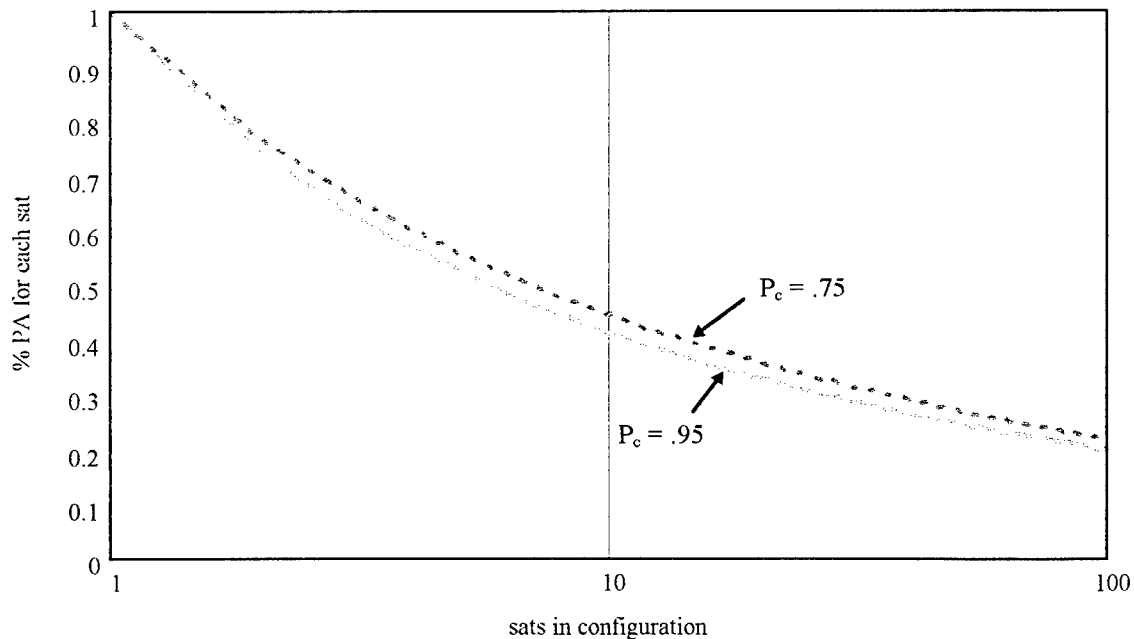
Figure 5.5: Power/Aperture reduction as a function of configuration size (normalized by n=1 configuration). Note that although there is a difference of more than 1 dB between P_c= .75 and .95, the effect on the power-aperture product is negligible compared to the effect of distribution.
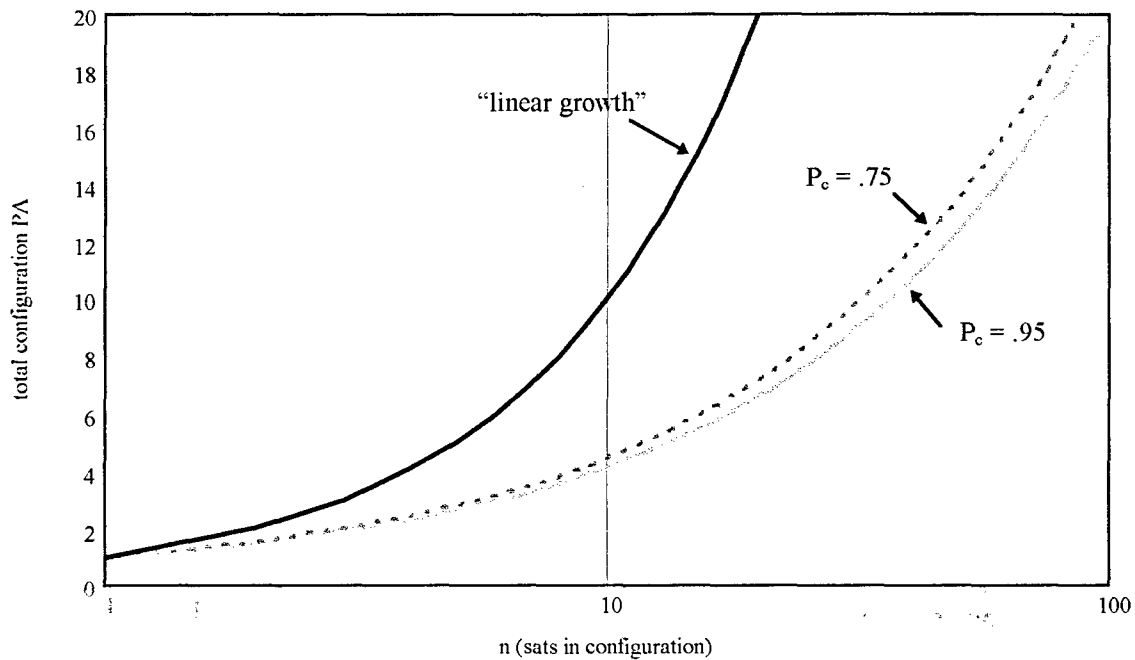
Figure 5.6: Total configuration increase in power-aperture as a function of configuration size (normalized by n=1 configuration size). Linear growth shown for comparison.

again illustrate how the concept of operations can significantly drive the system cost.

It should be noted that the signal to noise can be similarly reduced if a single radar searches n_s times faster. The probability of detection after n_s scans is now given by *eqn 5.10* where P_d should be interpreted as the probability of detection for a single scan. Since the scan rate has increased by a factor of n_s and n_s scans are required to achieve the desired detection capability, the area search rate and the mean time to detection have not effectively changed. This would of course be a much more efficient method for implementing the operations concept since only one satellite is required. As noted above, though, the operations concept does not justify the distribution. Distribution is cost effective for other reasons; the operations concept merely makes distribution to several satellites even more attractive.

5.3 Mean Time to Detection:

Mean time to detection is the average time it takes the distributed configuration to detect a target in the presence of noise. The noise is thermal noise in the individual receivers and does not include clutter return. Radar performance in clutter is analyzed in Chapter 7.

Mean time to detection is a metric for quantifying the performance of a search radar system. We develop it because the traditional search radar performance metrics do not directly apply to the distributed search operations concept and are difficult to size in terms of system requirements. For example, area search rate for a distributed system cannot be meaningfully

85

compared to the area search rate for a single radar. The probability of detection metric does not include a time or rate measure. Mean time to detection can be expressed as a function of these other, more traditional metrics. Mean time to detection is also easily understood physically and leads directly to the establishment of system requirements.

The mean time to detection metric is derived in Section 5.3.1. Reasonable system requirement specifications are established in Section 5.3.2.

5.3.1 Derivation of Mean Time to Detection

The search process can be modeled as two independent negative binomial processes. The negative binomial is a discrete probability distribution that gives the number of trials, y, until the rth successful trial is observed in a process with p probability of success in a single trial. The discrete density is given by

$$P(y) = \binom{y-1}{y-r} p^r (1-p)^{y-r} \qquad eqn\ 5.13$$

where $\binom{a}{b}$ represents the combination of a events taken b at a time – that is, "a choose b". The mean and variance of a negative binomial process are

$$\mu_y = \frac{r}{p} \qquad eqn\ 5.14$$

$$\sigma_y^2 = \frac{r(1-p)}{p^2} \qquad eqn\ 5.15$$

The first negative binomial process is the number of trials (scans) until m detections are made. This is necessary for an m-out-of-n detection scheme. Since the probability of detection in a scan is, P_d, the mean number of scans for m detections is

$$\mu_{scans} = \frac{m}{P_d} \qquad eqn\ 5.16$$

A target detection is declared when there have been m detections of a potential target. Thus the probability of detection for a single radar at any time is now P_d/m. The m detections can occur from a single radar or between radars. In general, the m detections will not occur simultaneously though they should occur within a sufficiently small window so that some given correlation between scanned positions can be achieved. Of course, the correlation is always limited by the resolution of the individual radars. A more detailed discussion of distributed radar correlation is given in section 5.4.

The second negative binomial process is the detection of targets throughout the search area. Let n_{tgt} targets be uniformly distributed in the area to be searched, A. This may represent a theatre, threat area, cruise missile fence, battle field, etc. Although a uniform distribution of

86

targets is likely to be infrequent, there is usually no *a priori* knowledge of target distribution. If there is some knowledge of target distribution within the search area, the area can be subdivided into smaller search areas compatible with the resolution of the known distribution. These sub-areas are then assumed to have a uniform target distribution. The analysis proceeds in the exact manner. For a uniform distribution of targets, the probability of illuminating a target is

$$P_{tgt \in A_f} = \frac{n_{tgt} n_s A_f}{A} \qquad \qquad eqn\ 5.17$$

Each trial in the second negative binomial process is a time unit in which n_s search radars scan their footprint areas, A_f. The mean number of trials until r targets have been illuminated is

$$\mu_{trials} = \frac{Ar}{A_f n_s n_{tgt}} \qquad \qquad eqn\ 5.18$$

The probability of success of a scan is the intersection of two independent probabilities: the probability that a target is in a footprint and the detection probability of the search radar, P_d/m. That is

$$p = \frac{n_{tgt} n_s A_f}{A} \frac{P_d}{m} \qquad \qquad eqn\ 5.19$$

The mean number of trials (time units) until r targets have been detected for a *m*-out-of-*n* detection scheme is then

$$\mu_t = \frac{rmA}{n_{tgt} n_s A_f P_d} \qquad \qquad eqn\ 5.20$$

The mean time to detect r targets, $T_d^{(r)}$, is then the mean number of trials times the time for 1 trial, A_f/ASR, where ASR is the area search rate.

$$T_d^{(r)} = \frac{rmA}{n_{tgt} n_s P_d ASR} \qquad \qquad eqn\ 5.21$$

The mean time to detect all targets is

$$T_d = \frac{mA}{n_s P_d ASR} \qquad \qquad eqn\ 5.22$$

eqn 5.22 gives a convenient method to trade detection probability, area search rate, and detection time in the cost and performance models. Typical SBR area search rates for AWACS tend to be incomprehensible numbers due to their magnitude. Mean time to detection is much easier to understand.

eqn 5.22 is also useful for determining a more optimal operations concept. In typical applications, there is a tradeoff between the range of detection and probability of detection. Targets are generally less threatening at longer ranges which results in decreased risk for a missed detection. By increasing the area search rate at low threat ranges, the probability of detection can

be reduced to a level commensurate with the risk. This increases the assets available for search at higher threat ranges and reduces the mean time to detection for the more threatening targets.

The parameter m can also be adjusted to give a more optimal search profile. False alarms are less desirable at low threat ranges. Similarly, missed detections are more costly at high threat ranges. Thus, m can be increased while searching low threat areas and decreased in high threat areas.

The detection model derived here is simplified in that the r successes are not guaranteed to span the n_{tgt} targets. Since we are sampling with replacement, there is a non-zero probability that the same target will be detected twice while a different target remains undetected. The m-out-of-n detection scheme decreases the probability of this type of model error, but it is still non-zero. Extending the process over time reduces the probability further. Nevertheless, the mean time to detection given by *eqn 5.22* probably slightly under-predicts the actual time required to detect all targets.

5.3.2 Mean Time to Detection: reasonable requirements for SBR

Since mean time to detection is the primary performance metric, it is also the primary system driver. Mean time to detection is the single performance specification that scales the entire system. Thus, it is important to establish a good, reasonable requirement for mean time to detection.

Targets must be detected quickly enough to allow adequate time to respond to the threat. The mean time to detection requirement then depends upon both the threat and the threat response. *Figure 5.7* depicts the threat detection and response process. The goal of the process is to react to the threat before the threat can carry out its objective. For example, suppose the threat is a flight of enemy fighter-bombers with the objective of bombing an airfield. The targets must be detected soon enough to allow time to decide an appropriate response, command the response action, and allow the action to take effect before the fighter-bombers can make their bombing run. A typical response may be to dispatch fighters or launch SAMs. If fighters must be scrambled, the response time could be minutes. If SAMs are fired, the response time could be on the order of the missile flight time (this is assuming missiles have adequate coverage of the targets). The command process is the most uncertain. If the scenario is an air battle, the command response is simply the target assignment and may require only seconds. In other scenarios, particularly when the threat is unexpected, the command response may be much longer. Thus, determining reasonable mean time to detection requirements also depends on the scenario.

The threat detection and response process is well characterized by the OODA loop: Observe, Orient, Decide, Act [37]. The OODA loop is a major tenet of Information Warfare (IW) and "describes a single iteration of the cycle proceeding from data acquisition, through information integration and decision making, to enaction of a response [38]". Operating inside or disrupting the enemy's OODA loop is the primary objective of IW. Target detection is analgous to the observe task; threat processing is the orient task; determine response is the decide task; and

target detection	→	process threat determine response command response	→	respond to threat

mean time to detection: $T_d = 10$ sec	processing time & command decision: seconds to minutes	response time: seconds to minutes
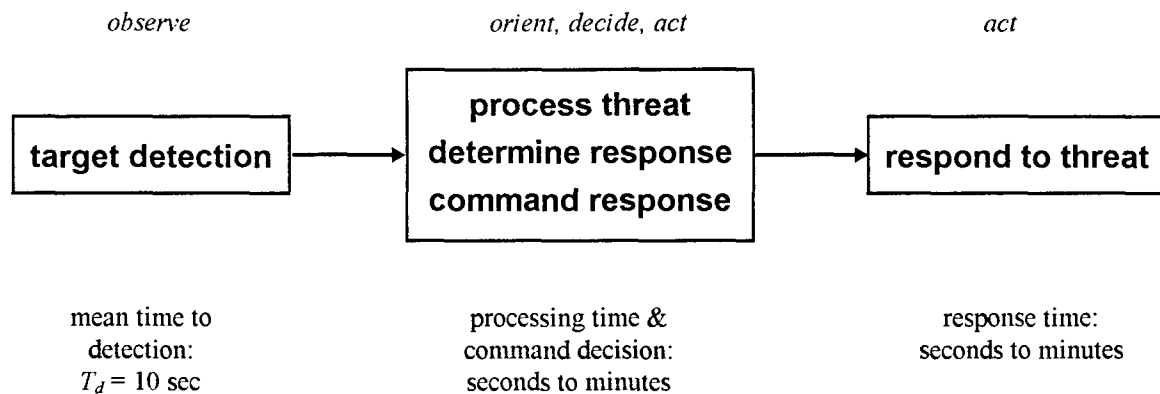
Figure 5.7: Threat detection and response process.

the command and response is the act task of the OODA loop.

None of the later steps in the threat detection and response can take place until the targets are detected. Thus, decreasing mean time to detection is important to start the other steps sooner. However, if it requires minutes to respond to the threat, cutting mean time to detection in half - to 5 sec - does not significantly reduce the total threat response time while the system cost has likely doubled.

In dialogue with Air Force Space and Missile Center, we have established a 10 sec mean time to detection requirement for the purpose of sizing the system [39]. This could actually represent an average mean time to detection over a theatre which includes high threat regions of 1 sec mean time to detection, moderate threat regions of 10 sec mean time to detection, and low threat areas of 1 min mean time to detection. It is interesting that a 10 sec requirement corresponds to an elbow in the system cost curve as a function of the mean time to detection. Detection times less than 10 sec result in significant increases in system cost and detection times less than 10 sec result in only marginal savings (Section 8.5).

Recall that there is some flexibility in mean time to detection, even after the system is deployed with a fixed capability. Mean time to detection can be reduced by decreasing the area to be searched. Thus, if it is necessary to detect targets within 1 sec in an air battle, but the system was designed to a requirement of 10 sec, the search area can be reduced by a factor of 10. Of course, this is not without a price, since the mean time to detection must be increased by a factor of 10 somewhere else in the theatre. Nevertheless, there is extensive flexibility to make the mean time to detection in a given area commensurate with the threat in the area.

5.4 Correlation between distributed radars.

The distributed operations concept reduces the probability of detection of a single satellite by distributing the detection function throughout the configuration. This leads to a decrease in
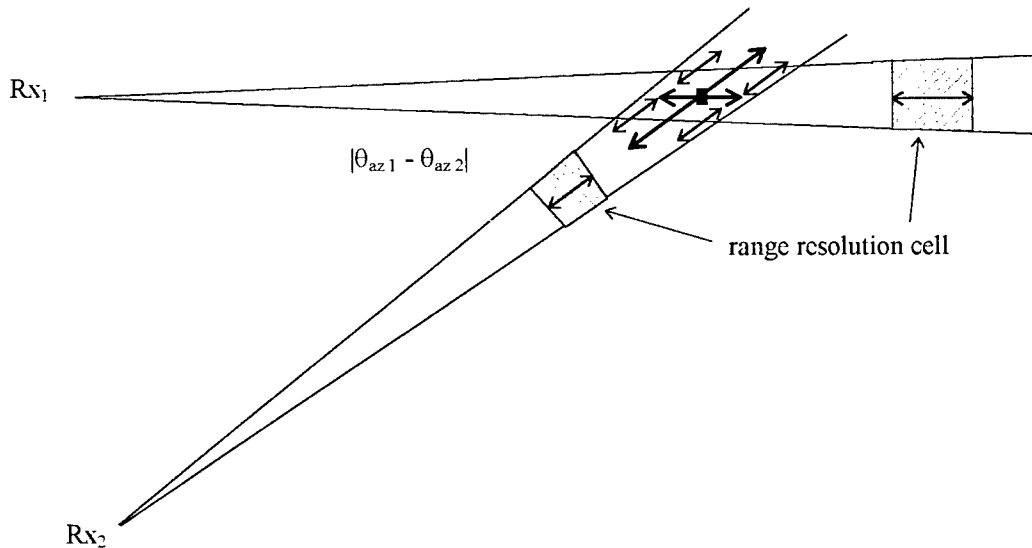
Figure 5.8: Range resolution relationship for two receivers viewing the same target. Rx_2 must increase the range at which it must hypothesis test to verify a target detected by Rx_1.

the signal to noise ratio required on a single satellite (*Figure 5.4*). Although the probability of false alarm on a single satellite does not change (the detection threshold is unchanged), the configuration probability of false alarm does increase in the same manner that the configuration probability of detection was increased. To avoid this, an *m*-out-of-*n* detection scheme is employed: *m* pulses out of *n* total pulses must exceed the detection threshold before a target detection is declared. For a distributed system of search radars, this *m*-out-of-*n* detection scheme can take place between satellites. That is, *m* pulses from any combination of satellites must exceed the threshold out of a possible *n* times for a declared detection. This requires the correlation of returns between the satellites: a potential positive detection from one satellite must be matched (correlated) with potential positive returns from other satellites in the configuration. The correlation error depends upon the geometry between the satellites and target. The correlation process seeks to match targets of a given velocity in absolute position.

Correlation of position is the most obvious target matching procedure. If two radars have potential target returns from the same location, the return is probably due to the same target. Note that correlating absolute position of the target requires knowing both radar positions and their beam aspect on the target. Radars can report the position of a target with respect to the azimuth of the transmitted beam and range from the radar to the target. Radar range resolution depends upon the width of the autocorrelation of the transmitted waveform and is finite since the waveform must have some pulse width to transmit energy. Correlation of a target between two radars in absolute space then involves the convolution of two range resolution cells. For example, in *Figure 5.8*, radar 1 detects a target with a finite range resolution. Radar 2 seeks to corroborate the detection by hypothesis testing the existence of a target in the range bin that corresponds to the same location in absolute space. The same location in absolute space is the convolution of the

90
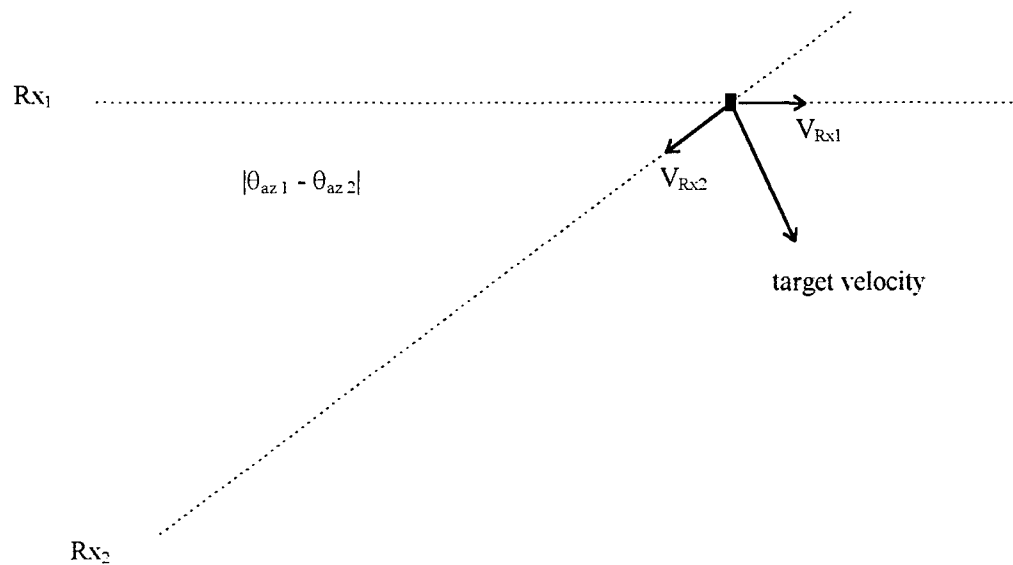
Figure 5.9: True target velocity and radial velocity measured by Rx₁ and Rx₂.

resolution cells of radar 1 and radar 2. Thus, radar 2's hypothesis testing must be done in a range bin that is larger than its range resolution cell. This results in an increased processing load and the chance that a return from another target will fold into the detection of the initial target. Over time, the estimates will converge in absolute space. This also increases the processing load since the position must be processed over time, but it makes correlation between radars possible.

A radar estimates the velocity of a target relative to the radar platform by the Doppler
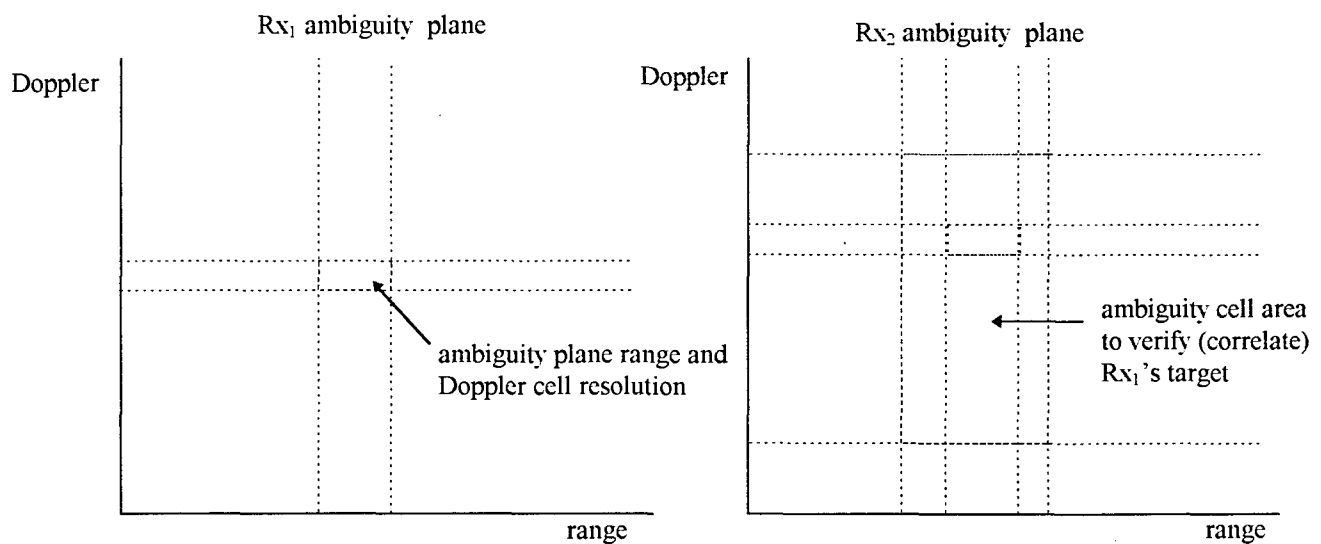
Figure 5.10: Ambiguity diagram for two receivers and the corresponding areas of correlation. Rx₁ detects a potential target and Rx₂ attempts to verify target. The hypothesis test must take place in a larger region of the range-Doppler plane.

shift of the returned signal. A measurement by one radar is insufficient to give the absolute velocity. A second measurement with an aspect different from the first will give an estimate of the absolute velocity in space. The resolution of the estimate depends on the frequency resolution of the transmitted waveform. Because the ambiguity in range and Doppler is conserved (range and frequency are Fourier conjugates), it is useful to analyze target correlation in the range-Doppler or ambiguity plane. *Figure 5.10* demonstrates the increase of area in the ambiguity plane that radar 2 must hypothesis test to correlate a return with a potential target detected by radar 1. Again, this increases the processing load, but does not preclude the feasibility of the distributed detection operations concept.

6. Constellation Design and Geometry

Constellation design determines the orbital elements of the satellites that satisfy the coverage constraints and requirements. Constellation size is a significant driver of the system cost – a decrease in the size required to meet coverage requirements is certainly desirable. The primary goal of constellation design is then to optimize the design for given coverage requirements. Additionally for the SBR system architecture analysis, the constellation design process determines several parameters required in the cost model including constellation size, slant range, duty and mission cycle, and coverage angles. Ideally, the constellation design process should lead to a constellation that is optimal in several senses. First, the constellation should consist of the minimum number of satellites capable of performing the mission. Second, the coverage should be the best achievable with the number of satellites in the constellation. Metrics of coverage include revisit time, the extent of multiple coverage, viewing angles, and slant range between targets and satellites. Range is particularly important in a radar system since the radar range equation (and hence the power aperture product) depends on the second power of range for search radars and the fourth power of range for track radars. As a third measure of optimality, coverage should be concentrated in areas that are likely to be of interest and minimal in regions that are of little interest. Other constellation design considerations are eclipse time, trapped radiation (Van Allen belts), orbit perturbations and subsequent orbit maintenance, orbital debris, and atmospheric drag.

A generalized approach to constellation optimization for a general set of coverage requirements has not yet been created. In part, this is due to the wide range of ways in which "optimality" may be measured. Several were listed above. As well, every mission is driven by unique characteristics that affect the constellation in different ways. Ultimately, though, the constellation dynamics and coverage relationships are extremely complex and nonlinear which makes optimization hard. Finding constellations that satisfy the requirements, however, is not generally difficult -- determining optimality is. The only currently known approach to optimizing the general constellation problem is a search over the parameter space. Two methods for searching the parameter space exist: exhaustive search and a genetic algorithm approach. Both methods are limited to the resolution of the parameter space. Exhaustive search techniques are practically limited by computational rates. For example, we initially implemented a coarse exhaustive search over eight parameters (the six orbital elements, number of planes, and phasing between satellites) that required computational runs on the order of days on an SGI workstation. This approach is too unwieldy to be used as a flexible design process.

A more common approach to constellation design for first order system analysis is to use one of the set of constellations for which coverage is simple to calculate or well known. Constellations in this class include Walker delta patterns [40,41,42], and Draim's eccentric orbit constellations [43,44,45]. Walker constellations attempt to equally distribute coverage about the Earth. Satellites in eccentric orbits spend a significant portion of their orbit near apogee which

can be used to extend the satellite's coverage over a given area. This is typical of Draim's constellations. A similar set of constellations to fill the nadir hole for radar coverage constraints (Section 6.1), are given by Rider [46,47,48].

Using constellations from these classes amounts to a heuristic approach because the constellations are generally built from a set of rules. Often, the mean coverage can be derived analytically. Thus, we call this technique for first order constellation design the analytically derived heuristic approach. For the SBR mission, Rider's approach is closest to meeting the desired coverage requirements. The problem with Rider's approach, however, is that it fills coverage holes at high latitudes and gives overdesigned coverage in polar regions. Thus, a new class of constellations is needed. We develop a set of *zenith hole constellations* that give continuous global coverage within a desired range of latitude. This approach falls in the class of analytically derived heuristics.

A final tool in constellation design is the analysis of the constellation through simulation. Simulation was useful for verifying the results developed for the set of zenith hole constellations. We also implemented the constellation simulation software in a quasi-monte carlo simulation with various, random constellation parameters. We then generated some empirical results by multivariable regression that complemented the zenith hole analytically derived heuristics. The constellation design approaches that we implemented are summarized in *Table 6.1*.

Section 6.1 discusses coverage and defines the satellite-target geometry. The zenith-hole design approach is detailed in Section 6.2 and the constellation sets are described in 6.3. Section 6.4 details the simulation methodology. The joint results of the constellation design process, including the altitude coverage function, are summarized in Section 6.5.

Table 6.1: SBR constellation design approach summary

	Constellation Design Approach	Result
1	an exhaustive search of the constellation parameter space (orbital elements including eccentricity, number of planes, plane spacing, satellites per plane, and inter-plane phasing) with a range squared and number of satellite metric over a diurnal period	good results but implementation was extremely slow and too inflexible to test a large range of constellations; feasible only with a small target set of interest
2	constellation simulation on a discretized, weighted globe	implemented for 3, 4, and 6 hr orbits; good results but slow with no gauge of optimality
3	zenith hole minimum satellite coverage (analytically derived heuristics)	implemented for 1.65, 1.75, 2, 3, 4, and 6 hr orbits; good results are fast and easy; with visual implementation is easy to gauge "optimality"

6.1 Coverage and Satellite-Target Geometry

A coverage constraint is a geometric relationship between the target and the satellite that must be satisfied for the satellite to perform its mission at the target. For radar, the coverage constraint is a maximum and minimum elevation angle, ε, at the target. Elevation is the angle between the target local horizon and the satellite direction (*Figure 6.1*). Below the minimum elevation, the transmitted signal is subject to atmospheric ducting and fading. Thus, the minimum elevation constraint is a propagation constraint. Above the maximum elevation, the target doppler shift is too low and the clutter return is too high to reasonably discriminate the target from the clutter. The maximum elevation constraint is a clutter constraint. For all constellation designs presented here, a minimum elevation of 5 deg and maximum elevation of 50 deg is used [49].

The coverage constraints result in a *zenith hole* from the perspective of the target (*Figure 6.2*). The satellite must be within two concentric cones defined by the constraint elevation central angles. When projected onto the surface of the Earth, any satellite with a subsatellite point (SSP) between the central angles associated with the maximum and minimum elevation constraints will be able to cover the target. For SSP's with central angles less than the central angle of the maximum elevation, the satellite does not meet the maximum elevation constraint. Thus, the maximum elevation constraint forms a *zenith hole* from which satellites cannot cover the target even though they are within view. From the satellite perspective, coverage is in the form of a nadir hole with uncoverable targets above a maximum grazing angle.
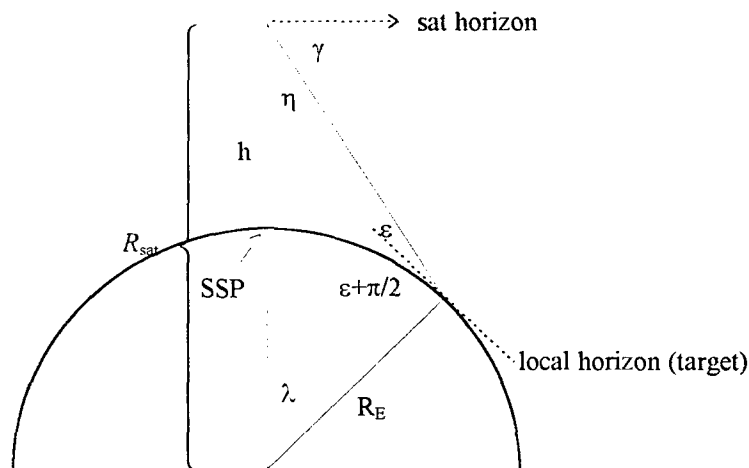
Figure 6.1: Coverage parameters: ε is the elevation angle of the satellite; λ is the Earth central angle; η is the nadir offset angle; γ is the satellite grazing angle; R_E is the radius of the Earth; R_{sat} is the orbital radius of the satellite; h is the altitude of the satellite. The parameters can be related using the law of sines and cosines for plane triangles.
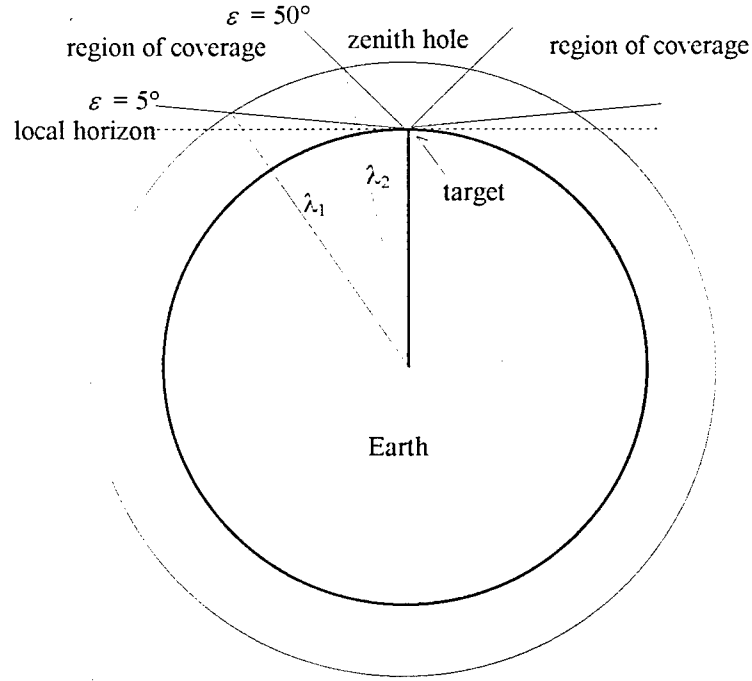
Figure 6.2: Zenith coverage hole geometry. The hole results from the radar elevation constraints but can be described from the geographic location. The difference between the central angles defines the size of the coverage region projected on the surface of the Earth. See Figure 6.3.

Define λ_1 as the central angle associated with the minimum elevation constraint. Similarly, λ_2 is the central angle associated with the maximum elevation constraint. Besides elevation, the central angle, λ, is a function of the orbital altitude, h. From a combination of the law of sines and cosines, the relationship is given by and plotted as a function of altitude in *Figure 6.3*.

$$\lambda = \frac{\pi}{2} - \varepsilon - \sin^{-1}\left(\frac{R_E}{h + R_E}\cos(\varepsilon)\right)$$

eqn 6.1

There is no need for a quadrant check for the inverse sine function in *eqn 6.1* since the quantity is the nadir offset angle and is always less than $\pi/2$. By the law of sines, nadir offset and elevation are related by

$$\frac{\sin(\eta_{tgt})}{R_E} = \frac{\cos(\varepsilon_{tgt})}{h + R_E}$$

eqn 6.2

and slant range to target is

$$R_{tgt} = \frac{\sin(\lambda_{tgt})}{\cos(\varepsilon_{tgt})}(h + R_E)$$

eqn 6.3

Orbital altitude, radius, and period are used interchangeably to define the same orbit. Altitude is orbital radius less the radius of the Earth. Period is a function of the orbit's semimajor
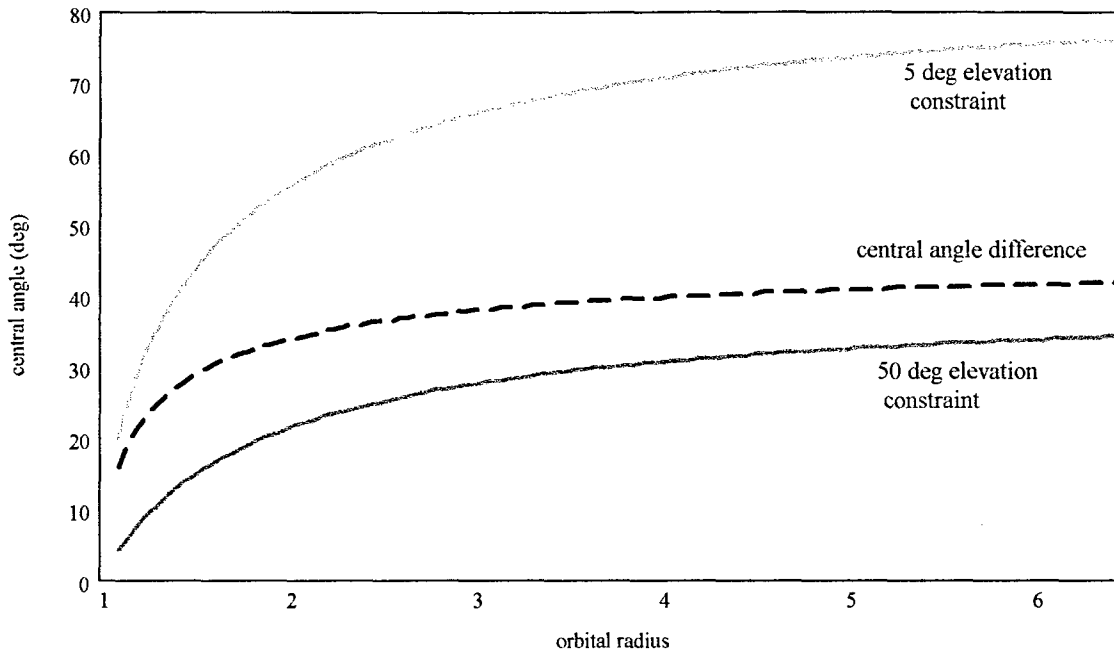
Figure 6.3: Central angles of the minimum and maximum elevation constraints as a function of orbital radius. The difference represents the angle between the concentric cones from within which a satellite can view the target. DU is a distance unit equal to one Earth radii.

axis, a, which is equal to the radius for circular orbits. The relationships for select values are given in *Table 6.2*.

$$Per = 2\pi\sqrt{\frac{a^3}{\mu}}$$

eqn 6.4

Table 6.2: Altitude, orbital radius, and orbital period relationship for selected values.

altitude (km)	altitude (nm)	radius (DU)	period (hr)	altitude (km)	altitude (nm)	radius (DU)	period (hr)
300	162	1.05	1.51	6,003	3,242	1.94	3.81
503	271	1.08	1.58	6,576	3,551	2.03	4.08
597	322	1.09	1.61	7,892	4,261	2.24	4.72
709	383	1.11	1.65	8,646	4,668	2.36	5.09
842	455	1.13	1.70	9,471	5,114	2.49	5.52
1,000	540	1.16	1.75	10,370	5,602	2.63	6.00
1,188	641	1.19	1.82	11,740	6,341	2.84	6.75
1,410	762	1.22	1.90	13,290	7,178	3.08	7.63
1,675	904	1.26	2.00	14,140	7,637	3.22	8.13
2,202	1,189	1.35	2.20	17,180	9,274	3.69	10.00
2,412	1,302	1.38	2.28	20,520	11,080	4.22	12.20
3,171	1,712	1.50	2.58	23,220	12,540	4.64	14.09
4,169	2,251	1.65	3.00	26,290	14,190	5.12	16.33
5,003	2,701	1.78	3.36	31,660	17,100	5.96	20.53
5,480	2,959	1.86	3.57	35,840	19,350	6.62	24.00

6.2 Zenith Hole Constellation Design:

The zenith hole constellation design approach uses the projected elevation constraints and satellite ground traces to determine empirically the number of planes and satellites per plane required to provide continuous global coverage within a specified range of latitudes. When projected on the celestial sphere (of unit radius) or the surface of the Earth, the maximum and minimum elevation coverage constraints are two concentric circles whose diameter is governed by the central angles. The central angles are a function of the target elevation limits and the orbital altitude being considered (*Figure 6.4*). Since the central angles (and subsequently the concentric circle constraints) change with orbital radius, the zenith hole constellation design approach will only work for circular orbits. It is worth noting that eccentric orbits do not improve the SBR mission coverage. This is demonstrated in Section 0.

We now show, by means of an example, why it is desirable to have more than one orbital plane within the region of coverage (between the concentric circles in *Figure 6.4*). Take an orbital radius of 1.655 DUs which corresponds to a 3 hour orbit. At this altitude, the central angles are 47 and 17 degrees. Let there be one orbital plane within the coverage region at the worst case target longitude -- that for which the satellite trajectory intercepts the target latitude and longitude. (The longitude is fixed in inertial space and will move west on the surface of a rotating Earth at 15 deg/hr.) With this geometry, two 30 deg sections of the orbit separated by 34 deg lie within the coverage region. Since each satellite provides 60 deg of coverage in an orbit, 6
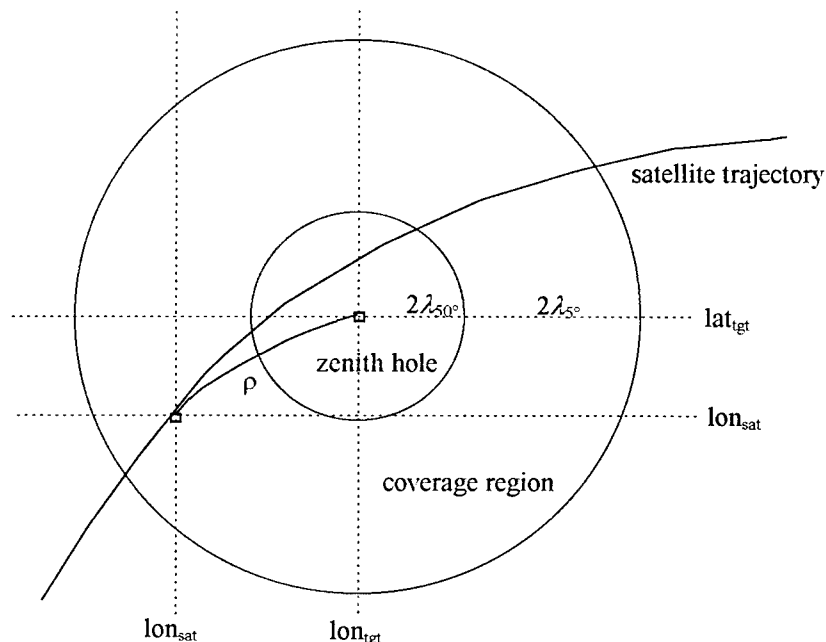
Figure 6.4: Coverage constraints and satellite-target geometry projected onto the surface of the Earth. $2\lambda_{5°}$ is the angular diameter of the coverage cone due to the minimum elevation constraint. Similarly $2\lambda_{50°}$ is the angular diameter of the maximum elevation constraint. The angular separation between the satellite and target, ρ, is a celestial sphere central angle.
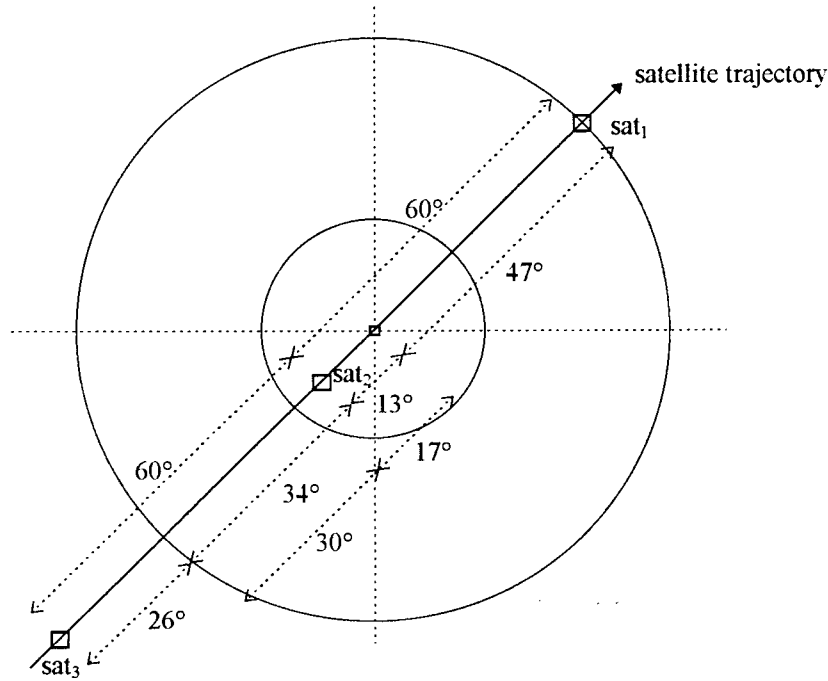
Figure 6.5: 2 hr orbit with 60 deg satellite phasing showing the 26 deg coverage gap when sat₁ leaves the coverage region and sat₂ is still within the zenith hole. This demonstrates the nature of the zenith hole – 60 deg of the orbit lies within the coverage region but 60 deg phasing between satellites will not give continuous coverage.

satellites should provide continuous coverage. The zenith hole makes it more complicated however, and phasing between the satellites must be checked to ensure that the target is always covered. *Figure 6.5* shows that 60 deg phasing between satellites is *not* sufficient to provide continuous coverage. The actual maximum allowable phasing is 47 deg (*Figure 6.6*). Since the number of satellites in a plane must be discrete, 8 satellites at 45 deg phasing gives the desired continuous coverage. Since several orbital planes will be required to provide continuous global coverage, the satellite per plane density should be as low as possible. This can be done if the orbital planes are spaced such that the coverage region of a target contains more than one orbital plane. For the 1.655 DU example, only 2 satellites are required per plane if 4 orbital planes are within the coverage region which reduces the total number of satellites required for *global* coverage.

A plane crossing is defined as the intersection of two planes. If a plane crossing occurs within the limits of λ_1, then both planes will lie within the coverage region. The number of plane crossings, N_{pc}, is the twice the combination of the number of planes taken two at a time.

$$N_{pc} = \frac{N_p!}{\left(N_p - 2\right)!}$$

eqn 6.5

where N_p is the number of planes. The longitude between plane crossings at a constant latitude is
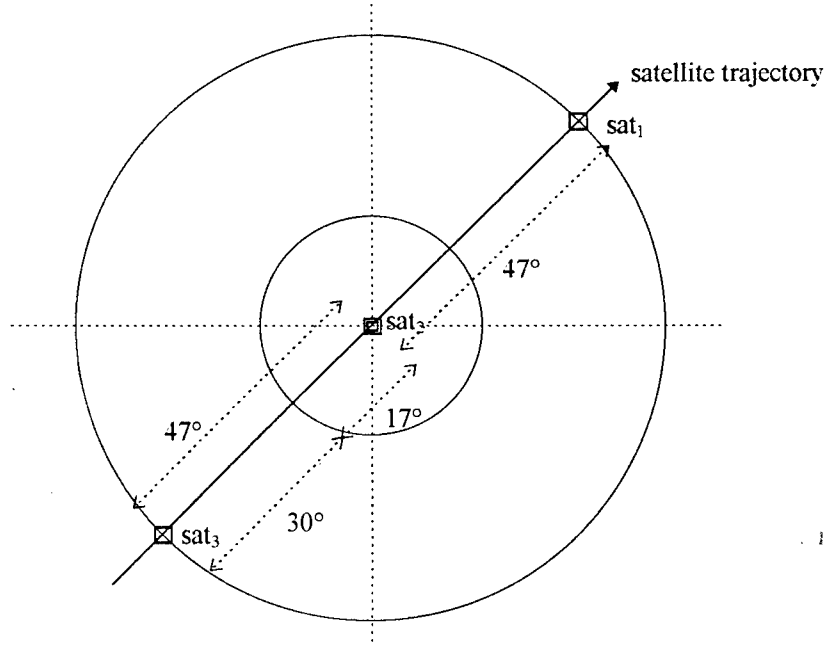
Figure 6.6: 2 hr orbit with 47 deg satellite phasing that provides continuous coverage of the target.

$$\Delta\Omega_{pc} = \frac{2\pi}{N_p} \qquad\qquad eqn\ 6.6$$

The quotient of N_{pc}/N_p gives the number of different latitudes at which plane crossings occur. As the number of planes increases, the latitude of the plane crossings increases (with the exception of the equatorial plane crossings, if they exist). The latitude of the plane crossings will never be greater than the inclination of the orbit.

One of the zenith hole approach heuristics is that the number of planes, N_p, should be chosen such that the longitude between the plane crossings is on the order of $2\lambda_l$, the central angle associated with the minimum elevation constraint. This ensures that at least one plane crossing occurs within the coverage region and at least two planes are within the coverage region. Depending on the latitude of the target and the orbital altitude, there will often be even more planes within the region of coverage than predicted by the number of plane crossings.

At low altitudes λ_l is small and to obtain coverage for moderate to high latitudes requires steeply inclined orbits. The number of planes required also increases and the plane crossings get pushed to relatively high latitudes. It is more efficient for these orbital altitudes to add a second tier of planes at a lower inclination to give plane crossings at the lower altitudes. For 5 deg and 50 deg elevation constraints, this bifurcation approximately occurs for orbits below 2 hr or 1600 km altitude orbits. This two tiered approach unavoidably results in over designed coverage at the low latitudes. Without two tiers, however, the constellation is over-designed for high latitudes. It is generally best to first design the constellation for the high latitude targets and then use shallow inclination orbits to fill in the low latitude coverage. Usually only a few relatively sparse shallow inclination planes are required.

Figure 6.7: Satellite latitude and longitude geometry in inertial space. Longitude as depicted and defined here is more commonly called right ascension. *Similarly, latitude is more commonly known as* decl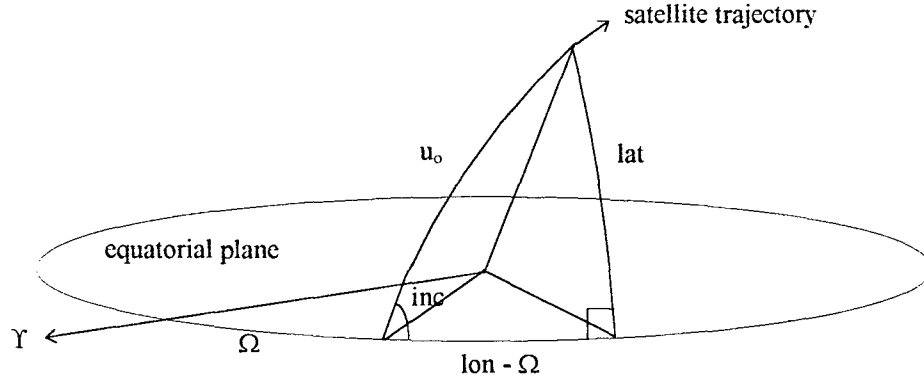ination. ϓ *is the first point of Aries (historic Vernal equinox – see footnote page 102), the principal axis of the inertial coordinate system;* Ω *is the longitude of the ascending node;* lat *is the latitude of the satellite on the celestial sphere at time* t_o *and* lon *is the longitude of the satellite on the celestial sphere at* t_o; inc *is the inclination of the satellite plane to the celestial equator;* u_o *is the argument of latitude at* t_o.

The zenith hole approach does not account for the rotating earth. Rather, everything is described in inertial coordinates and then the Earth reference can be calculated if desired. This is normally not required, however. Since the constellation must provide continuous global coverage, the diurnal variation in coverage due to Earth rotation is captured by requiring coverage for all longitudes. The latitude and longitude as defined in *Figure 6.7* are inertial latitude and longitude, more commonly known as declination and right ascension. We use these interchangeably since the Earth rotation need not be modeled. In inertial coordinates, the projected ground trace (in inertial latitude and longitude) of a circular orbit depends only on the inertial longitude since the longitude of the ascending node. To demonstrate this, consider the spherical triangle in *Figure 6.7*. By the law of sines for spherical triangles (*inc* ≠ 0)

$$\frac{\sin(lat)}{\sin(inc)} = \frac{\sin(u_o)}{1} \qquad \qquad eqn\ 6.7$$

where *lat* is the latitude of the satellite on the celestial sphere at time t_o; *inc* is the inclination of the satellite plane to the celestial equator; and u_o is the argument of latitude at t_o. If *inc* is 0 then *lat* is 0. Note that by convention, longitudes are measured in the equatorial plane and arguments are measured in the orbital plane. The argument of latitude replaces the argument of perigee and true anomaly for circular orbits. Additionally, by the law of cosines for spherical triangles

$$\cos(u_0) = \cos(lon - \Omega)\cos(lat) + \sin(lon - \Omega)\sin(lat)\cos(\pi/2) \qquad eqn\ 6.8$$

where Ω is the longitude of the ascending node; and *lon* is the longitude of the satellite on the celestial sphere at t_o. All longitude measurements are angles between the intersection of the local

101
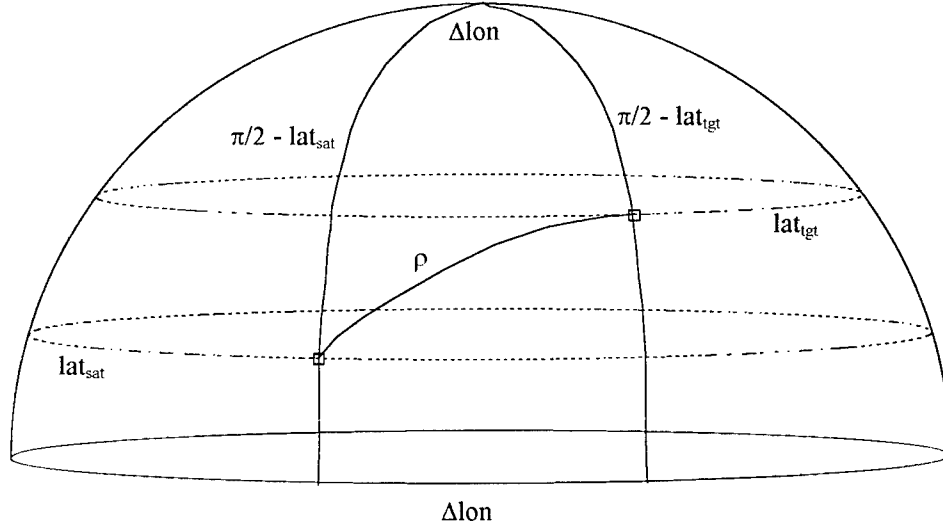
Figure 6.8: Spherical geometry for finding the angular distance between the satellite and target.

meridian with the celestial equator and the direction of the Vernal equinox[1].

Squaring eqn 6.7 and eqn 6.8 and eliminating u_o between them and then solving for lat gives (inc $\neq$ 0)

$$lat(lon) = \begin{cases} \sin^{-1}\left[\sqrt{\dfrac{\sin^2(inc)\sin^2(lon-\Omega)}{1-\sin^2(inc)\cos^2(lon-\Omega)}}\right] & lon \leq \pi + \Omega \\ -\sin^{-1}\left[\sqrt{\dfrac{\sin^2(inc)\sin^2(lon-\Omega)}{1-\sin^2(inc)\cos^2(lon-\Omega)}}\right] & lon > \pi + \Omega \end{cases} \qquad eqn\ 6.9$$

eqn 6.9 is the projection of the satellite orbit onto the celestial sphere. Plotting *lat(lon)* against *lon* from 0 to 2π will give the satellite ground trace for a non-rotating Earth.

The angular distance (on the unit celestial sphere) between the satellite and the target is given by

$$\rho = \cos^{-1}\left[\sin\left(lat_{sat}\right)\sin\left(lat_{tgt}\right) + \cos\left(lat_{sat}\right)\cos\left(lat_{tgt}\right)\cos\left(lon_{sat} - lon_{tgt}\right)\right] \qquad eqn\ 6.10$$

If $\rho < \lambda_1$ and $\rho > \lambda_2$ then the satellite lies within the coverage region of the target. *eqn 6.10* is important for determining and implementing the zenith hole constellation design process.

If the two points of latitude and longitude are constrained to lie in the orbital plane as in *Figure 6.9*, then the angular distance between the points is also the difference in true anomaly, δv, for the orbit.

1. The Vernal equinox is commonly called the first point of Aries, which was the direction of the Vernal Equinox 2000 years ago when its position was first determined by the Greeks. Since then, the Vernal equinox has precessed through Pisces and is on the verge of entering Aquarius.
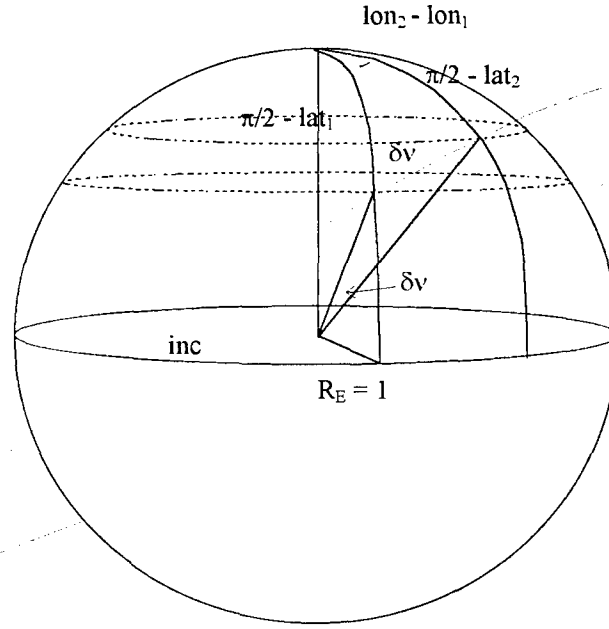
$$\text{lon}_2 - \text{lon}_1$$

Figure 6.9: Relationship between 2 points in an orbit and the true anomaly through which the satellite has passed.

$$\delta v = \cos^{-1}\left[\sin\left(lat_1\right)\sin\left(lat_2\right) + \cos\left(lat_1\right)\cos\left(lat_2\right)\cos\left(lon_2 - lon_1\right)\right] \qquad eqn\ 6.11$$

Having defined the constellation and coverage geometry, the zenith hole constellation design process is straightforward. It helps to do much of the design graphically with the aid of a computer to generate the satellite ground trace and the coverage constraint circles on the surface of the unit celestial sphere.

Zenith Hole Constellation design approach:

1. Specify the maximum latitude for which continuous coverage is desired.

2. For the given orbital altitude, determine the Earth central angles λ_1 and λ_2 that correspond to the elevation constraints.

3. Determine the inclination and number of orbital planes such that the highest latitude plane crossing is at the maximum target latitude minus the central angle λ_1. The ground traces can be projected on the celestial sphere with *eqn 6.9*.

4. Determine the number of orbital planes required to give a plane crossing within λ_1 for any longitude and to the maximum latitude, *eqn 6.5*, *eqn 6.6*. The ground traces can be projected on the celestial sphere with *eqn 6.9*.

5. Repeat steps 3 and 4 until the number of planes is minimized. For orbits with period less than 2 hours, a two tiered set of orbital planes is required.

6. From the ground trace and elevation constraint plots, identify the latitudes and

longitudes with the most limited coverage – those points with the smallest portion of ground trace within the central angle concentric circles. These are typically high latitude targets or points for which the concentric constraints do not contain a plane crossing. For these, determine the intra-plane phasing required to provide continuous coverage.

7. Check the condition that the product of the intra-plane phasing and the number of satellites is at least 2π. If not, the number of satellites per plane must be increased.

8. Check the condition that the product of the inter-plane phasing and the number of planes must be at least 2π. If not, there are insufficient satellites to provide global coverage – repeat step 3 with a different plane arrangement.

9. The number of satellites required for coverage is the product of the number of satellites per plane times and the number of planes. Iterate through the process to minimize the number of satellites. Various plane arrangements should be tried from step 3. In some instances, increasing the number of planes and decreasing the satellites per plane will decrease the overall constellation size.

The zenith hole design process cannot guarantee an optimal design since there is no criteria by which to stop the iteration process. For a given altitude, though, it usually only requires a few cycles to get a good estimate for what an upper bound on the optimal constellation must be.

The zenith hole design process could be modified to give constellations that satisfy a requirement for a maximum duration coverage gap instead of continuous coverage. Step 6 above would be modified so that the number of satellites per plane never exceed the coverage gap. This approach is more sensitive to the inter-plane phasing of satellites which would then have to be determined.

6.3 Zenith Hole Constellations for SBR

The zenith hole heuristics outlined above were used to create a set of constellations of circular orbits with periods of 1.65, 1.75, 2, 3, 4, and 6 hours. Although each constellation was designed to give continuous, global coverage between ± 75 deg latitude, short coverage gaps on the order of a few minutes exist for some constellations at some latitudes. The gaps could have been compensated for by increasing the size of the constellation, but the next increment in constellation size would have given significantly overdesigned coverage. For distributed system architectures, these short gaps could be easily filled by one of the other satellites in the configuration. The coverage specification for a n_s size configuration (n_s distribution) requires n_s satellites continuously visible everywhere on the globe. A short coverage gap of one satellite at some latitude means that n_s -1 satellites are visible for the duration of the gap. This is not a significant coverage problem.

The set of zenith hole constellations are summarized in *Table 6.3* and the ground tracks are plotted in *Figure 6.10* through *Figure 6.15*. The constellation coverage statistics are summarized in *Table 6.6* and plotted in *Figure 6.29* through *Figure 6.34*. Coverage as a function of orbital radius is plotted in *Figure 6.20*.

Table 6.3: Zenith hole constellations.

period (hr)	radius (DU)	altitude (km)	total sats	planes	sats per plane	plane inc
1.65	1.11	700	40	8	4	65°
				4	2	20°
1.75	1.155	1000	33	6	24	67°
				3	3	30°
2	1.263	1700	24	6	4	55°
3	1.655	4200	12	6	2	45°
4	2.005	6400	9	3	3	45°
6	2.627	10,400	7	7	1	35°

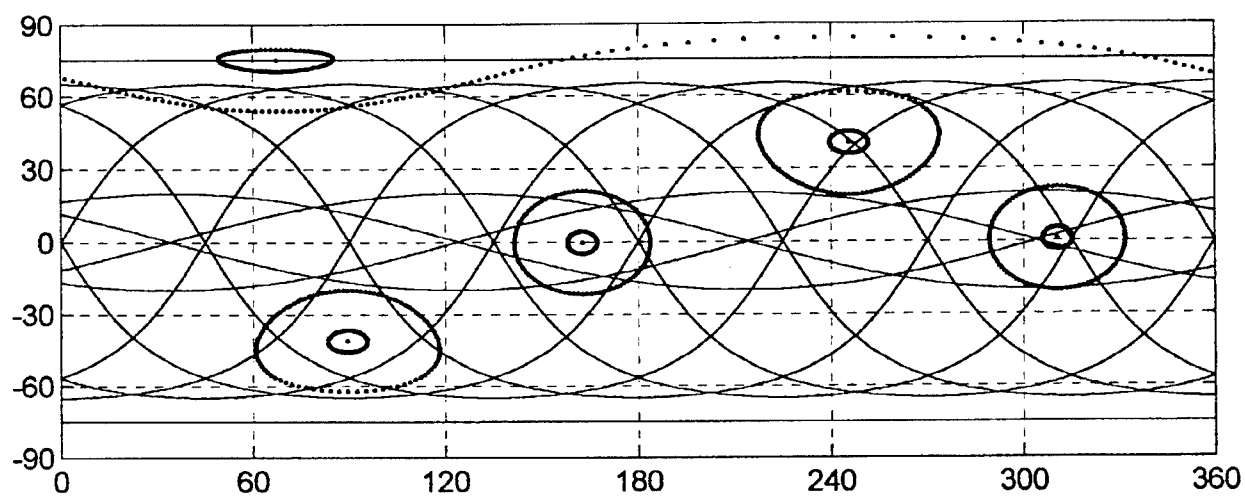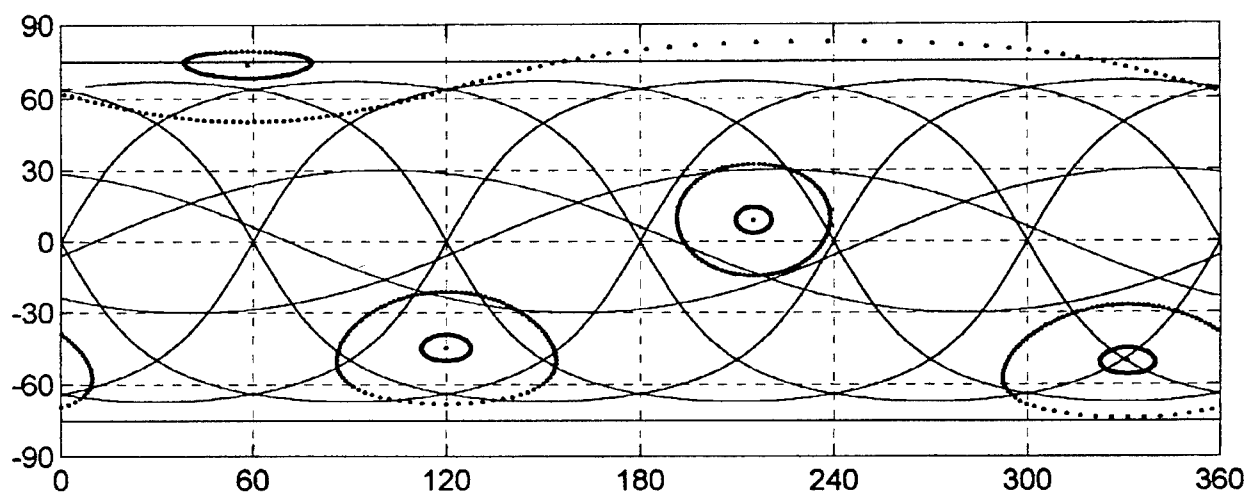*Figure 6.10: 1.65 hour zenith hole constellation ground track.*

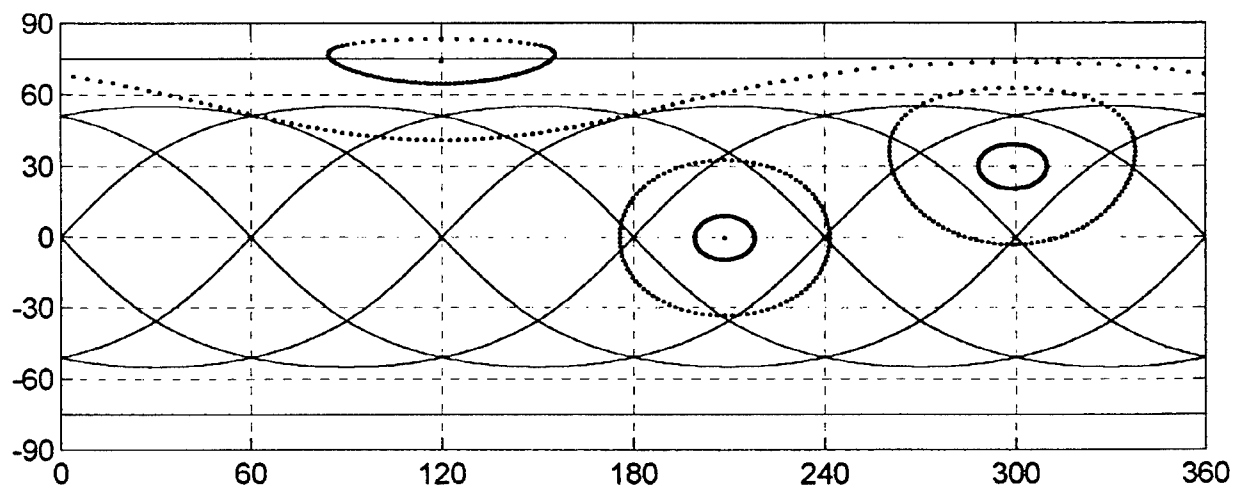*Figure 6.11: 1.75 hour zenith hole constellation ground track.*

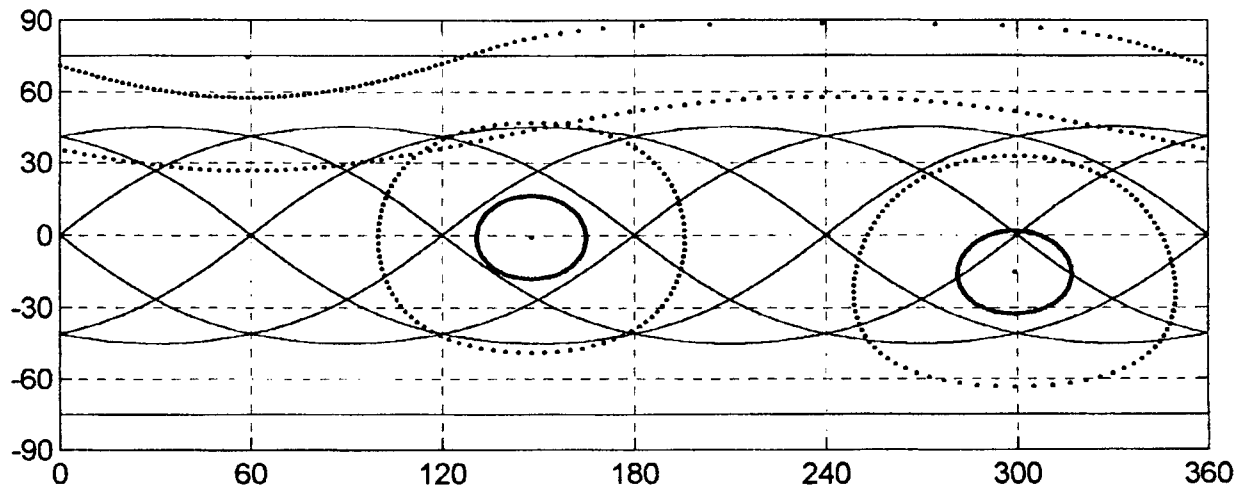*Figure 6.12: 2 hour zenith hole constellation ground track.*

Figure 6.13: *3 hour zenith hole constellation ground track.*
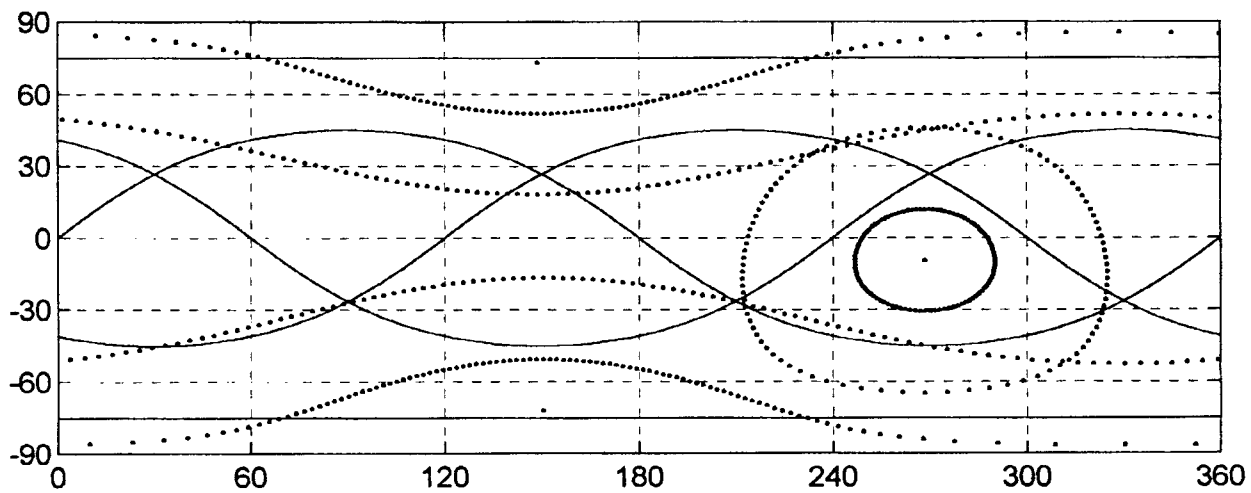
Figure 6.14: *4 hour zenith hole constellation ground track.*
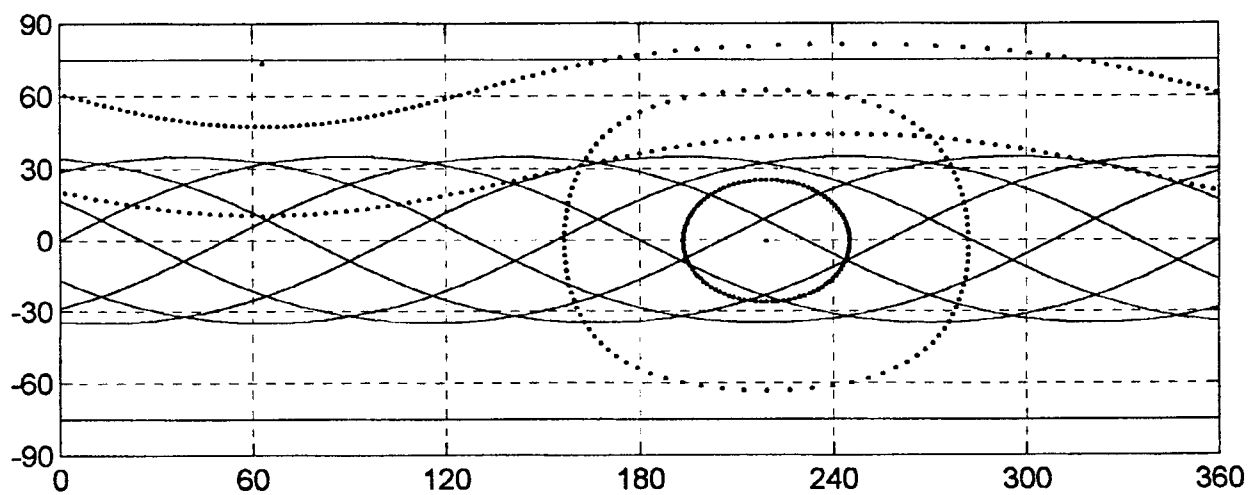
Figure 6.15: *6 hour zenith hole constellation ground track.*

6.4 Constellation Simulation

A constellation simulation process was the second SBR constellation design approach implemented. Although the initial goal was to design constellations from simulation results, the zenith hole analysis process later led to a more effective and flexible design methodology. Nevertheless, some of the empirical results from the simulation process are used to validate important relationships predicted by other methods. The simulation results are particularly important in deriving the altitude coverage function, Section 6.5.1.

The constellation simulation process uses the same orbit propagating code developed for the exhaustive search method. Constellations of arbitrary configuration size are simulated for a diurnal period of 24 hrs and the desired coverage statistics are recorded at every time step for every latitude and longitude on the simulation grid. Many arbitrary sets of orbital elements are implemented for a given configuration size and those with the best coverage statistics are kept. The coverage statistics for a configuration size are averaged and multivariable regressive techniques then characterize constellation size to give the *average* coverage as a function of constellation and configuration size. Regression is also used to characterize the minimum coverage relationships. The relationship between the model for average and minimum coverage is important since it validates a conjecture made in deriving the altitude coverage function below. The conjecture is that minimum coverage can be predicted by an application of the central limit theorem to the average coverage statistics.
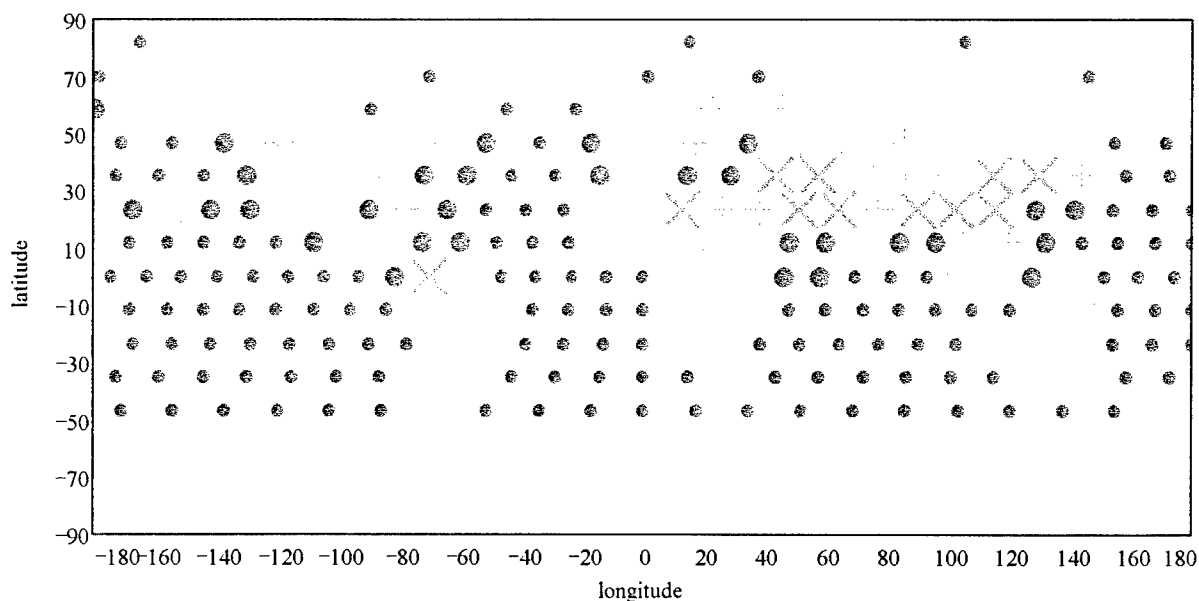
Figure 6.16: Discretized world. Cylindrical projection of grid consisting of 299 equally spaced points (11.7 deg grid separation). The grid points are weighted corresponding to their priority/threat for the radar mission. blank: no priority •: low priority ocean •: mod/high priority ocean (coast, shipping, oil routes)o: low priority land +: mod priority land (coastline defense) x: hi priority/threat land

Note that the multivariable regression relationships are based on averaged statistics. This was mostly done because of the enormous amount of data from a diurnal length simulation. A diurnal length simulation was initially deemed necessary because the constellation coverage of key targets changed during the course of a day. The zenith hole analysis seems to indicate that this is not a significant concern. Nevertheless, the averaging process can be justified based on the most important result of the constellation simulation: average coverage scales linearly with constellation size. This is a necessary relationship in the derivation of the altitude coverage function and is empirically proven by the simulation results. The relationship is discussed in more detail below.

In determining the best coverage set of orbital elements, a discrete world grid was used that was weighted according to threat. The goal of course is to make coverage commensurate with risk. The weighting on a discretized world is shown in *Figure 6.16*.

6.5 Results

The results given below are important parameters in the cost model developed in Chapter 3. The results include the altitude coverage function (section 6.5.1), the satellite coverage (section 0), the range to target function (section 6.5.4), and the duty cycle expectations (section 6.5.5).

6.5.1 Altitude Coverage Function

The altitude coverage function gives the number of satellites required for continuous coverage to ± 75 deg latitude as a function of the orbital radius. The altitude coverage function is empirically derived - power law regressed on number of satellites and radius - from the zenith hole constellations. The simulation process was used to confirm the results predicted by the altitude coverage function. The empirically derived altitude coverage function has a 98% coefficient of correlation as measured by the residues for the *eqn 6.12* power law. The parameters are given in *Table 6.4*.

$$acf(a) = \beta_1 a^{\alpha_{acf}} + \beta_0 \qquad\qquad eqn\ 6.12$$

when the orbital radius, a, must be in DUs. Recall that a DU is a distance unit equal to one Earth radii -- 6378 km.

Since the zenith hole constellations were designed to give minimum constellations for continuous coverage, *eqn 6.12* gives the number of satellites for a minimum coverage of one

Table 6.4: Altitude coverage coefficients for eqn 6.12. (Units in DUs)

α_{acf}	-5
β_0	7.1
β_1	54.8
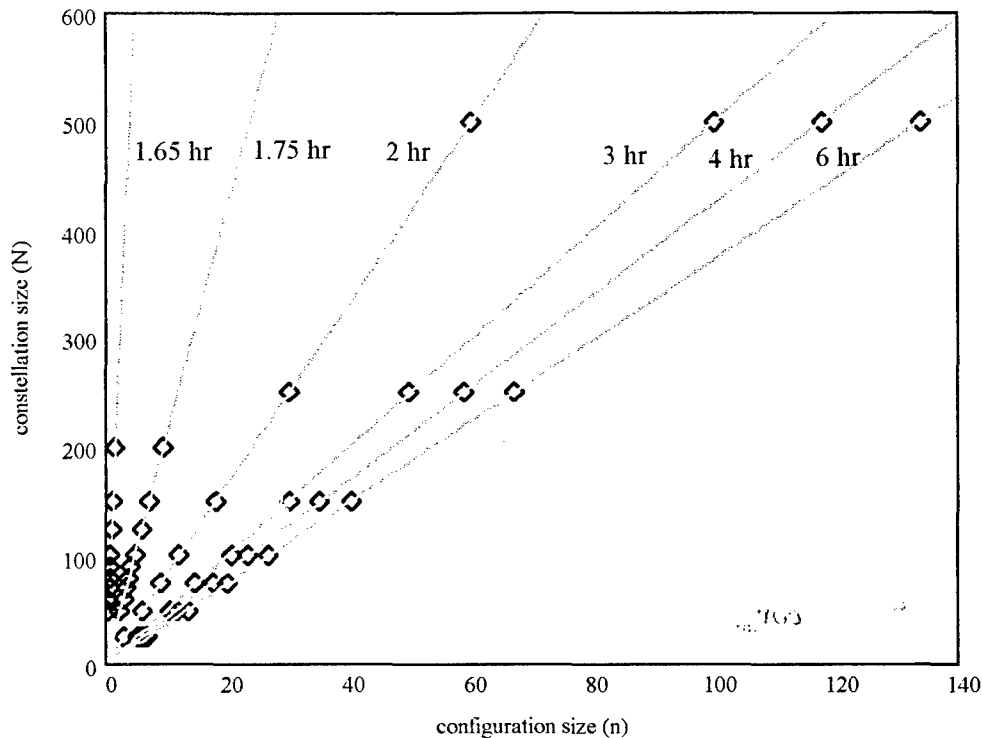
Figure 6.17: Constellation simulation results that demonstrate the linear relationship between constellation size and average configuration size. Constellations of arbitrary size were simulated and the average configuration size determined from the coverage results.

satellite at any time. Average coverage is often much better than the minimum coverage of one satellite as shown by the average coverage statistics in *Figure 6.20*. Average coverage also scales more clearly with constellation size than minimum coverage does. One empirical result of the constellation simulation is that average configuration size[1] is a linear function of overall constellation size (*Figure 6.17*). The converse is true as well: if an N size constellation gives an average of n coverage, $2n$ average coverage requires a $2N$ size constellation.

Although average coverage scales linearly with constellation size, minimum coverage does not for the following reason. For a configuration size of n_s, a minimum of n_s-fold coverage is required. As the constellation size increases, coverage is more uniform and the minimum coverage for a given latitude will approach the average latitude coverage in accordance with the central limit theorem. A reasonable conjecture is that the minimum coverage actually represents the average coverage less 2 or 3 standard deviations. Two standard deviations will explain over 95% of the observations if the underlying distribution is Gaussian. Three standard deviations will explain over 99% of the observations. Another aspect of the central limit theorem is that averaging processes approximate normal distributions for sufficiently large samples. Since the coverage statistics are averaged over a diurnal period of 24 hours, a Gaussian distribution should

---

[1] Recall that configuration size is the level of distribution.

110

approximate the coverage distribution as a function of latitude reasonably well. Thus, the conjectured relationship between average and minimum coverage as a function of distribution and latitude is

$$cov_{\min}(n_s, lat) = cov_{avg}(n_s, lat) - \frac{2\sigma_{cov}(lat)}{\sqrt{n_s}}$$

eqn 6.13

where cov_{min} gives the minimum coverage statistic as a function of latitude and cov_{avg} gives the average coverage statistic. This conjecture is actually verified by the constellation simulation results. This is done below.

Since constellation size scales linearly with average coverage (see above), the ratio between constellation size and average coverage should be a constant - the slope of the linear relationship - for a given orbital altitude. This coverage constant is

$$k_{cov}(a) = \frac{acf(a)}{cov_{avg}(a)}$$

eqn 6.14

The constellation size, N, now scales as a function of average coverage according to

$$N(a, cov_{avg}) = k_{cov}(a)cov_{avg}$$

eqn 6.15

where cov_{avg} is now the independent variable. Substituting *eqn 6.13*, constellation size is a function of minimum coverage.
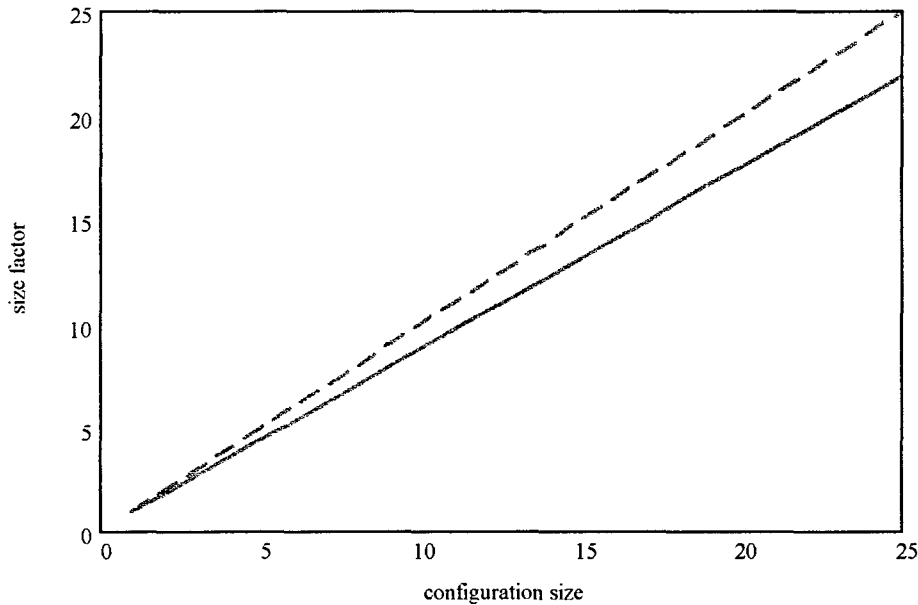
Figure 6.18: Configuration size factor (second set of parenthesis of eqn 6.17) as a function of configuration size. The dashed line is a strict linear increase for comparison.
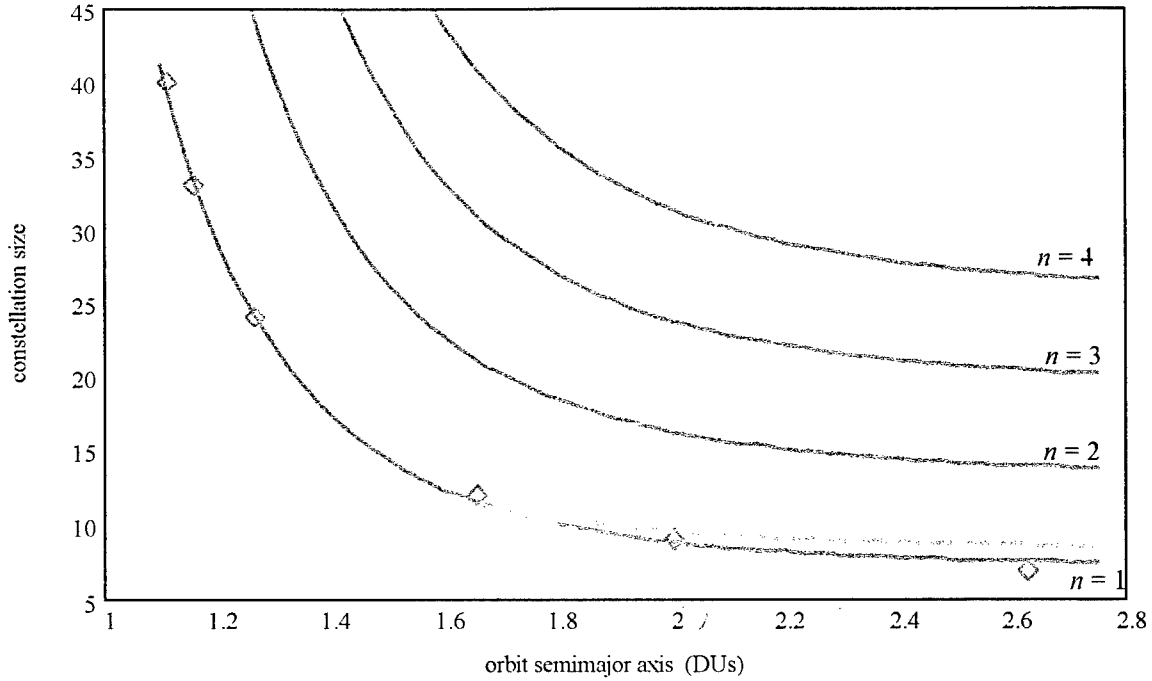
Figure 6.19: Altitude coverage function for several configuration sizes. The ◇ data points are for the zenith hole constellations. The dashed line is the simulation minimum coverage function (eqn 6.21) showing the agreement in constellation size.

$$N(a, cov_{min}) = k_{cov}(a)\left(cov_{min} + \frac{2\sigma_{cov}}{\sqrt{n}} \right)$$

eqn 6.16

When cov_{min} is taken as the desired configuration size, *eqn 6.16* is the altitude coverage function can be revised and expressed as a function of configuration size. That is

$$acf(a, n, lat) = \frac{acf(a,1)}{cov_{avg}(a, lat)}\left(cov_{min} + \frac{2\sigma_{cov}(lat)}{\sqrt{n}} \right)$$
$$= (\beta_1 a^{\alpha_{acf}} + \beta_0)\left(\frac{n}{cov_{avg}(a, lat)} + \frac{2\sigma_{cov}(lat)}{\sqrt{n}cov_{avg}(a, lat)} \right)$$

eqn 6.17

Note that the altitude coverage in this form is also a function of latitude since coverage is very latitude dependent (*Figure 6.29* through *Figure 6.34*). If continuous coverage is satisfied at the latitude of worst coverage, then it is also achieved at all other latitudes that have better average and minimum coverage. Thus, the cov_{avg} function in *eqn 6.17* should be evaluated at the latitude of minimum average coverage. For the zenith hole constellations, these latitudes of minimum mean coverage are listed in *Table 6.6*. The minimum mean coverage latitude should be used instead of the minimum coverage latitude since the altitude coverage function scales as a function of average coverage.

The term in the second set of parentheses of *eqn 6.17* should be unity for a configuration

112

size of 1 (see *eqn 6.13*). It is a factor that alters the altitude coverage function as a function of configuration size. In fact, the factor usually varies from unity by 5 to 10% -- note that the factor of 2 cannot always be exact as most times (< 95%) the random variable is not two standard deviations away from the mean. The altitude coverage function factor is only meant to capture the effects of distribution and the nature of the central limit theorem scaling is not changed by normalizing the altitude coverage factor by average coverage. The revised altitude coverage function is then

$$acf(a,n) = \left(\frac{\left(\beta_1 a^{\alpha_{acf}} + \beta_0\right)}{1 + 2\sigma_{cov}\left(lat_{min,mean}\right)} \right)\left(n + \frac{2\sigma_{cov}\left(lat_{min,mean}\right)}{\sqrt{n}} \right) \qquad eqn\ 6.18$$

For the zenith hole parameters given in *Table 6.6* and *Table 6.4*, the altitude coverage function as a function of orbit semimajor axis and configuration size is

$$acf(a,n) = \left(54.8a^{-5} + 7.1\right)\left(n + \frac{.142}{\sqrt{n}} \right).876 \qquad eqn\ 6.19$$

eqn 6.19 is plotted in *Figure 6.19*.

6.5.2 Altitude coverage function verification by simulation results

The form of the altitude coverage function given in *eqn 6.18* and *eqn 6.19* can be verified by the results of the constellation simulation process. Particularly important is the way that the coverage function scales with orbital radius - power law exponent of -5 - and the way the coverage function scales with configuration size - central limit theorem.

The altitude coverage function derived from the simulation data gives average coverage because of the regression approach used to process the simulation results. One of the important relationships already used above is readily apparent from *Figure 6.17*: average coverage (equivalent to average configuration size) scales linearly with constellation size. The altitude average coverage function for the simulation results is

$$acf_{avg}^{(sim)}\left(a,n_{avg}\right) = n_{avg}\left(\beta_{avg_1} a^{\alpha_{acf_{avg}}} + \beta_{avg_0} \right) \qquad eqn\ 6.20$$

The regressed power law exponent is -4.7 which closely corresponds to the -5 power law exponent regressed above. This is well within the accuracy of the simulation process.

To put *eqn 6.20* in the form of *eqn 6.12* which gives the number of satellites required for continuous coverage (not average coverage), eqn 6.20 must be multiplied by a factor that corresponds to infimum of the coverage standard deviation. The factor, which then gives N_{min}, was derived from the simulation data.

$$N_{min}^{(sim)}(a,n) = \left(\beta_{min2} n^{\alpha_n} + \beta_{min3}\right)acf_{avg}(a,n) \qquad eqn\ 6.21$$

The exponent on *n* was a free parameter in the regression: $n^{-1/2}$ corresponds exactly to the

113

Table 6.5: Simulation acf parameters. (Units are DUs)

simulation average acf	eqn 6.20
$\alpha_{acf\,avg}$	-4.7
$\beta_{avg\,0}$	3.57
$\beta_{avg\,1}$	14.2
simulation min coverage	eqn 6.21
α_n	-1/2
$\beta_{min\,2}$	1.337
$\beta_{min\,3}$	.92

conjecture above that the minimum coverage result˘ from application of the central limit theorem to the average coverage. *eqn 6.21* is plotted against the altitude coverage function in *Figure 6.19*. The simulation parameters are given in *Table 6.5*.

6.5.3 Satellite Coverage

Satellite coverage is an important metric and is easy to tally during the zenith hole constellation design process. *Table 6.6* summarizes the constellation coverage statistics. Average and minimum coverage is plotted in *Figure 6.29* through *Figure 6.34*. Coverage as a function of orbital radius is plotted in *Figure 6.20*.

Two important aspects of coverage were used in the derivation of the altitude coverage function. They are restated here:

1. Coverage averaged over latitude scales linearly with constellation size.

2. With increasing coverage - that is, increasing configuration size - the minimum coverage at any latitude approaches the average coverage at that latitude in accordance with the central limit theorem.

Table 6.6: Zenith hole constellations coverage statistics.

period (hr)	mean coverage over latitude	coverage standard dev	min mean coverage	min mean cov std dev	lat of min mean cov	minimum coverage	lat of min coverage
1.65	2.47	.8934	1.485	.0772	$\pm 75°$	.84	$\pm 40°$
1.75	2.45	.4827	1.741	.0481	$\pm 75°$	1.21	$\pm 5°$
2	1.77	.2255	1.393	.1109	$\pm 75°$	.95	$\pm 5°$
3	1.68	.2805	1.274	.0854	$\pm 75°$	1.15	$\pm 45°$
4	1.61	.2504	1.380	.0779	$\pm 50°$	1.06	$\pm 25°$
6	1.51	.2232	1.232	.0250	$\pm 75°$	1.15	$\pm 70°$
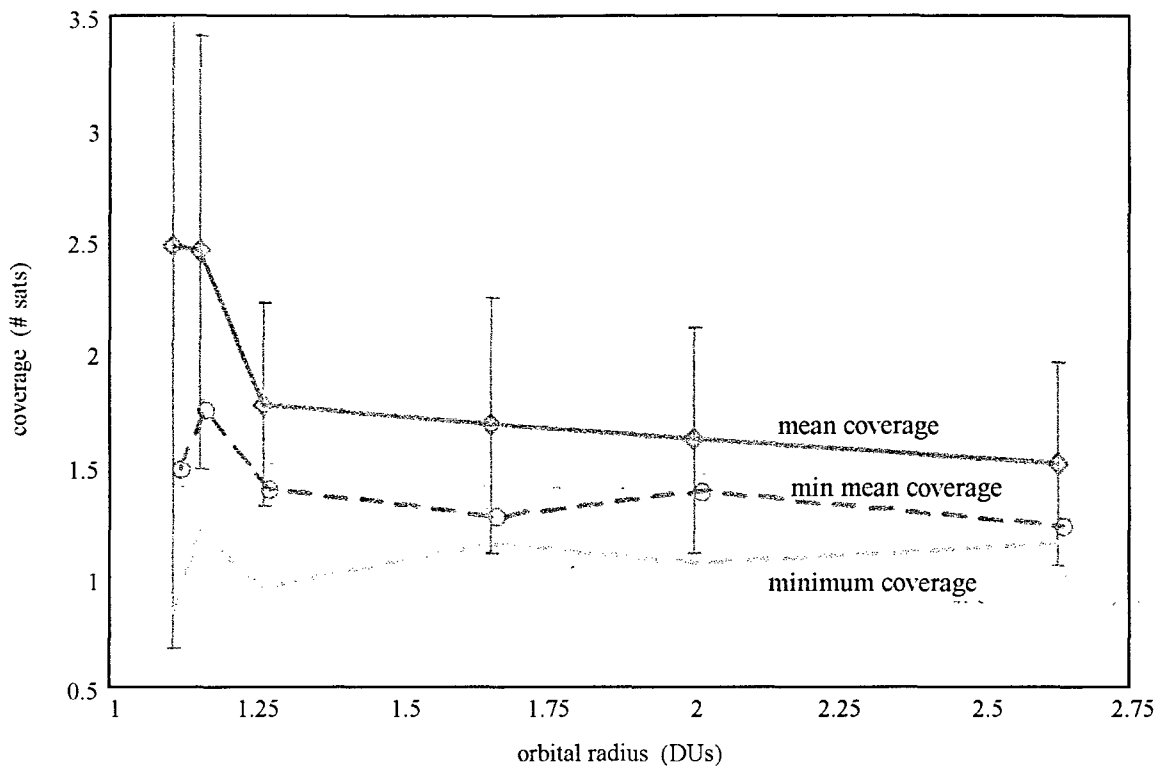
Figure 6.20: Zenith hole constellations coverage statistics as a function of orbital radius. Mean coverage is the diurnal coverage averaged over latitude. Minimum mean coverage is the diurnal coverage at the latitude with the minimum average coverage. Minimum coverage is the minimum coverage over the day and over all latitudes within ±75°. Both the mean coverage and minimum mean coverage are plotted with 1-sigma error bars.

6.5.4 Range to Target

We derive two expressions for radar slant range to target. The first is based on the zenith hole constellation range calculations. The second expression is based on an analytic derivation for range to target statistics and does not use the actual zenith hole constellation coverage patterns. Nevertheless, the second expression is probably more reasonable due to its general nature and analytic derivation.

Target slant range is a very significant parameter in the cost model developed in Chapter 3 because the search radar power aperture product scales with squared range. As with coverage, average range is easy to tally during the zenith hole constellation design process. Range statistics are also a function of constellation. Constellations with more orbital planes will have satellites closer to targets between orbital ground tracks. Average and maximum range for the zenith hole constellations is plotted in *Figure 6.35* through *Figure 6.40*. The latitude averaged maximum range to targets is plotted as a function of orbit semimajor axis in *Figure 6.21*.

115
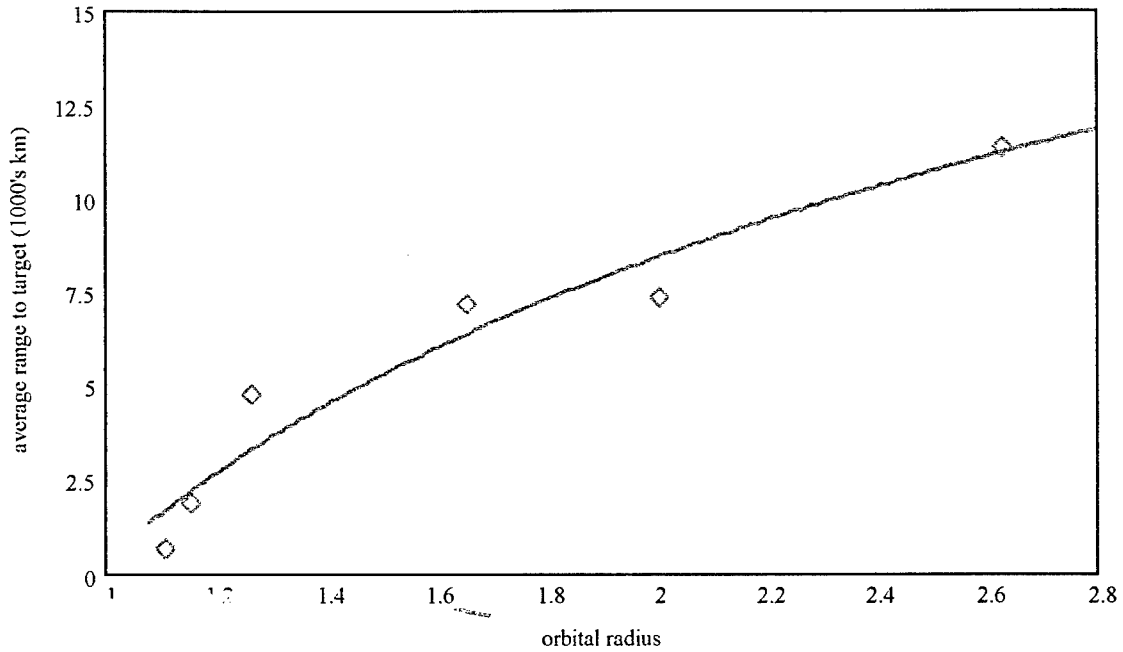
Figure 6.21: Average range to target as a function of orbital radius. The average is a diurnal average over longitudes and an average over latitudes between -75 deg and 75 deg.

The best fit average range to target curve plotted in *Figure 6.21* is the power law

$$rng_{zh\,data}(a) = \left(7.976 - 7.914a^{-1.4}\right) (DU) \qquad eqn\ 6.22$$

when the orbital radius, a, must be in DUs and the range is in units of DUs. Recall that a DU is a distance unit equal to one Earth radii -- 6378 km. This fit has a 66% correlation coefficient.

Instead of using the range statistics from the zenith hole constellations, we derive an analytic relationship for the maximum range to target. As with minimum coverage, maximum range as a function of distribution is also governed by the central limit theorem.

$$rng_{max} = rng_{avg} + \frac{2\sigma_{rng}}{\sqrt{n}} \qquad eqn\ 6.23$$

Recall that two standard deviations will explain 95% of the observations of a Gaussian random variable. Since targets can appear anywhere within the satellites nadir-hole coverage region, the range temporal average will approximate a Gaussian distribution in accordance with the second conclusion of the central limit theorem.

To calculate the average maximum range for use in the cost model, the average range and range variance must be characterized as a function of target distribution. Assuming that targets are uniformly distributed throughout the satellite coverage region (given by the elevation requirements - see Section 6.1), the average target elevation is given by

116
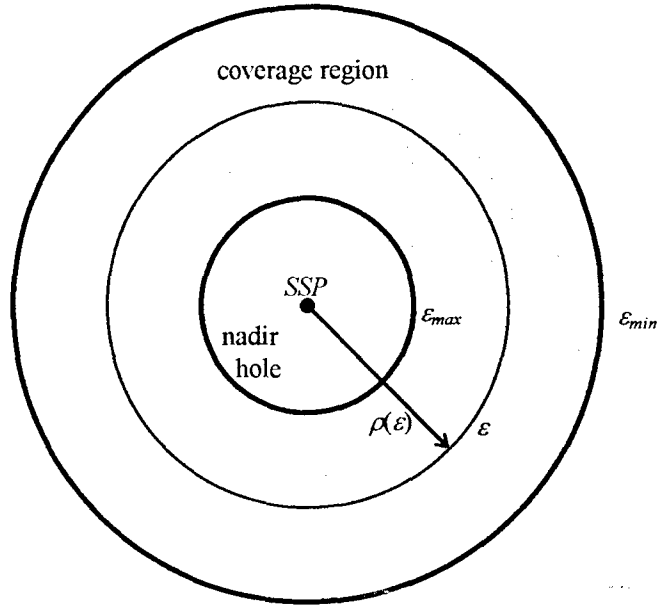
Figure 6.22: Average elevation derivation. Each elevation in the region of coverage is weighted by the circumference of the radius of elevation.

$$\left\langle \varepsilon_{tgt} \right\rangle = \frac{\displaystyle\int_{\varepsilon_{max}}^{\varepsilon_{min}} 2\pi\rho(\varepsilon)\varepsilon d\varepsilon}{\displaystyle\int_{\varepsilon_{max}}^{\varepsilon_{min}} 2\pi\rho(\varepsilon)d\varepsilon} \qquad\qquad eqn\ 6.24$$

where $\rho(\varepsilon)$ is the radius of elevation: on the Earth's surface, the arc-distance from the sub-satellite point to the target elevation. *eqn 6.24* is the elevation weighted average where the weight is the incremental amount of area for each elevation in the target area. Similarly, the elevation variance is

$$\sigma_\varepsilon^2 = \frac{\displaystyle\int_{\varepsilon_{max}}^{\varepsilon_{min}} 2\pi\rho(\varepsilon)\left(\varepsilon - \left\langle \varepsilon_{tgt} \right\rangle\right)d\varepsilon}{\displaystyle\int_{\varepsilon_{max}}^{\varepsilon_{min}} 2\pi\rho(\varepsilon)d\varepsilon} \qquad\qquad eqn\ 6.25$$

Average elevation and elevation standard deviation are plotted as a function of orbit semimajor axis in *Figure 6.23* and *Figure 6.24*. The standard deviation is not very dependant on orbital altitude and averages about 12.5 degrees.

Average range to target can be calculated in an analogous way to average elevation.
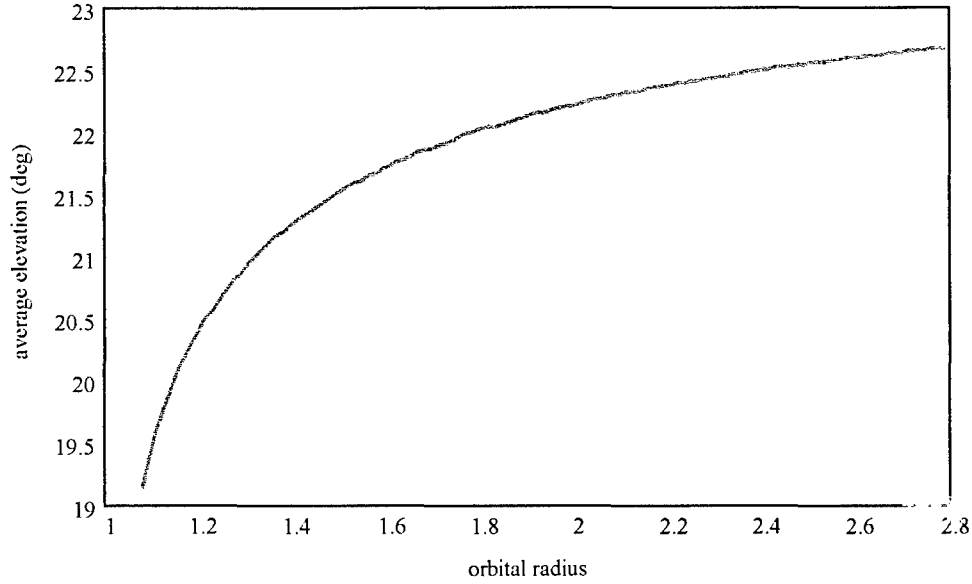
Figure 6.23: Average elevation as a function of orbital radius.

$$\left\langle rng_{tgt} \right\rangle = \frac{\int\limits_{\varepsilon_{max}}^{\varepsilon_{min}} 2\pi\rho(\varepsilon)R_{tgt}(\varepsilon)d\varepsilon}{\int\limits_{\varepsilon_{max}}^{\varepsilon_{min}} 2\pi\rho(\varepsilon)d\varepsilon}$$ *eqn 6.26*

$$\sigma_{rng}^2 = \frac{\int\limits_{\varepsilon_{max}}^{\varepsilon_{min}} 2\pi\rho(\varepsilon)\left(R_{tgt}(\varepsilon) - \left\langle R_{tgt}(\varepsilon) \right\rangle\right)d\varepsilon}{\int\limits_{\varepsilon_{max}}^{\varepsilon_{min}} 2\pi\rho(\varepsilon)d\varepsilon}$$ *eqn 6.27*
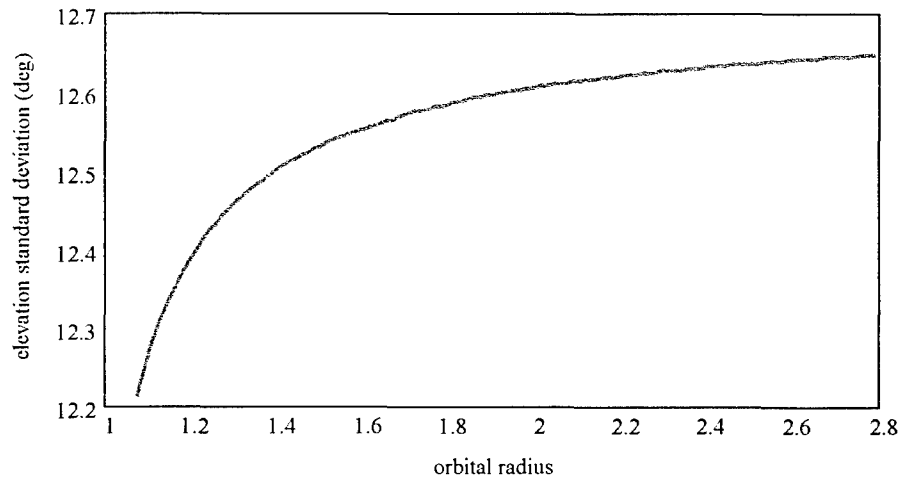
Figure 6.24: Elevation standard deviation as a function of orbital radius.

118
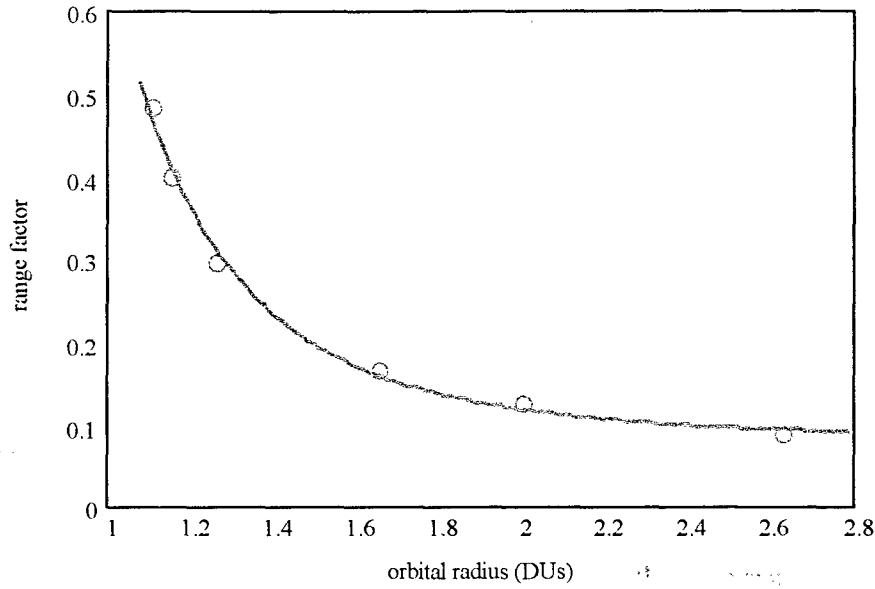
Figure 6.25: Range factor data and best fit power law.

where target slant range as a function of elevation is given by *eqn 6.3*. The difference between using *eqn 6.23* to calculate average range or using *eqn 6.3* with the average elevation (*eqn 6.24*) to compute average range is very slight (< 5%) and makes little difference in the cost model. Since it is easier to position targets by elevation rather than by radar slant range, we use the latter.

Define the range factor, k_{rng}, as

$$k_{rng}(a) = \frac{rng_{max}(a)}{rng_{avg}(a)} - 1 = \frac{R_{tgt}(\varepsilon_{min}, a)}{R_{tgt}(\varepsilon_{avg}, a)} - 1 \qquad\qquad eqn\ 6.28$$

so that

$$rng_{max}(a) = \left(1 + \frac{k_{rng}(a)}{\sqrt{n}}\right) rng_{avg}(a) \qquad\qquad eqn\ 6.29$$

The range factor is determined empirically by power law regression. The range factor data and best fit are plotted in *Figure 6.25*.

$$k_{rng}(a) = .082 + .587a^{-4} \qquad\qquad eqn\ 6.30$$

where a is expressed in DUs. The fit has a coefficient of correlation better than 98%.

Target average slant range is also given as an empirical relationship to avoid having to repeatedly solve *eqn 6.24* or *eqn 6.23* when implementing range in the cost model. The data and best fit power law are plotted in *Figure 6.26*. They are plotted against the zenith hole range data for comparison and are within one standard deviation of the data points. The empirical average

119
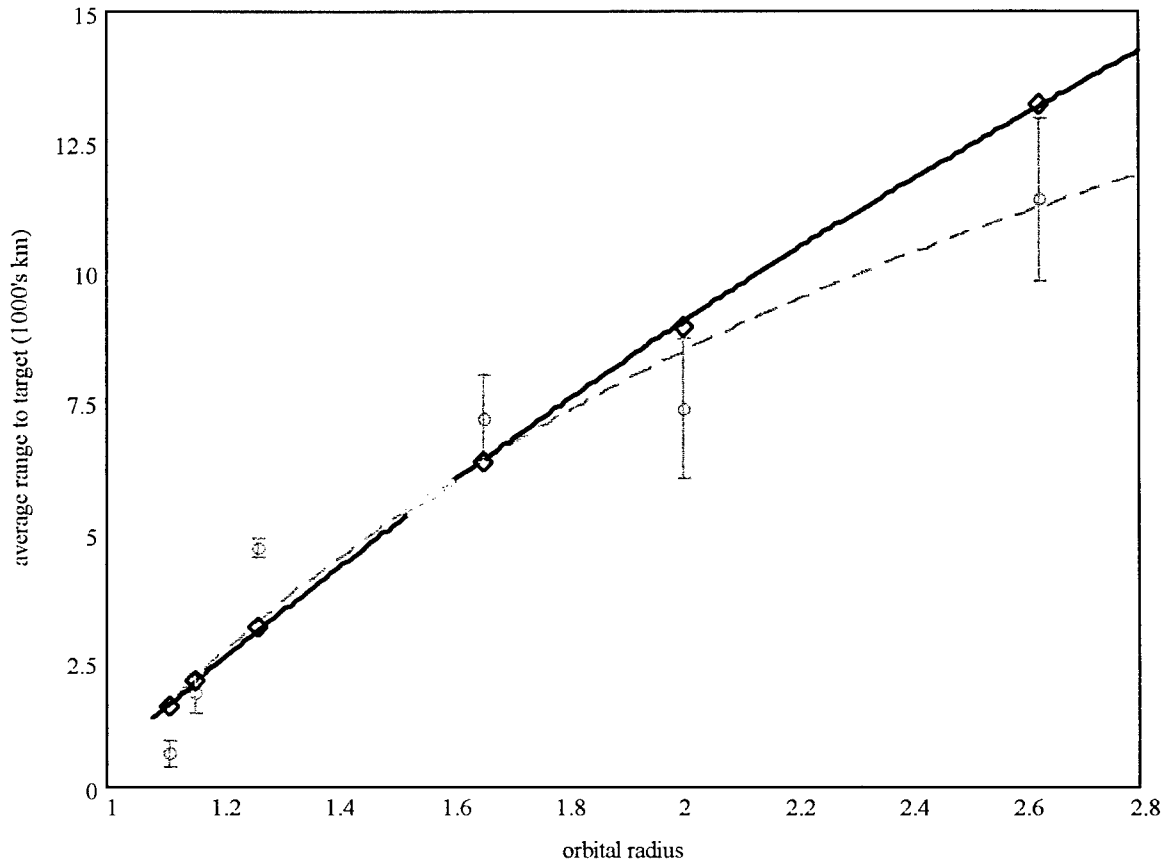
Figure 6.26: Average range data and best fit power law. The dashed curve and accompanying points and error bars are the zenith hole average range data. (See eqn 6.22 and Figure 6.21.)

range power law is

$$rng_{avg}(a) = \left(-3.1 + 3.188a^{1.2}\right) DU \qquad\qquad eqn\ 6.31$$

where a is again expressed in DUs. *6.26:* Average range data and best fit power law. The dashed curve and accompanying points and error bars are the zenith hole average range data. (See eqn 6.22 and Figure 6.21.)

$$rng_{avg}(a) = \left(-3.1 + 3.188a^{1.2}\right) DU$$ eqn 6.31 has a 90% correlation coefficient for describing average range as given by *eqn 6.24* or *eqn 6.26*.

The two sigma maximum slant range to target used in the cost model is then

$$R_t(a,n) = \left(1 + \frac{.082 + .587a^{-4}}{\sqrt{n}}\right)\left(-3.1 + 3.188a^{1.2}\right)\ DU \qquad eqn\ 6.32$$

120

where both slant range and orbital radius are expressed in DUs. *eqn 6.32* is plotted in *Figure 6.27* for a range of distributions.
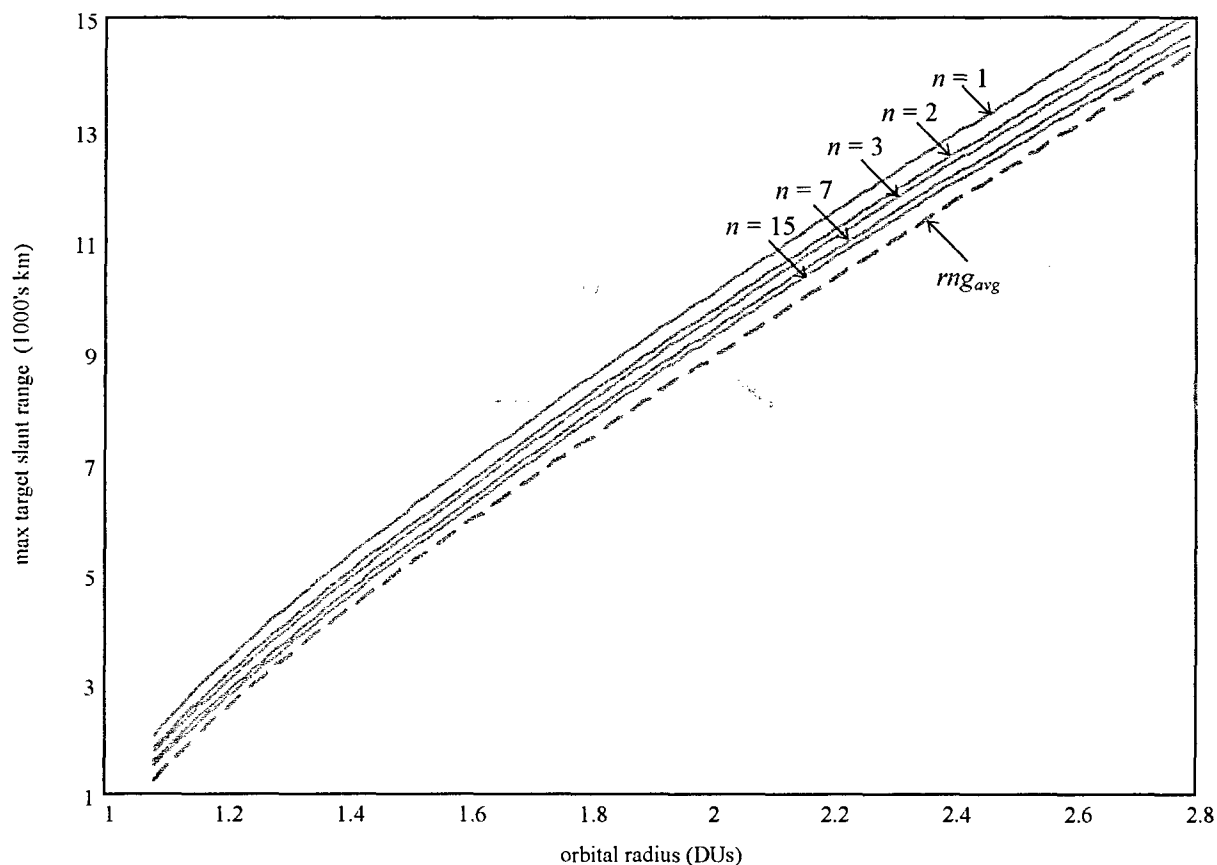
Figure 6.27: Two sigma maximum slant range to target as a function of orbital radius.

6.5.5 Duty Cycle

Satellite duty cycle is the percent time that the satellite must be performing its primary mission. (Radar duty cycle is the percent time that the transmitter is on while the radar is operating.) We define mission cycle as the duty cycle expressed as fraction of orbital period. Mission cycle is a parameter that is important for determining the power mass density since it governs the necessary secondary power storage capacity (see Chapter 3.2). Decreased duty cycle implies less required storage capacity but it also implies less efficient use of assets on orbit.

The duty cycle results are from the constellation simulation. There are three main duty cycle results:
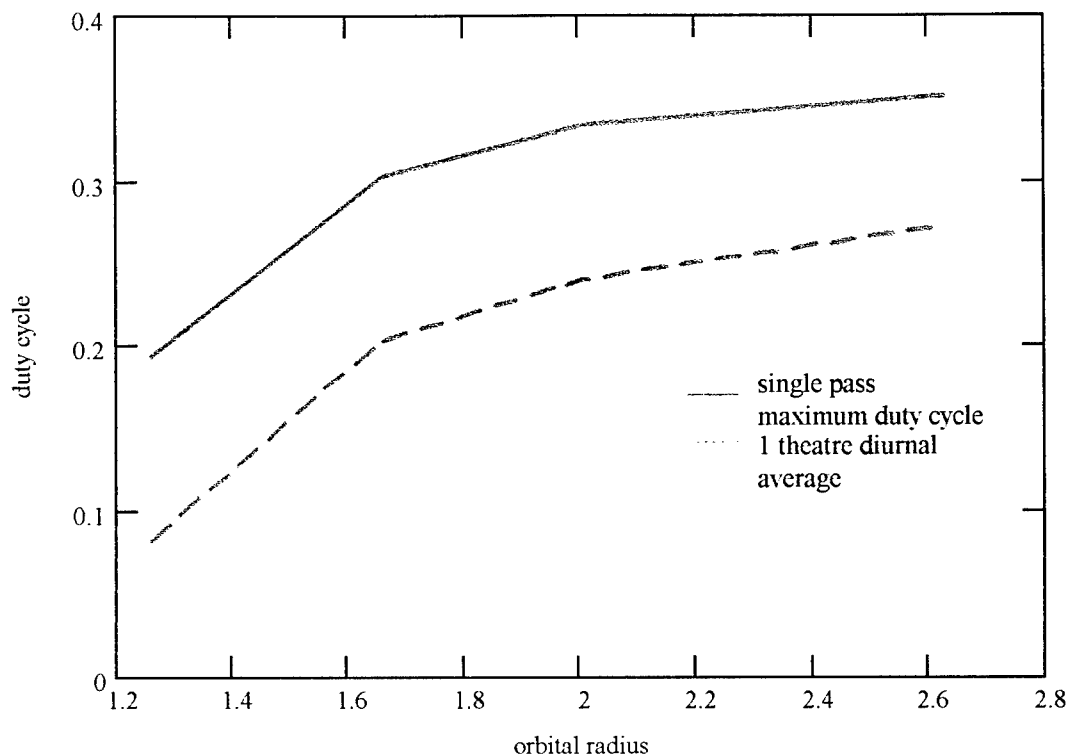
Figure 6.28: Duty cycle for a single theatre as a function of orbital radius.

1. Distribution has no effect on duty cycle. Duty cycle is independent of configuration size.

2. Duty cycle increases with orbital altitude. Mission cycle variation between low earth orbit and higher orbits is about 75%. Duty cycle averaged over a day (diurnal average) can vary by 150% between leo and higher orbits. (*Figure 6.28*).

3. Maximum mission cycle for a two theatre scenario is two times the single theatre mission cycle.

6.5.6 Effect of Eccentricity

The majority of land mass is situated in the Northern Hemisphere. When a differentiation is made for where likely future conflicts are to occur, the distinction is even stronger -- the Northern Hemisphere includes all regions that have been historically strife (Middle East, India/Pakistan, China, Korea) and those that are likely to become so (based on the fastest developing countries and those with emerging nuclear technologies). Constellations that concentrate coverage over the Northern Hemisphere are then desirable. Satellites in eccentric orbits with perigee located in the Southern Hemisphere spend greater fractions of their orbit over

northern latitudes. Simulation results for eccentricity values of 0, .25, .5 did result in a slight decrease in the number of satellites required for coverage. However, this was offset by an increase in range and by poorer coverage for low latitude, Northern Hemisphere targets. Thus, there does not seem to be a significant advantage for going to eccentric orbits, particularly if the mission specifications require global coverage. For these reasons and because is greatly simplifies the analysis, circular orbits are assumed for the system architecture cost model.
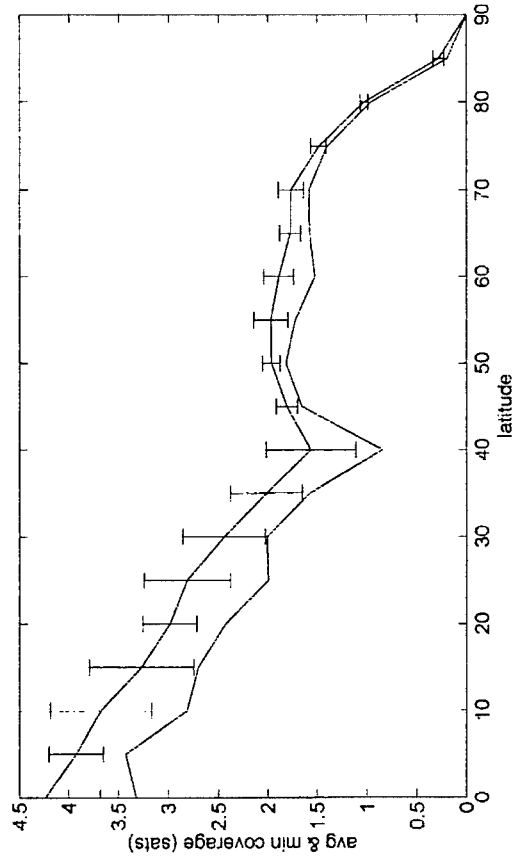
Figure 6.29: Coverage for 1.65 hour zenith hole constellation.
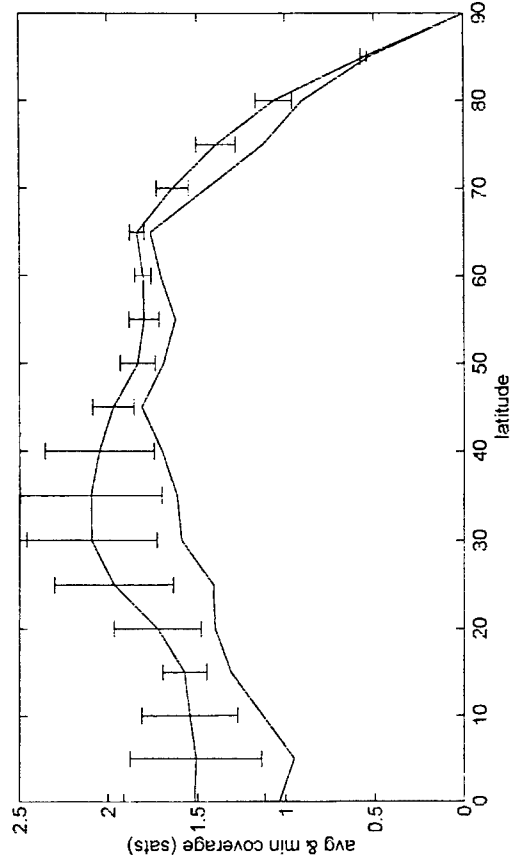
Figure 6.31: Coverage for 2 hour zenith hole constellation.
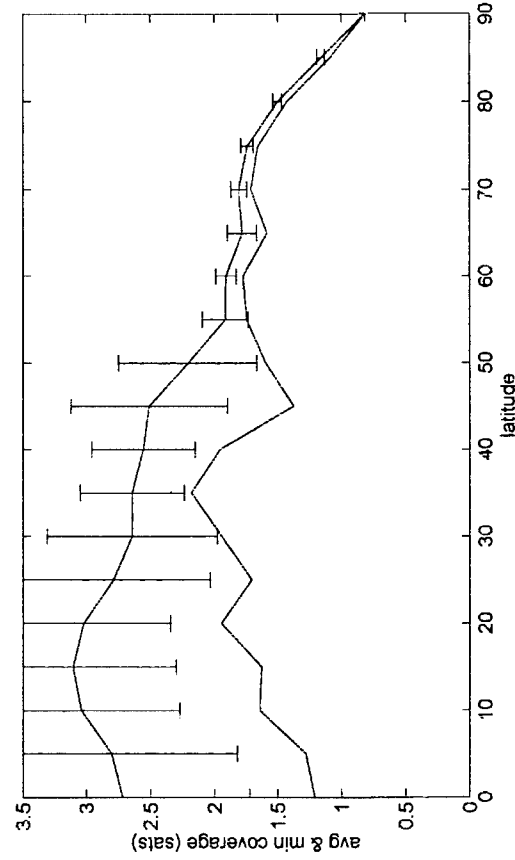
Figure 6.30: Coverage for 1.75 hour zenith hole constellation.

Figure 6.32: Coverage for 3 hour zenith hole constellation. Average coverage is given with 1-sigma error bars.

Minimum and average coverage are plotted as a function of latitude. Average coverage and average coverage are plotted as a function of latitude. Average coverage is given with 1-sigma error bars.
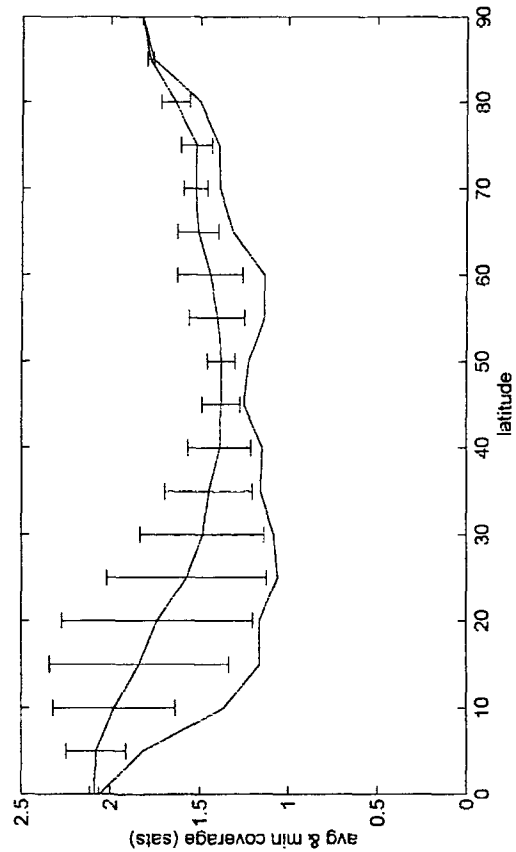
Figure 6.33: Coverage for 4 hour zenith hole constellation.
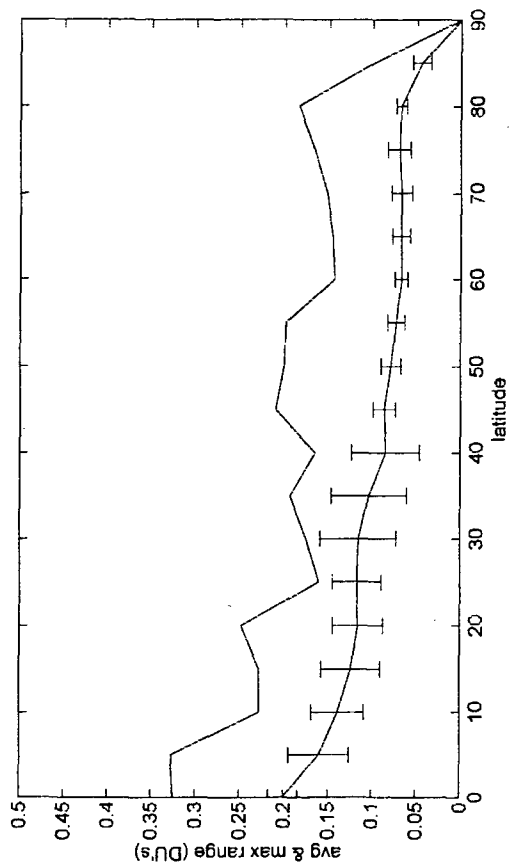
Figure 6.35: Range profile for 1.65 hour zenith hole constellation.

Figure 6.34: Coverage for 6 hour zenith hole constellation.

Figure 6.36: Range profile for 1.75 hour zenith hole constellation.

125

Figure 6.37: Range profile for 2 hour zenith hole constellation.

Figure 6.38: Range profile for 3 hour zenith hole constellation.

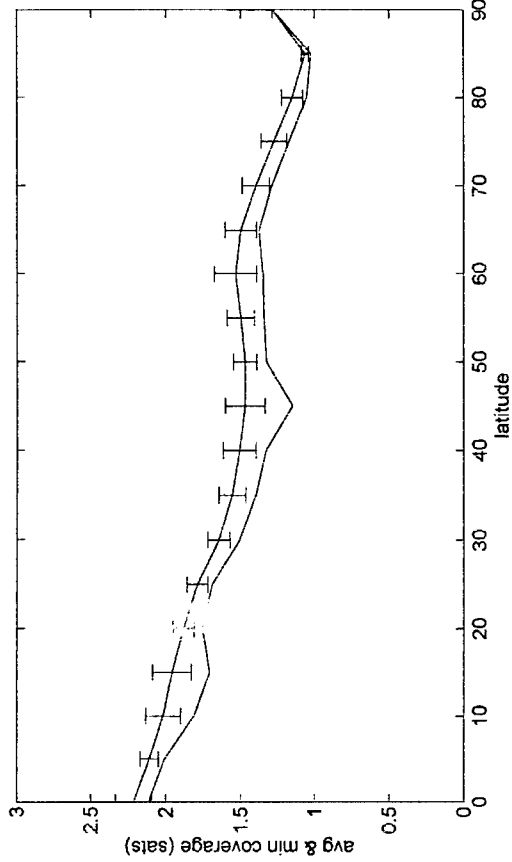*Figure 6.39: Range profile for 4 hour zenith hole constellation.*
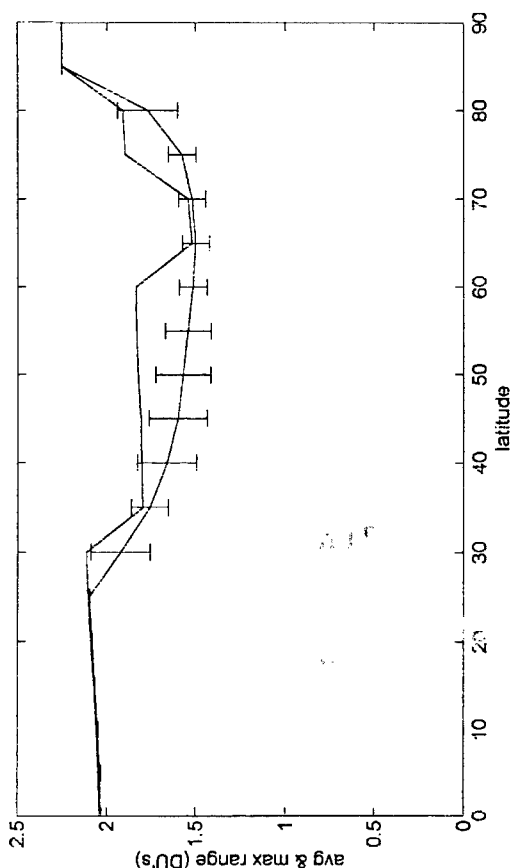
Figure 6.40: Range profile for 6 hour zenith hole constellation.

Maximum and average slant range are plotted as a function of latitude. Average range is given with 1-sigma error bars.
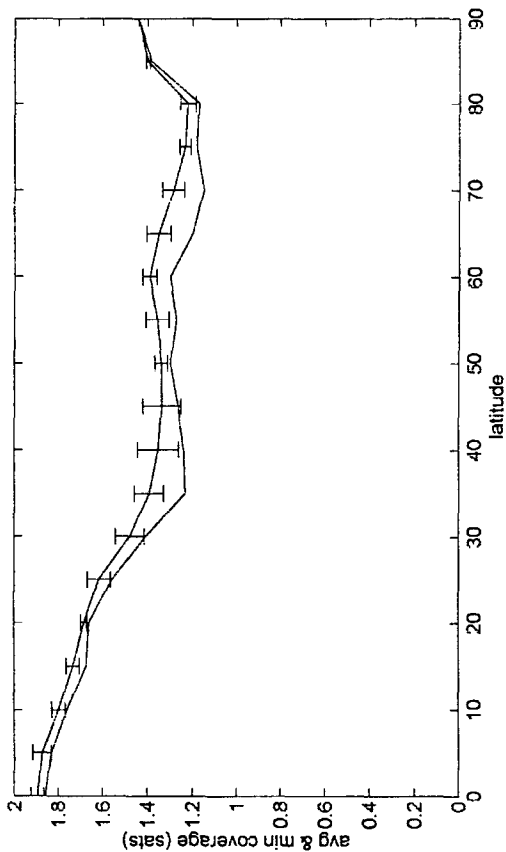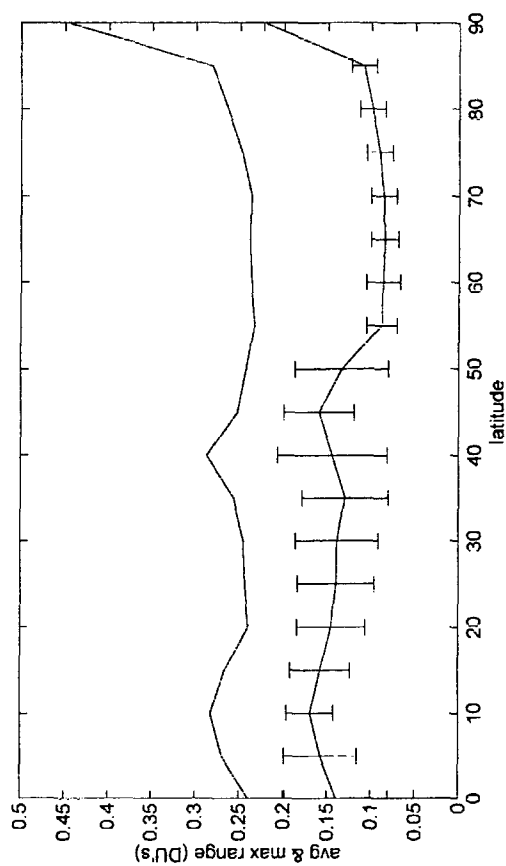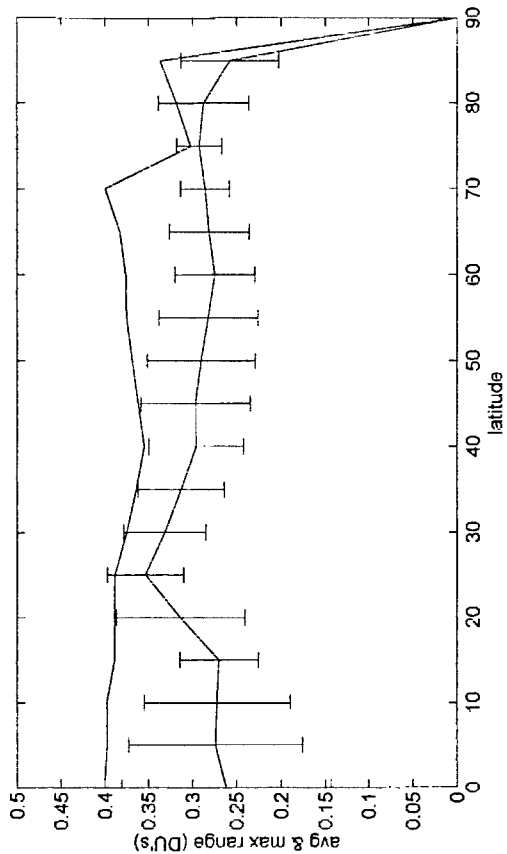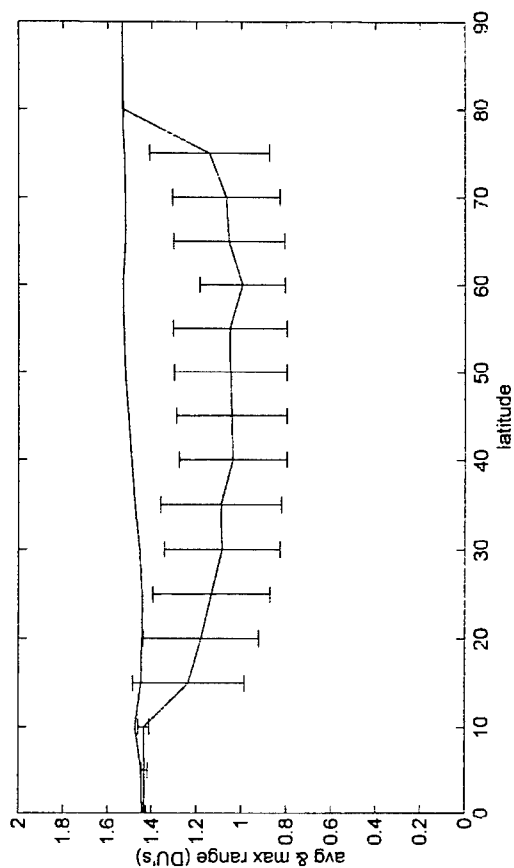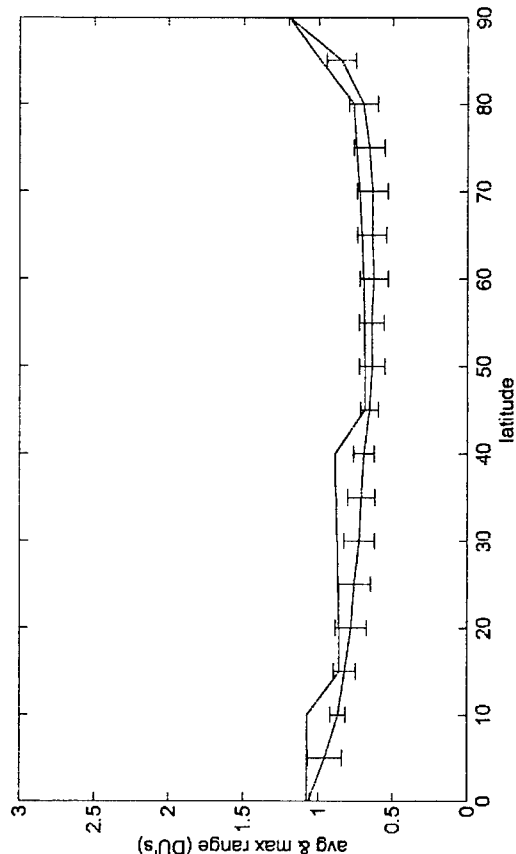
7. SBR Performance in Clutter

Clutter is the unwanted signal return from the ground, ocean, airborne particles, and weather. Although the individual clutter elements that reflect the transmitted signal typically have a smaller RCS than the desired targets, the large number or area of these clutter elements represents a combined received energy that exceeds the energy from legitimate targets. Thus, the clutter signal typically buries the target signal and returned signal strength is insufficient to separate the desired target from clutter.

Targets, though, often move with respect to the clutter background. The relative velocity between targets and clutter results in different frequency spectra between the targets and clutter. This is the most widely used form of moving target indicator (MTI) radar. It is sometimes known as pulse-doppler MTI radar. The achievable resolution in the radar receiver for separating targets and clutter by spectrum depends on the transmitted waveform and the radar-target-clutter geometry and relative motion. For a given frequency resolution, a moving target is detectable if its velocity with respect to the clutter background results in a Doppler shift that moves the target spectrum outside of the clutter spectrum. This implies that there exists a minimum detectable velocity (MDV) for detecting moving targets with MTI. MDV is a useful metric for quantifying the performance of a radar in clutter. This is *clutter performance*.

In addition to pulse-doppler MTI radar, there are several signal processing techniques for canceling clutter return. These include displaced phase center array (DPCA) and adaptive array processing. DPCA processing compensates for the clutter spread that is due to the radar platform's velocity by making the antenna appear stationary between pulses. Because of this, it can only be used when the beam is directed to a region around broadside and not along-track. This represents a serious limitation in terms of coverage of a space based platform. It is conceivable to use DPCA for broadside coverage and pulse-doppler for along-track coverage. Clutter cancellation also really only becomes necessary for GMTI (ground MTI). For these reasons, and because we are primarily interested in SBR performance of the AWACS mission, we only model pulse-doppler MTI.

In the analysis of the previous sections, we have designed the system architecture with respect to the time to detection performance metric. Mean time to detection is the performance requirement input and system IOC cost is the cost function for the system architecture optimization accomplished in Section 8. The system clutter performance is also an important performance metric and requirement. However, setting both detection and clutter performance requirements prior to system architecture optimization would create an overconstrained problem. We have chosen detection performance as the primary performance metric since it ultimately drives the system response time in battlefield situations (Section 5.3.2).

Although clutter performance is important, it is not as critical for several reasons. First, clutter performance varies widely depending upon the radar-target geometry. For example, Doppler shift depends upon the radial velocity between the radar and target. When the target

velocity vector approaches right angles with the radar beam direction, the radial velocity approaches zero and the target is undetectable regardless of the capability of the radar system. Thus specifying a MDV system requirement becomes important only for a small subset of target geometry. Additionally, for a distributed architecture, some satellites tend to have very favorable coverage geometry's and a decreased MDV requirement would give the same configuration performance. However, this cannot be captured in the requirements definition.

Clutter performance is not critical, but it is still an important performance consideration. Thus, we design the system to meet detection performance requirements, and then verify that the clutter performance is acceptable. In this chapter we develop the nature of clutter return (Section 7.1) and then develop a radar waveform well suited to clutter processing from a space based radar platform (Section 7.2). The clutter processing requirements are developed in Section 7.3 and the clutter performance for a distributed operations concept is developed in Section 7.3.1. Finally, the clutter performance for the optimum SBR systems derived in Chapter 8 is analyzed as a function of frequency, altitude, and distribution in Section 8. 6.

7.1 Clutter Return

Clutter is the signal return of reflections from anything other than the desired target. This includes reflected energy from the ground, ocean, clouds and precipitation, airborne particles, and even birds and insects. The most significant return in terms of total energy is from the ocean and ground. Although the typical RCS for both the sea and nonmountainous, unpopulated land is -20 to -40 dB per square meter (10^{-1} to 10^{-2} m^2 per m^2), a typical SBR footprint area is on the order of 10^2 to 10^4 km^2 depending on frequency. Thus, the total energy reflected from the clutter surface is significant – on the order of 120 to 180 dB. For comparison, a 100 m^2 RCS target will only reflect 40 dB.

The reflected signal has a frequency spectrum that depends on the spectrum of the transmitted waveform, the radar platform velocity, and the radar-clutter geometry. The waveform spectrum is discussed in Section 7.2. For a circular orbit, the orbital velocity or radar platform velocity with respect to inertial space is

$$v_{orb}(a) = \sqrt{\frac{\mu}{a}}$$

eqn 7.1

where a is the orbit semi-major axis and μ is the Earth's gravitational constant, 3.98 10^5 km^3/sec^2.

The radar-clutter geometry influences clutter spectrum in several ways. First, the clutter Doppler shift depends on the relative velocity between the radar platform and the clutter which in turn depends on the orbital velocity, the radar-clutter elevation, and the beam azimuth to the clutter measured from the radar velocity vector. The radar-clutter geometry also affects the clutter spread due to the range of elevation through footprint length, and the range of azimuth through footprint width. We start by characterizing the footprint geometry.
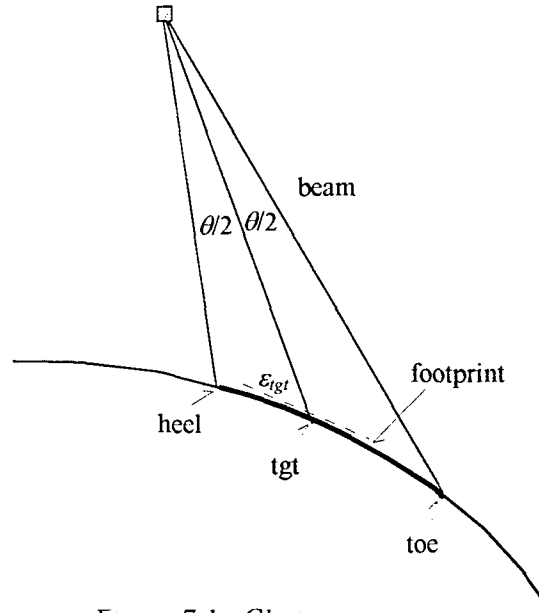
Figure 7.1: Clutter geometry.

If the center of the beam is directed at a point on the surface with elevation ε_{tgt}, then the beam nadir offset is

$$n_{tgt} = \sin^{-1}\left(\frac{R_E}{h+R_E}\sin\left(\varepsilon_{tgt}+\frac{\pi}{2}\right)\right)$$

eqn 7.2

where R_E is the radius of the Earth, and h is the orbital altitude.

The half-power beamwidth in the range dimension for a diffraction limited aperture with range dimension D_r is λ/D_r where λ is the wavelength of the transmitted signal. The beam nadir offset for the footprint toe is then

$$\eta_{toe} = \eta_{tgt} + \frac{c}{2fD_r}$$

eqn 7.3

and the beam nadir offset for the footprint heel is

$$\eta_{heel} = \eta_{tgt} - \frac{c}{2fD_r}$$

eqn 7.4

where c is the speed of light, f is the signal frequency and η_{tgt} is given by *eqn 7.2*. Often, nadir offset is expressed as the grazing angle, γ, where

$$\gamma = \frac{\pi}{2} - \eta$$

eqn 7.5

Given a nadir offset angle η, the corresponding Earth central angle from altitude h is

$$\lambda(\eta,h) = \frac{\pi}{2} - \eta - \cos^{-1}\left(\frac{h+R_E}{R_E}\sin(\eta)\right)$$ *eqn 7.6*

The slant range to point with nadir offset η and central angle λ is

$$rng = R_E \frac{\sin(\lambda)}{\sin(\eta)}$$ *eqn 7.7*

The clutter source is the region of the Earth laying within the half-power beam footprint. The length of the footprint is

$$l_{foot} = R_E\left(\lambda_{toe} - \lambda_{heel}\right)$$ *eqn 7.8*

and the width of the footprint at nadir offset η is

$$w_{foot} = rng\left(\eta,\lambda(\eta)\right)\frac{c}{fD_{cr}}$$ *eqn 7.9*

where D_{cr} is the array dimension that gives cross-range resolution, c is the speed of light, and f is the transmitter frequency. If γ_{toe} is greater than zero, the entire footprint lies on the surface of the Earth and forms an ellipse with major and minor axes given by the footprint length and width at midpoint nadir offset. The footprint area is then

$$A_{foot} = \pi l_{foot} w_{foot}$$ *eqn 7.10*

Range is also important in radar since it represents the time delay between the transmitted and received signals. Clutter returns from the heel of the footprint will arrive before return from the toe. The heel and toe also have different Doppler frequency shifts due to the difference in grazing angle. Doppler shift also depends on azimuth – the leading edge of the footprint has an increased shift compared to the footprint trailing edge. Thus, we can characterize the clutter ambiguity plane by four points: the toe, heel, leading edge, and trailing edge. For a beam center point with elevation ε_{tgt} and azimuth β_{tgt}, the ambiguity plane ordered pairs are

$$\left(t_d,f_d\right)_{toe} = \left(\frac{2rng\left(\lambda_{toe},\eta_{toe}\right)}{c}, 2fv_{orb}\frac{\sin\left(\eta_{toe}\right)\cos\left(\beta_{tgt}\right)}{c}\right)$$

$$\left(t_d,f_d\right)_{heel} = \left(\frac{2rng\left(\lambda_{heel},\eta_{heel}\right)}{c}, 2fv_{orb}\frac{\sin\left(\eta_{heel}\right)\cos\left(\beta_{tgt}\right)}{c}\right)$$

eqn 7.11

$$\left(t_d,f_d\right)_{le} = \left(\frac{2rng\left(\lambda_{tgt},\eta_{tgt}\right)}{c}, \frac{2fv_{orb}}{c}\sin\left(\eta_{tgt}\right)\cos\left(\beta_{tgt} + \frac{c}{2fD_{cr}}\right)\right)$$

$$\left(t_d,f_d\right)_{te} = \left(\frac{2rng\left(\lambda_{tgt},\eta_{tgt}\right)}{c}, \frac{2fv_{orb}}{c}\sin\left(\eta_{tgt}\right)\cos\left(\beta_{tgt} - \frac{c}{2fD_{cr}}\right)\right)$$
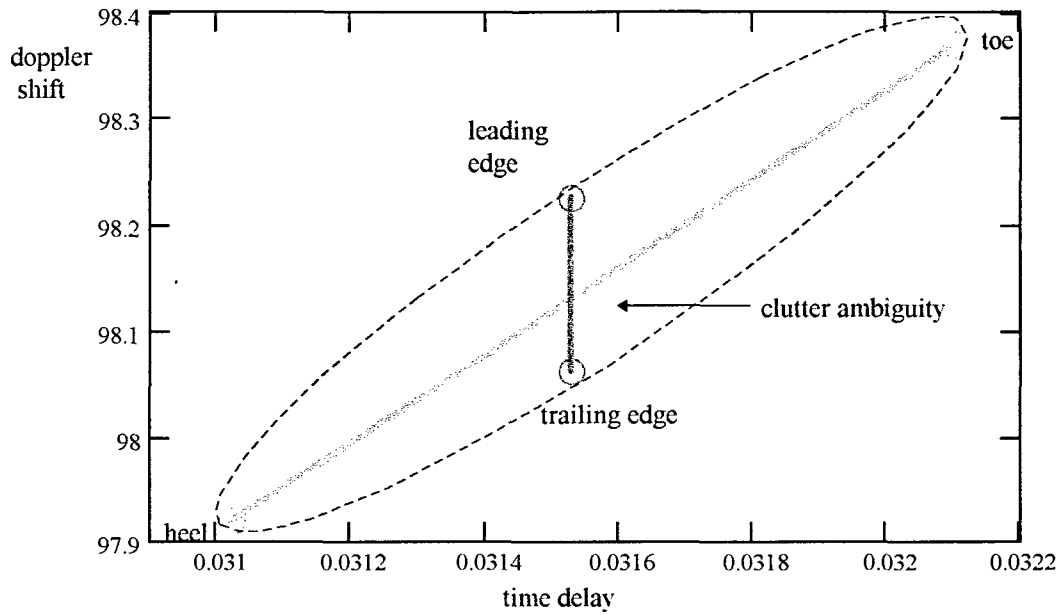
Figure 7.2: Clutter ambiguity plane characterization. The endpoints are given by eqn 7.11.

where t_d is the time delay and f_d is the Doppler shift. D_{cr} is the cross-range dimension aperture. These four points characterize the clutter ambiguity plane in the sense that they mark the approximate major and minor axes of the clutter ambiguity ellipse. Due to the trigonometric relationships in *eqn 7.11* for range, the ambiguity is not a perfect ellipse. Nevertheless, it is convenient to approximate it as an ellipse defined by the points in *eqn 7.11*.

Range resolution can eliminate the clutter spread between the footprint toe and heel. Cross-range or azimuth clutter spread however cannot be altered except by changing the footprint width. Since cross-range clutter spread is significant limitation to MTI for SBR, the footprint width should be minimized. This requires increasing the aperture length parallel to the cross-range dimension. For a fixed array area, this will result in a longer footprint in range which will then require increased unambiguous range.

Range resolution partitions the range axis into finite range cells or range bins. Thus, range bins have a dimension in range given by the waveform range resolution and a dimension in cross-range given by the half-power beamwidth in azimuth. The clutter spectrum of the bin is spread from the change in elevation due to range dimension and the change in azimuth due to cross-range dimension. The clutter spectrum of each bin convoluted with the waveform spectrum governs the minimum detectable velocity and clutter performance.

7.2 Pulse Train Waveform

The pulse train is a commonly used waveform in radars. It is motivated chiefly by the range-doppler uncertainty principle. Because range (measured by time delay) is the Fourier

131
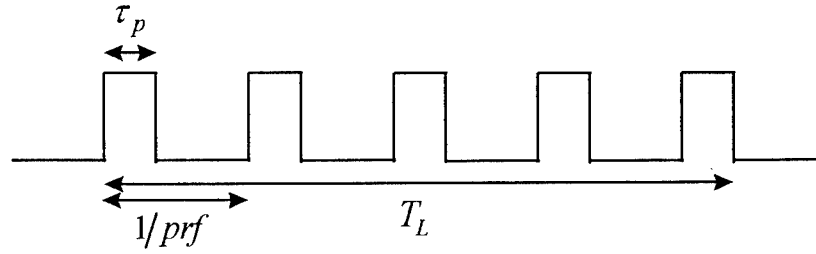
Figure 7.3: Pulse train definition

conjugate of frequency, it is impossible to increase resolution in range without sacrificing resolution in frequency. Short pulses give narrow autocorrelations and hence good range resolution. Long pulses have narrow Fourier transforms which gives good frequency resolution.

The pulse train is a hybrid between short and narrow pulses (*Figure 7.3*). A long series of short pulses can simultaneously improve both time and frequency resolution. Ambiguity is always conserved. Therefore, the cost of improving both time and frequency resolution is that the ambiguity is pushed from the mainlobe into other regions of the ambiguity plane. That is, the pulse train results in both ambiguous ranges and blind velocities.

Frequency resolution is given by the width of the main lobe of the waveform's Fourier transform (*Figure 7.6*). If the total pulse train length is T_L, the frequency resolution is

$$\Delta f_d = \frac{1}{T_L} \qquad\qquad eqn\ 7.12$$

Frequency ambiguity is the aliasing due to pulse repetition frequency. In the frequency domain, it appears as the convolution of a sinc function with a *prf* periodic function (*Figure 7.6*). The frequency ambiguity is

$$\Delta f_{am} = prf \qquad\qquad eqn\ 7.13$$
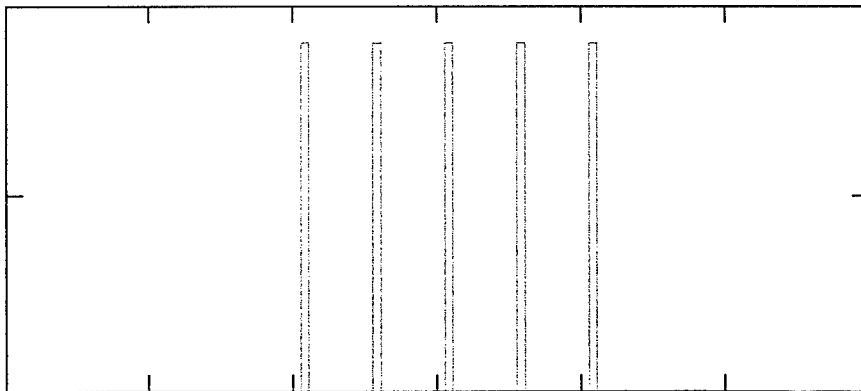
Figure 7.4: Pulse train example.

132

-3/prf -2/prf -1/prf 0 1/prf 2/prf 3/prf   time delay
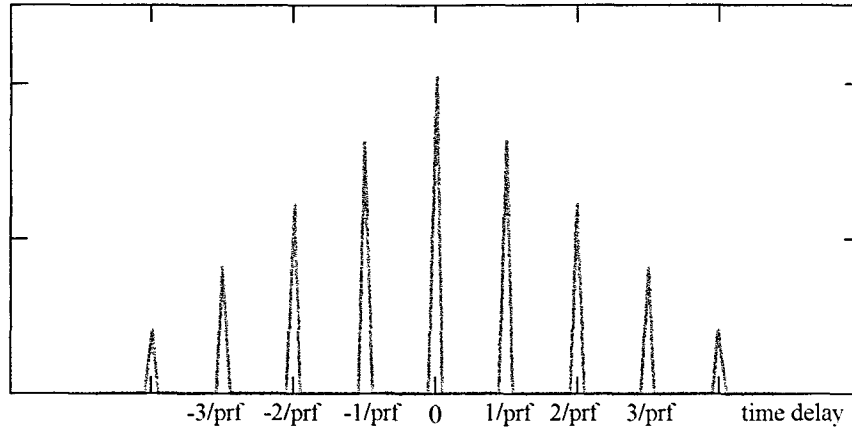
Figure 7.5: Autocorrelation of pulse train in Figure 7.4.

where *prf* is the pulse repetition frequency. Ambiguous frequencies result in blind velocities in an MTI radar. This is because the clutter frequency spectrum is aliased at multiples of the pulse repetition frequency. Targets whose velocity results in a Doppler shift obscured by aliased clutter are undetectable. The effects of blind velocities in processing are often avoided by altering the pulse repetition frequency over time. The nth blind velocity is

$$v_{blind}^{(n)} = \frac{n \, \lambda \, prf}{2}$$ *eqn 7.14*

Range resolution is given by the width of the autocorrelation of the waveform. This is because matched filter detectors and correlation detectors work by correlating a delayed replica of the transmitted waveform with the received waveform. The time delay corresponding to a high correlation gives the range to the target. For a boxcar pulse, the autocorrelation is a triangle function. For the pulse train, the autocorrelation is a series of triangles with decreasing amplitude (*Figure 7.5*). Two points can be resolved if they are separated by more than the main lobe width

$f_d - 2prf$ $f_d - prf$ f_d $f_d + prf$ $f_d + 2prf$
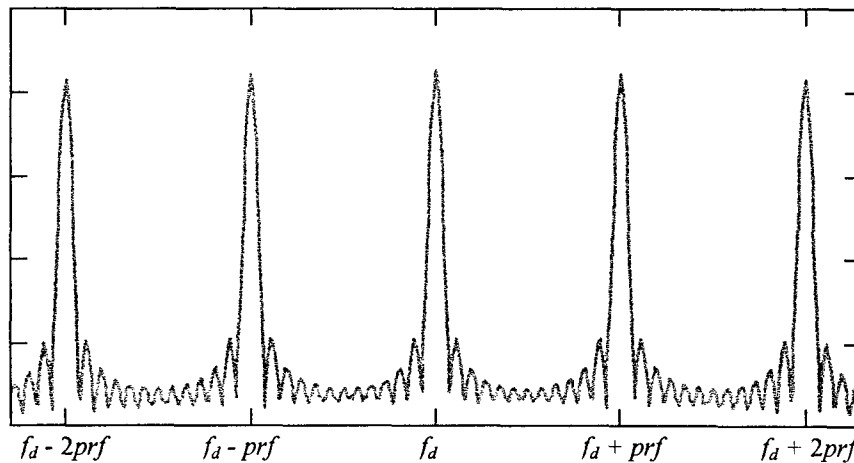
Figure 7.6: Magnitude of the Fourier transform of the pulse train in Figure 7.4.

133
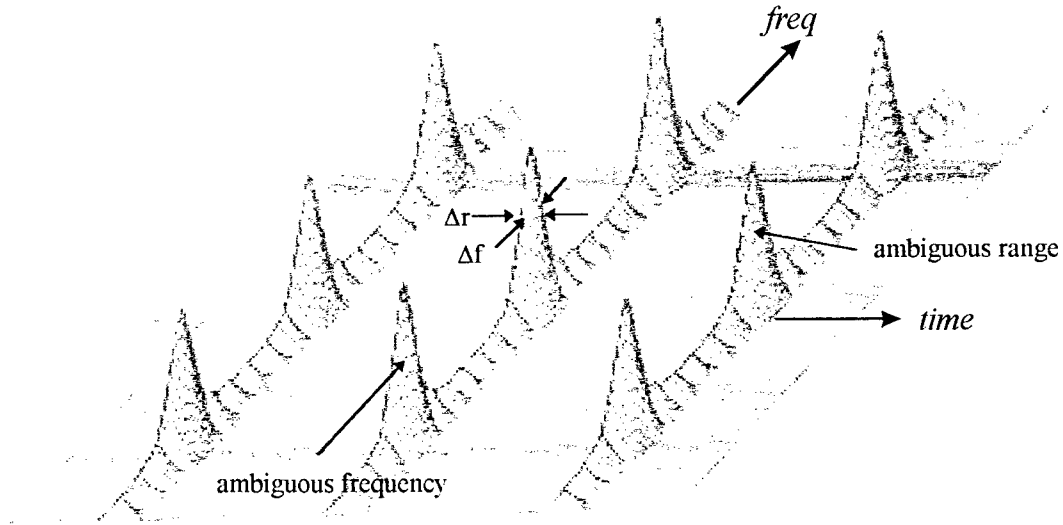
Figure 7.7: Pulse train ambiguity diagram. [13]

of the autocorrelation function. For the pulse train, if the individual pulse length is τ_p, then the range resolution is

$$\Delta r = c\tau_p \qquad\qquad eqn\ 7.15$$

where c is the speed of light. Similarly, the range ambiguity is

$$\Delta r_{am} = \frac{c}{prf} \qquad\qquad eqn\ 7.16$$

Ambiguous ranges fold long range reflections into close range bins. This is a problem for SBR due to the significant clutter frequency dependance on range (*Figure 7.2*). For the clutter processing, we require a *prf* such that the entire footprint is range unambiguous. This does not mean that the range to the footprint must be range unambiguous since no other reflections are expected except from within the footprint. The length of the footprint is the desired unambiguous range.

7.3 Minimum Detectable Velocity

The Doppler shift for a clutter element with platform geometry defined by beam azimuth, β_{sat}, and nadir offset, η_{tgt}, is

$$f_{d,clutter} = \frac{2f_0}{c}v_{sat}\cos\beta_{sat}\sin(\eta_{tgt}) \qquad\qquad eqn\ 7.17$$

Similarly, target Doppler shift with the same geometry and target velocity v_{tgt} is

$$f_{d,tgt} = \frac{2f_0}{c}\left(v_{sat}\cos\beta_{sat} + v_{tgt}\cos\beta_{tgt}\right)\sin(\eta_{tgt}) \qquad\qquad eqn\ 7.18$$

134
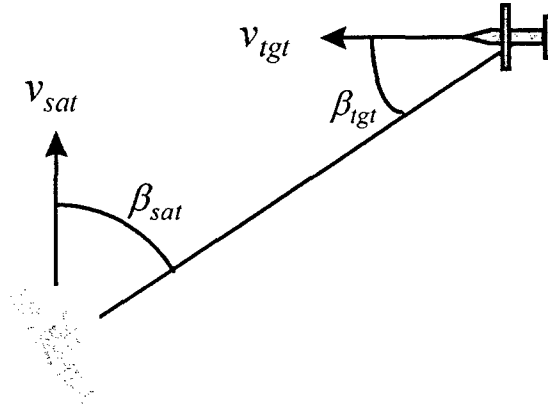
Figure 7.8: Clutter-target geometry.

where β_{tgt} is the target azimuth: the angle between the target velocity vector and the beam direction projected into the plane tangent to the surface of the Earth at the target point.

MTI radar can detect a target if its velocity in *eqn 7.18* is sufficient to shift the target frequency spectrum out of the clutter spectrum. There are four significant factors that effect the clutter spectrum width: cross-range or azimuth spreading, range cell spreading, waveform spectrum, and clutter motion. The minimum detectable Doppler shift is one half the clutter spectrum width. Therefore, the minimum detectable relative velocity is

$$v_{min,rel} = \frac{c}{2f} \frac{\Delta f_{az} + \Delta f_{rc} + \Delta f_{wv} + \Delta f_{cm}}{2} \qquad eqn\ 7.19$$

where f is the transmitted frequency, Δf_{az} is the azimuth clutter spread, Δf_{rc} is the range cell clutter spread, and Δf_{wv} is the waveform frequency resolution. Δf_{cm} is the frequency spread due to clutter motion such as ocean surface waves, wind blown trees, etc. Clutter motion is a random variable that depends on the type of clutter background and may at times be significant.. For the purposes of modeling clutter performance, we do not include the clutter spectrum cue to clutter motion. For a target with beam azimuth β_{tgt} and grazing angle γ_{tgt}, the minimum detectable target velocity (*mdv*) is then

$$v_{min} = \frac{c}{4f} \frac{\Delta f_{az} + \Delta f_{rc} + \Delta f_{wv}}{\cos(\beta_{tgt})\cos(\gamma_{tgt})} \qquad eqn\ 7.20$$

Note that azimuth is measured from the radar's velocity vector to the beam direction projected into the plane of the radar's velocity vector. For target elevation, ε_{tgt}, the grazing angle is given by *eqn 7.2* and *eqn 7.5*.

The maximum azimuth clutter spread occurs at the widest point of the footprint - between the leading and trailing footprint edges - and is due to the variation in azimuth. Thus, the azimuth clutter spread is bounded above by the difference in Doppler shift from *eqn 7.11*. Azimuth spread may also be differentially or variationally derived. The differential Doppler shift due to azimuth is
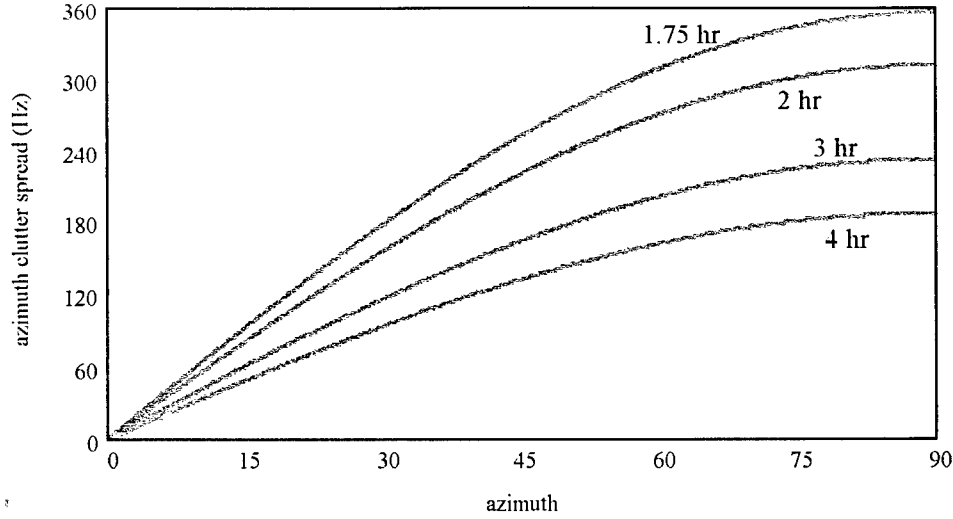
Figure 7.9: Azimuth clutter spread as a function of azimuth for an optimal aperture SBR system at UHF.

$$\partial f_d = \frac{2f v_{orb}}{c}\sin(\beta)\partial\beta \qquad\qquad eqn\ 7.21$$

so that

$$\Delta f_{az} = \frac{2f v_{orb}}{c}\sin(\eta_{tgt})\sin(\beta_{tgt})\frac{c}{fD_{cr}} \qquad\qquad eqn\ 7.22$$

Azimuth clutter spread as a function of azimuth varies significantly. The spread is greatest at broadside when the clutter Doppler shift due to platform motion is a minimum. Spread due to azimuth is minimum along track when the clutter shift due to platform velocity is a maximum. For determining system performance, target azimuth is an important consideration. An expression for mean β_{tgt} as a function of distribution is given below in Section 7.3.1.

There is a degree of control over the range cell clutter spread and the waveform clutter spread by selection of the parameters in the waveform. The individual pulse length in the pulse train governs range resolution. Range cell clutter spread is the clutter frequency shift across a range resolution cell due to variation in elevation or grazing angle of the radar-clutter geometry. The range cell clutter spread is approximately

$$\Delta f_{rc} = \left|\frac{c\tau_p}{l_{foot}}\frac{2f v_{orb}\cos(\beta_{tgt})}{c}\left[\cos(\gamma(\eta_{toe}))-\cos(\gamma(\eta_{heel}))\right]\right| \qquad\qquad eqn\ 7.23$$

Decreasing τ_p improves range resolution and decreases the range cell clutter spread. The individual pulse length is limited below only by the ability of the transmitter to transmit short pulses. Integrated signal energy does not change with decreased pulse length because the peak transmitted power increases for a constant average power assumption.

136

The waveform clutter spread is the result of a finite duration waveform. Only continuous wave transmitters have infinite frequency resolution. For the pulse train, the waveform spread is equivalent to the waveform frequency resolution. Thus,

$$\Delta f_{wv} = \frac{1}{T_L}$$

where T_L is the total pulse integration period.

7.3.1 Mean target azimuth

Beam azimuth to target is an important factor in minimum detectable velocity. *Figure 7.10* shows that the clutter spread varies significantly as a function of target azimuth. It is useful to be able to characterize the statistics of expected target azimuth, particularly as a function of distribution when determining system clutter performance.

For radar platforms and targets that are randomly spaced, the target azimuth is a random variable. For the purpose of a model, assume a uniform azimuth distribution between 0 and $\pi/2$ radians. Note that for the projection of the radar and target velocity vectors onto the beam axis, a distribution between 0 and $\pi/2$ is equivalent to a distribution between 0 and 2π. We limit the distribution to the $[0, \pi/2]$ range since we require a single valued function to find the inverse below.

For a distribution of n_s radar platforms, we are interested in the probability density of the
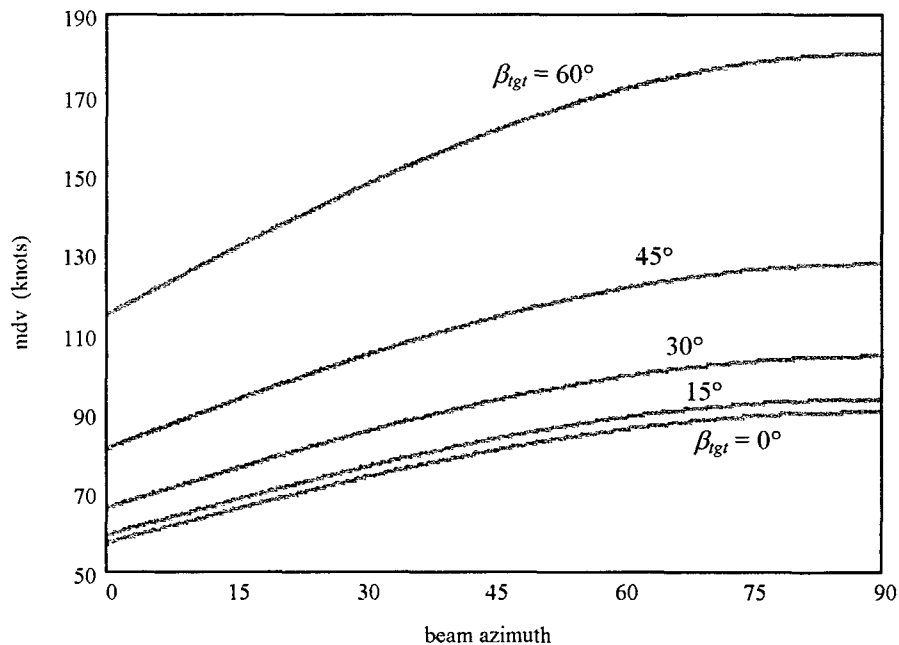
Figure 7.10: Example S-band system MDV as a function of radar beam azimuth. The level curves are for different target azimuths.

137

platform with the minimum azimuth since this gives the minimum azimuth clutter spread. Order statistics are used to characterize the ordered sets of observations. That is, for the set of observations $\{x_1, x_2, ..., x_n\}$ of the random variable x, the ordered set is $\{x_{(1)}, x_{(2)}, ..., x_{(n)}\}$ where $x_{(1)}$ is the minimum element of the set $\{x_1, x_2, ..., x_n\}$, $x_{(2)}$ is the minimum element of the set $\{x_1, x_2, ..., x_n\} - \{x_{(1)}\}$, and $x_{(n)}$ is the minimum element of the set $\{x_1, x_2, ..., x_n\} - \{x_{(1)}, x_{(2)}, ... x_{(n-1)}\}$. Note that equivalently, $x_{(n)}$ is the maximum element of the set $\{x_1, x_2, ..., x_n\}$. The probability density of the kth ordered element is

$$f_{(k)}(x) = \frac{n!}{(k-1)!(n-k)!}[F(x)]^{k-1} f(x)[1-F(x)]^{n-k} \qquad eqn\ 7.25$$

where $f(x)$ is the probability density for the random variable x and $F(x)$ is the cumulative probability function for x. Thus, the first ordered element probability density is

$$f_{(1)}(x) = \frac{n!}{(n-1)!} f(x)[1-F(x)]^{n-1} \qquad eqn\ 7.26$$

For the random variable θ distributed uniformly on the interval $[0, \pi/2]$, the first minimum element of n observations is

$$f_{(1)}(\theta) = n\frac{2}{\pi}\left[1 - \frac{2}{\pi}\theta\right]^{n-1} \qquad eqn\ 7.27$$

The mean value of the minimum observation is then

$$E[\vartheta_{(1)}] = \int_0^{\pi 2} n\frac{2}{\pi}\left(1 - \frac{2}{\pi}\theta\right)^{n-1} \theta\,d\theta = \frac{1}{n+1}\frac{\pi}{2} \qquad eqn\ 7.28$$

Therefore, the mean target azimuth as a function of distribution is

$$\mu_{\beta tgt}(n_s) = \frac{1}{n+1}\frac{\pi}{2} \qquad eqn\ 7.29$$

138

8. SBR Architecture Design and Optimization Results

The analysis from the proceeding chapters is combined into a single model. The model was implemented[1] in MATHCAD® 6.0, a software tool that is particularly suited for parametric and functional analysis. A summary of the system architecture design and optimization process - how the model is implemented and the analysis of the proceeding chapters is synthesized - is presented in Section 8.2. Because of the parametric nature of the cost model, it is a powerful tool by which to examine a wide range of scenarios and possibilities. We solve for the optimum values of four primary architecture variables (Section 8.3) to determine the optimal system architecture: power, aperture, orbital altitude, and distribution. All other system parameters, e.g. constellation size, mass, probability of detection, are derived from these optimum values with the system relationships developed in the preceding chapters. The optimal values for power and aperture, orbital altitude, and distribution are closely examined in Section 8.3.1, Section, 8.3.2, and Section 8.3.3. The radar transmitter frequency tradeoff is also important, Section 8.3.4. One very useful capability of the analysis approach is that it allows characterization of system cost as a function of system performance, Section 8.5, which can be used in developing the system performance requirements. Cost sensitivity to various system parameters, Section 8.4, may also identify key technology drivers and indicate where research and development money should be invested. Finally, we also examine the system performance in a clutter background, Section 8.6.

Table 8.1: SBR example theatre task list. From Phillips Lab SBR IPT requirements definition [51].

task	search area	update interval	time constant	notes
aircraft detection	$2.5 \ 10^5 \ \mathrm{km}^2$	10 sec	45 min	track all aircraft in theatre
monitor armored fighting vehicles	$4 \ 10^4 \ \mathrm{km}^2$	4 hrs	2-50 hrs	tank transporters and armored vehicles (6-80 km/hr)
monitor naval activity	$5 \ 10^4 \ \mathrm{km}^2$	4 hrs	constant	port activity, surface ships, subs
monitor air activity	$2.5 \ 10^5 \ \mathrm{km}^2$	.5 hrs	constant	monitor airfields. hand-off departing aircraft
monitor fixed ground sites	$2.5 \ 10^5 \ \mathrm{km}^2$	4 hrs	constant	bridges, roads, rail, Army areas of assembly
detect missile sites	$250 \ \mathrm{km}^2$	30 min	constant	SAR image, recognize targets
weather	$1 \ 10^6 \ \mathrm{km}^2$	6 hrs	constant	environmental conditions

[1] Requests for electronic copies of the model should be directed to Prof Daniel Hastings, MIT Department of Aeronautics and Astronautics, 77 Massachusetts Ave, Cambridge, MA, 02139, or hastings@mit.edu.

Table 8.2: SBR mission task for AWACS replacement.

mission task	search area	update interval
aircraft detection	$1\ 10^6\ \mathrm{km}^2$	10 sec

8.1 System Requirements

The scale of a system depends ultimately on the system requirements. Requirements definition is one of the most important steps in the systems engineering process. For SBR, the system requirements are still largely uncertain since no consensus has been reached as to the desired mission. A space based radar could perform an AWACS type mission (AMTI), a JSTARS type mission (GMTI), a SAR imaging mission, or some combination of all or some of these. An example of an SBR theatre task list being used by AF Phillips Lab is given in _Table 8.1_.

The most important critical element of the Phillips Lab SBR IPT's requirements is the aircraft detection and tracking. Their operations concept calls for continuous track of all airborne targets in the theatre. Traditional track from a space based platform is very difficult for the reasons outlined in Section 5.1. A better paradigm which gives the same performance is a track while search with a search mean time to detection on the order of the track update rate. This is the approach that we adopt for establishing the system requirements. The requirements given in _Table 8.2_ are based on discussions [39] with the USAF Space and Missile Center (SMC) development branch (SMC/XR). Note that the top level requirements given in _Table 8.2_ consist

Table 8.3: Baseline system requirements and parameters.

System requirements			source
area to search	A_{search}	$10^6\ \mathrm{km}^2$	SMC/XR
mean time to detection	T_d	10 sec	SMC/XR
radar cross section	RCS	$1\ \mathrm{m}^2$	SMC/XR
configuration probability of detection	P_c	.95	reasonable
min acceptable performance degradation	P_e	.85 and .9	reasonable
max acceptable performance degradation	$1 - P_e$	.15 and .1	reasonable
probability of false alarm	P_{fa}	10^{-10}	reasonable
false alarm rate per MHz bandwidth	far	.36 per hr	reasonable
m-of-n detection criteria	m,n	3,5	reasonable
payload (power aperture) mass fraction	m_{PA}/m	.5	reasonable
mission reliability/availability	R_{ms}	.99	reasonable
baseline spacecraft reliability	R_{sc}	.75	reasonable

of only the aircraft detection task but over an area 4 times the requirement given in *Table 8.1*.

The baseline mission requirements are given in *Table 8.3*. The area to be searched (A_{search}), the mean time to detection (T_d), and the radar cross section (*RCS*) requirements were established by SMC/XR for sizing SBR concept studies [39,52]. The other parameters were selected as reasonable values for the purpose of sizing the complete SBR system.

8.2 System Architecture Design and Optimization Process

This section summarizes the implementation of the cost model. Although all of the analysis has been described in the proceeding Chapters, how each of the parts and submodels are integrated in the system architecture optimization process has not. The results presented in this Chapter employ the model as described below. When adapting the analysis for other systems or even other studies of space based radar systems, subsections of the model can be used or eliminated as needed to study different effects of the system architecture design. For a detailed explanation of the equations or parameters, refer to the indicated section or equation.

System Architecture Design and Optimization Process:

1. Establish system performance requirements (Section 8.1):
   mean time to detection, T_d; probability of false alarm, P_{fa}; area to be searched, A_{search}; and minimum radar cross section, σ.

2. Establish radar system parameters (Section 2.4):
   radar system losses (6 dB), receiver antenna temperature (300 K)

3. Calculate power mass density parameter (*eqn 3.7*): $\rho_P = \dfrac{\rho_{SA}(1+\mu_{ms})}{(1-\mu_e)\eta_{SA}\eta_{store}} + \dfrac{\rho_c(1+\mu_{ms})Per}{\eta_{store}DOD}$

   (nominal value: 1/25 kg/W)

4. Calculate aperture mass density parameter (*eqn 3.11*): $\rho_A = \dfrac{4m_{T/R}f^2}{c^2}$

   (nominal value: 15 kg/m^2)

5. Create two-sigma maximum range to target function (*eqn 6.32*):

   $R_t(a,n) = \left(1 + \dfrac{.082+.587a^{-4}}{\sqrt{n}}\right)\left(-3.1+3.188a^{12}\right)$ DU

6. Create ellipse factor function (*eqn 2.51*): $k_A = \dfrac{R_E\left[\lambda(\varepsilon_{toe}) - \lambda(\varepsilon_{heel})\right]}{R_{tgt}\Delta\theta}$

7. Create the radar range proportionality constant function, K_s (*eqn 2.57*): $K_s = \dfrac{\sigma \eta_{int} k_A \pi}{4^3 R^2 kTB_n \tau L}$

8. Calculate the reliability cost factor function (*eqn 4.15*):

$$\phi_R(n_s) = \begin{cases} 1 + \dfrac{\ln\left(1 - R_m(n_s)\right) - \ln\left(1 - R_{sc}\right)}{\ln\left(1 - R_{sc}\right)} & if\ R_{sc} < R_m(n_s) \\ 1 & if\ R_{sc} \geq R_m(n_s) \end{cases}$$

where (*eqn 4.16*) $R_m(n_s) = \left\{ p \ni R_{ms} - \displaystyle\sum_{k=m(n_s)}^{n_s} \dfrac{n_s!}{k!(n_s - k)!} p^k (1-p)^{n_s - k} = 0 \right\}$

and where (*eqn 4.17*) $m(n_s) = \left\{ m' \ni P_e - \dfrac{1 - (1 - P_c)^{m' n_s}}{P_c} = 0 \right\}$

9. Calculate the area search rate (*eqn 5.21*): $ASR = \dfrac{m A_{search}}{n_s P_d T_d}$

10. Calculate the signal to noise ratio for the configuration probability of detection (by iterative process) (*eqn 5.9*):

$$P_d = \tfrac{1}{2}\left(1 - erf\left(\sqrt{\ln\left(\dfrac{1}{P_{fa}}\right)} - \sqrt{\dfrac{S}{N}}\right)\right) + \dfrac{\exp\left(-\left(\sqrt{\ln\left(\dfrac{1}{P_{fa}}\right)} - \sqrt{\dfrac{S}{N}}\right)^2\right)}{4\pi\sqrt{\dfrac{S}{N}}}\left[\dfrac{3}{4} - \dfrac{\sqrt{\ln\left(\dfrac{1}{P_{fa}}\right)}}{4\sqrt{\dfrac{S}{N}}} + \dfrac{1 + \left(\sqrt{2\ln\left(\dfrac{1}{P_{fa}}\right)} - \sqrt{2\dfrac{S}{N}}\right)^2}{16\sqrt{\dfrac{S}{N}}} - \cdots\right]$$

11. Establish power subsystem development cost factor (4) and create power cost multiplier function (*eqn 3.3*): $\phi_P = \phi_R\left(n_s, \{R_i\}\right)\phi_{P_0} + \phi_d(a)$

12. Calculate optimal power (*eqn 3.30*): $P^* = \left(\dfrac{\alpha_A \phi_A \rho_A^{\alpha_A}}{2\alpha_P \phi_P \rho_P^{\alpha_P}}\right)^{\frac{1}{2\alpha_P + \alpha_A}} \left(\dfrac{ASR\ S\ N}{K_s}\right)^{\frac{\alpha_A}{2\alpha_P + \alpha_A}}$

13. Calculate power cost function (*eqn 3.2*): $cost_{power} = \phi_P(P_t, \{C_i\}, \{R_i\})\left(\left[P_t(\rho_P P_t)\right]^{\alpha_P} + \beta_P\right)$

14. Establish aperture subsystem development cost factor (4) and create aperture cost multiplier function (*eqn 3.9*): $\phi_A = \phi_R\left(n_s, \{R_i\}\right)\phi_{A_0} + \phi_d(a)$

15. Calculate optimal aperture (*eqn 3.31*): $A^* = \left(\dfrac{2\alpha_P \phi_P \rho_P^{\alpha_P}}{\alpha_A \phi_A \rho_A^{\alpha_A}}\right)^{\frac{1}{2\alpha_P + \alpha_A}} \left(\dfrac{ASR\ S\ N}{K_s}\right)^{\frac{2\alpha_P}{2\alpha_P + \alpha_A}}$

16. Calculate aperture cost function (*eqn 3.8*): $cost_{aperture} = \phi_A\left(A,\{C_i\},\{R_i\}\right)\left(\left[\rho_A A\right]^{\alpha_A} + \beta_A\right)$

17. Establish bus subsystem development cost factor (4) and create bus cost multiplier function (*eqn 3.13*): $\phi_{bus} = \phi_R\left(n_s,\{R_i\}\right)\phi_{bus_0} + \phi_d(a)$

18. Calculate bus mass (*eqn 3.17*): $m_{bus} = \left[1 + \mu_{hard}\phi_{hard}(a)\right]\dfrac{1-\mu_{PA}}{\mu_{PA}}\left(\rho_P P_t + \rho_A A\right)$

19. Calculate bus cost function (*eqn 3.12*):

$$cost_{bus} = \phi_{bus}\left(\{C_i\},\{R_i\}\right)\left(\left[m_{bus}\left(P_t,A,\{C_i\}\right)\left(1+\mu_{hard}\phi_{hard}(a)\right)\right]^{\alpha_{bus}} + \beta_{bus}\right)$$

20. Create launch cost multiplier function (*eqn 3.20*): $\phi_l = \phi_R\left(n_s,\{R_i\}\right)\phi_{l_0}$

21. Create booster mass factor function (*eqn 3.23*):

$$\phi_{boost}(a) = \exp\left[\left(\sqrt{\frac{\mu}{a}} - \sqrt{\frac{\mu}{a_{leo}}} + \sqrt{\frac{2\mu}{a_{leo}} - \frac{2\mu}{a_{leo}+a}} - \sqrt{\frac{2\mu}{a} - \frac{2\mu}{a_{leo}+a}}\right)\frac{1}{gI_{sp}}\right]$$

22. Calculate launch cost (*eqn 3.19*): $cost_{launch} = \phi_l\left(m_{S/C},\{C_i\},\{R_i\}\right)\left(\left[\phi_{boost}(a)\left(m_{S/C}\right)\right]^{\alpha_l} + \beta_l\right)$

23. Design constellations to meet coverage requirements (Section 6.2) and create altitude coverage function (*eqn 6.18*): $acf(a,n) = \left(\dfrac{\left(\beta_1 a^{\alpha_{cf}} + \beta_0\right)}{1 + 2\sigma_{cov}\left(lat_{min,mean}\right)}\right)\left(n + \dfrac{2\sigma_{cov}\left(lat_{min,mean}\right)}{\sqrt{n}}\right)$

24. Calculate constellation size multiplier (*eqn 3.35*): $\phi_N = \left[acf(a,n_s)\right]^{1+\ln(1-dis)\cdot\ln(2)}$

25. Calculate total system IOC cost (*eqn 3.1*):

$$IOC\ cost = \phi(N)\ cost_{sat} = \phi(N)\left(cost_{power} + cost_{aperture} + cost_{bus} + cost_{launch}\right)$$

26. Iterate over range of distribution, orbital altitude to determine optimal system architecture.

27. Check optimal system performance in clutter (*eqn 7.20*): $v_{min} = \dfrac{c}{4f}\dfrac{\Delta f_{az} + \Delta f_{rc} + \Delta f_{wv}}{\cos\left(\beta_{tgt}\right)\cos\left(\gamma_{tgt}\right)}$

where (*eqn 7.22*) $\Delta f_{az} = \dfrac{2fv_{orb}}{c}\sin\left(\eta_{tgt}\right)\sin\left(\beta_{tgt}\right)\dfrac{c}{fD_{cr}}$

where (*eqn 7.23*) $\Delta f_{rc} = \left|\dfrac{c\tau_p}{l_{foot}}\dfrac{2fv_{orb}\cos\left(\beta_{tgt}\right)}{c}\left[\cos\left(\gamma\left(\eta_{toe}\right)\right) - \cos\left(\gamma\left(\eta_{heel}\right)\right)\right]\right|$

where (*eqn 7.24*) $\Delta f_{wv} = \dfrac{1}{T_L}$

and where (*eqn 7.29*) $\mu_{\beta tgt}(n_s) = \dfrac{1}{n+1}\dfrac{\pi}{2}$

8.3 Optimal System Architecture

The optimal system architecture is determined by iterating through the design process described in Section 8.2. Four architecture optimal variables are determined: power, aperture, orbital altitude, and distribution. For examination, we size two different systems for the two reasonable performance effectiveness parameters given in *Table 8.3*: .90 and .85. (In Section 8.3.3, system cost and distribution as a function of performance degradation is analyzed.) The system with 15% performance degradation is designated SSL 1; and the system with 10% performance degradation is designated SSL 2. The optimal systems are summarized in *Table 8.4*. Detailed analysis of the optimum altitude and distribution is given in Section 8.3.2 and Section 8.3.3 respectively. The optimu.. trade off between power and aperture is determined by the power aperture optimality condition which was derived in Section 3.6.

One of the powerful uses of the system cost model is that different concepts can be quantitatively compared on the same basis. The optimal systems SSL 1 and SSL 2 are compared with four other SBR concepts in *Table 8.5* that have roughly similar missions. SMC 1 and SMC

Table 8.4: Optimal system architecture summary for SSL 1 and SSL 2. (*'s indicate optimal solutions; all other quantities are derived from the optimum values.)

	SSL 1	SSL 2
IOC cost (B)	$7.27	$7.35
power (kW), P*	6.4	8.2
aperture (m^2), A*	50	64
dry mass (kg)	2000	2600
wet mass (kg)	2500	3200
semimajor axis $(DU)^2$, a*	1.11	1.11
altitude (km)	700	700
period (hr)	1.65	1.65
configuration size, n*	5	3
constellation size, N	176	107
sat prob of detection	.45	.63
sat S/N (dB)	8.5	9.5

[2] A distance unit (DU) is equal to one Earth radii, or 6378 km.

144

Table 8.5: SBR concepts system and performance comparison. SMC proposals are the from the SMC Space Sensor Study. SPEAR is the Space Electronically Agile Radar concept developed by the Phillips Lab SBR IPT. STARLITE is a DARPA SBR concept for providing tactical synthetic aperture radar images. Iridium is shown only for comparison of the cost model output.

	alt (km)	P (hr)	P (kW)	A (m²)	T_d (s)	ASR (km²/s)	constellation size (spec)	SSL model	system cost (SSL model)
SMC 1	2600	2.35	25	2.275	4.1	2.6×10^5	(26)	17 (1x)	$61 B
SMC 2	10.000	6	30	12.300	3.6	2.9×10^5	(9)	8 (1x)	$34 B
SMC 1 resized	2600	2.35	40	300	10	1.0×10^5		17 (1x)	$15 B
SMC 2 resized	10.000	6	100	800	10	1.0×10^5		8 (1x)	$11 B
SPEAR	850	1.7	20	132	19.8	5.3×10^4	(7)	36	$9.7 B
STARLITE	770	1.67	6	15	520	2.0×10^3	(24)	38	$3.1 B
SSL 1	700	1.65	6.4	49	10	1.3×10^5		176	$7.27 B
SSL 2	700	1.65	8.2	64	10	1.6×10^5		107	$7.35 B
Iridium	775	1.675	.9	6	claimed	cost: $1B	66		$2 B

2 are the proposals from the SMC Space Sensor Study for the 1995 Corona conference [52,7]. SPEAR is the Space Electronically Agile Radar concept developed by the Phillips Lab SBR IPT [51]. STARLITE is a DARPA SBR concept for providing tactical synthetic aperture radar (SAR) images with air surveillance as a possible mission adjunct. Iridium is the 66 satellite constellation that will provide global cellular phone service and is included only for comparison. Iridium claims to be able to procure and deploy the entire constellation for approximately $1 billion. The SBR cost model predicts $2 billion IOC cost for an Iridium type system. Although accuracy in the absolute cost predicted by the model was never desired, the model gives reasonable results. Nevertheless, the IOC cost model should only be used to compare systems on a relative basis.

Note the system optimization process resulted in systems that give a 75-85% reduction in system cost compared to the SMC concepts. Part of this is due to the relaxed mean time to detection requirement. The SMC concepts were sized on a requirement for area search rate based upon a certain number of AWACS coverage regions updated at a certain rate. This resulted in a slightly over-capable system. However, even when the SMC concepts are resized to a 10 sec mean time to detection requirement (and employing the power aperture optimality condition), the reduction in cost ranges from 35-50%.

Table 8.5 very clearly shows the utility of system architecture optimization process. The optimized systems, SSL 1 and SSL 2, give cost savings of 35-50% for the same level of performance compared to the SMC architectures. The SPEAR concept architecture has roughly half the performance capability of the SSL architectures and costs almost 33% more. The STARLITE concept costs less than half than the SSL systems but at the unacceptable price of a 50-fold decrease in performance. Although SSL 1 and SSL 2 were optimized with respect to the cost model and thus are expected to be best under this particularly metric, the cost model can

easily be changed and the optimization reaccomplished. The point is that the system architecture optimization process is a very powerful analysis tool in developing the system concept that gives the best performance for cost.

Area search rate is not a good metric for comparing systems with distributed operations concepts, e.g. SSL 1 and SSL 2. The quantity given in *Table 8.5* is the distributed area search rate given by *eqn 5.22*. The configuration total area search rate would be the product of the configuration size - the number of distributed satellites - and area search rate. These problems in definition can be avoided by using mean time to detection as the performance metric. This is the convention we have adopted for discussion.

SSL 2 is more expensive than SSL 1 because it is a more capable system. Although both have the same baseline requirements, SSL 1 has a higher allowable performance degradation and thus has an effectively decreased level of performance. The only difference between SSL 1 and SSL 2 is the allowable degradation, but this changes the optimum distribution and subsequently the required power aperture product.

8.3.1 Optimal Power and Aperture

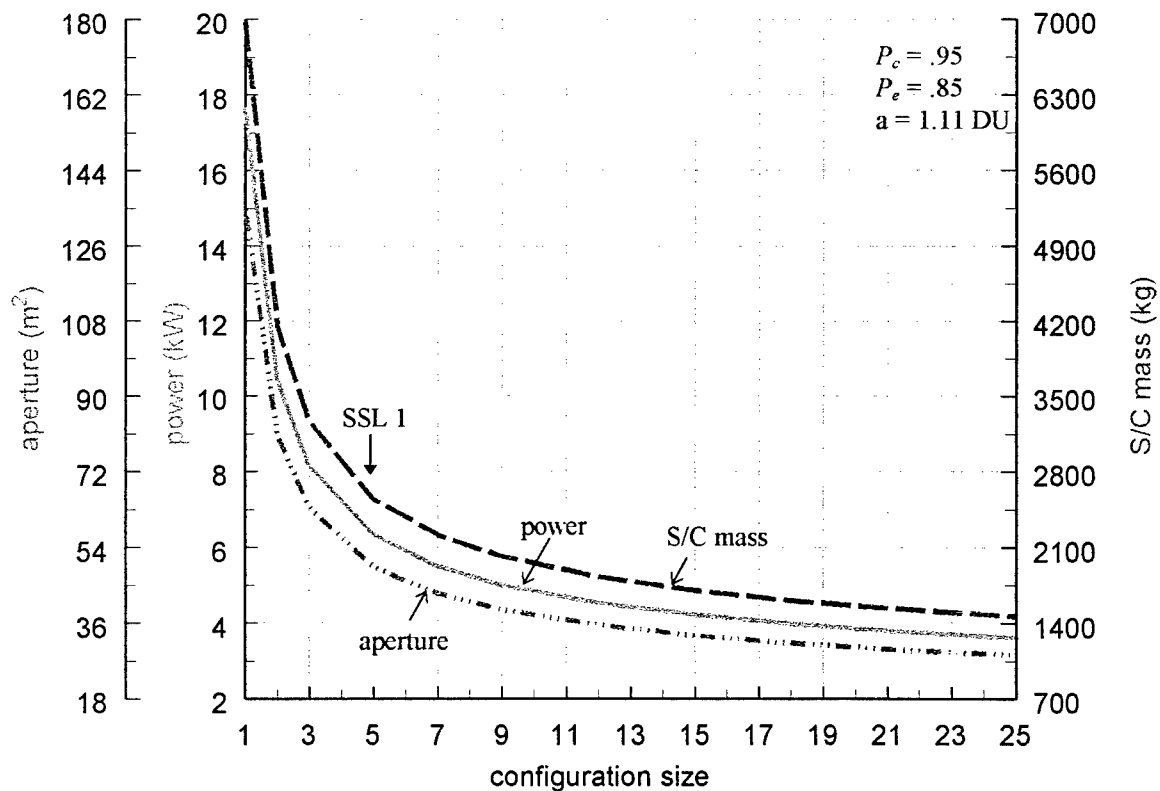The optimum power and aperture is determined by the power aperture optimality condition

Figure 8.1: Optimal power, aperture, and spacecraft mass as a function of the configuration size. Configuration size is a the number of distributed satellites.

derived in *Section 3.6*. The optimum power, aperture, and spacecraft mass as a function of distribution for the 15% performance degradation specification is given in *Figure 8.1*.

As the distribution decreases, the power and aperture required on each satellite increases. For distributions of less than three satellites, the increase in power, aperture, and mass is substantial. For distributions greater than seven satellites, the marginal decrease in power, aperture, and mass quickly diminishes. Indeed, all optimal systems examined below have optimum distributions of between three and seven satellites. It is in this region of distribution where the benefits due to decreased power, aperture, and mass balance with disadvantages due to increased constellation size and operational complexity. This is proven below but is already intuitively apparent in *Figure 8.1*.

Because power and aperture are used extensively to characterize clutter performance, it is useful to determine the empirical optimal power and aperture as a function of altitude and distribution. This obviates the need to recompute the iterative solution to the cost metric every time power and aperture are desired. The empirical, optimum power and aperture are regressed to

$$P^*(a,n) = \left(\beta_1^{(P)} + \frac{\beta_2^{(P)}}{n} \right) + \left(\beta_3^{(P)} + \frac{\beta_4^{(P)}}{n} \right) a^{1\,2} \qquad\qquad eqn\ 8.1$$

$$A^*(a,n) = \left(\beta_1^{(A)} + \frac{\beta_2^{(A)}}{n} \right) + \left(\beta_3^{(A)} + \frac{\beta_4^{(A)}}{n} \right) a^{1\,2} \qquad\qquad eqn\ 8.2$$

with better than 90% coefficient of correlation. The parameter values are given in . Note that optimum power and aperture scale with the same semi-major axis power law exponent as radar slant range to target *eqn 6.32*. (That is, both the empirical slant range and optimum power/aperture power laws scale as the square root of semi-major axis.) The semi-major axis power law exponent in *eqn 8.1* and *eqn 8.2* is in fact due to the scaling of *eqn 6.32*. Since power and aperture increase with the square root of orbital radius, the power aperture product increases linearly with orbit semimajor axis. It is this linear increase that drives the optimal orbit into low Earth orbit (*Section 8.3.2*).

Table 8.6: Empirical parameters for optimum power and aperture relationships.

|	$P*$	$A*$
β_1	-42.4	-320.6
β_2	-104.8	-760.7
β_3	43.5	330.6
β_4	113.3	827.6

8.3.2 Optimum Altitude

The optimum - minimum cost - altitude is not a function of mission requirements or distribution in that the optimal point does not vary with these. System cost as a function of orbit semimajor axis is plotted in *Figure 8.2*. In the cost model, there are primarily five factors that drive cost as a function of orbit semimajor axis:

1. Increased drag with decreased altitude for leo constellations.

2. Increased constellation size at decreased altitude.

3. Power and aperture increase with the square root of orbital radius. The power-aperture product increases linearly with orbital radius. (Section 8.3.1)

4. Launch costs increase with orbital altitude.

5. Radiation hardening is required for orbits in the radiation belts.

The effects of each of these five factors is apparent in *Figure 8.2*. The increased propellant mass required to counter atmospheric drag results in the significant increase in system cost below 1.11 DU. From 1.1 DU to approximately 1.25 DU, the increasing cost is due
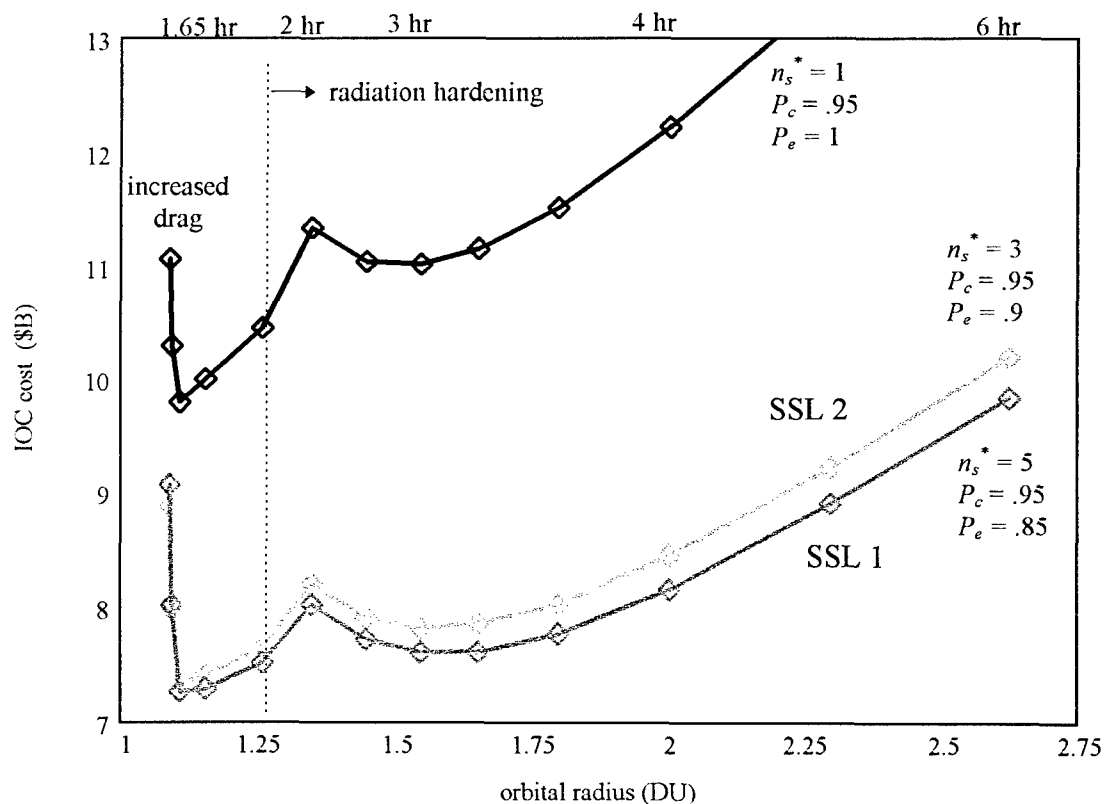
Figure 8.2: Cost as a function of orbital radius. The optimum levels of distribution are used. Learning curves are 15%. The minimum cost occurs at 1.11 DUs (700 km altitude). Cost for a system with no distribution is shown for comparison.

primarily to the linear increase in power-aperture product with orbital radius. There is a slight cost reduction due to the decreased number of satellites required to provide coverage from higher altitudes. The sharp increase in system cost from 1.25 DU to peak at 1.3 DU is due to the increased hardening required for orbits within the Van Allen belts. The inner peak of the radiation belts occurs at 1.3 DU (Section 3.4). The system cost decreases to a local minimum at 1.66 DU (3 hr orbits) as the required spacecraft hardening mass decreases from the peak. The increase above 1.66 DU is again due to the linearly increasing power-aperture product.

8.3.3 Optimum Distribution

Distribution is the most significant contribution to system optimality. Optimal distribution is a function of orbital altitude, learning curve discount, configuration probability of detection, and desired performance effectiveness. Except for extremely stringent requirements for performance degradation, *some distribution is always advantageous*. The level of "stringent" depends on performance capability and is quantified below. The factors that drive cost as a function of distribution (*Figure 8.4*) are

1. Reliability/survivability cost decreases with increasing distribution.

2. Increased coverage and increased probability for favorable target viewing angles with increased distribution.

3. Design efficiency is increased with distribution due to decrease in target range variability.

4. Decreased power-aperture product per satellites with increased distribution.

5. Number of satellites and launch costs increase with distribution.

6. Marginal decrease in production learning curves with increasing distribution.

Some of the advantages due to distribution are obvious, e.g., increased reliability and decreased power-aperture product per satellite. The are also several subtle advantages that nonetheless contribute to the reduction in system cost. Variability of range to target decreases with increased distribution (Section 6.5.4) which results in a more efficient design since the power aperture product is determined from the square of the expected range. Coverage efficiency improves in a similar manner (Section 6.5.3). For the same minimum coverage requirements (configuration probability of detection specification), constellations with higher levels of distribution require slightly smaller constellations than implied by linear scaling. Distribution also increases the target perspectives which improves performance in clutter (Section 8.6).

Performance effectiveness is a substantial driver of distribution. Optimal distribution as a function of performance effectiveness for a 95% configuration probability of detection
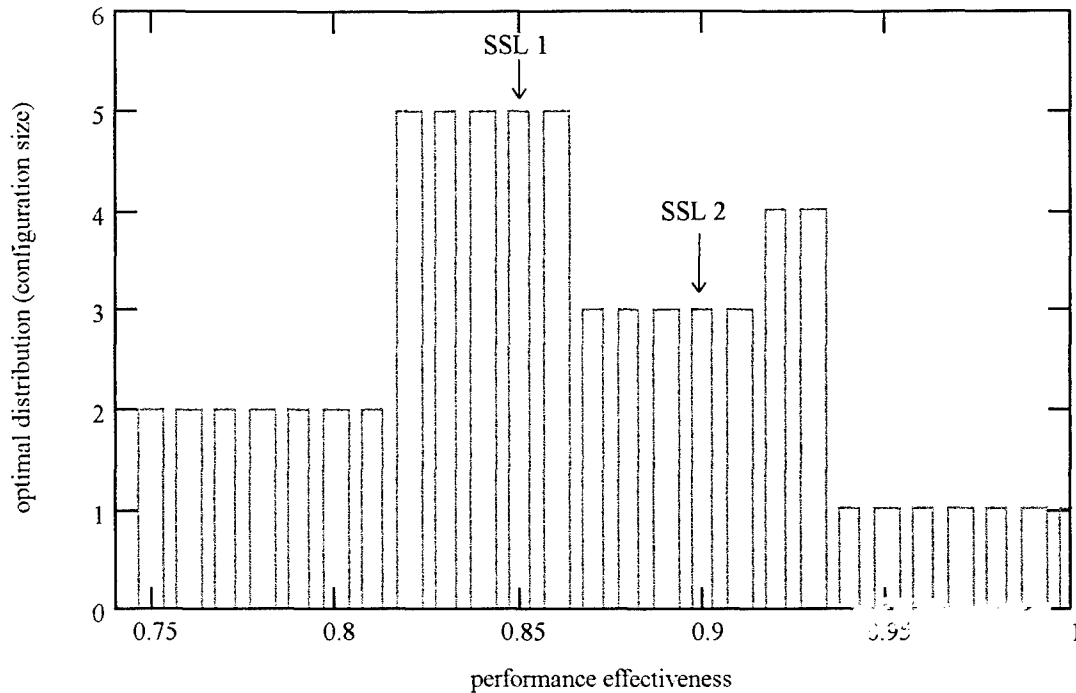
Figure 8.3: Optimal distribution as a function of performance effectiveness for a 95% configuration probability of detection capability.

requirement is plotted in *Figure 8.3*. The step nature of the relationship is due to cost plateaus in the reliability cost factor (Section 4.4). The optimum distribution ranges from 2 to 5 satellites for all levels of performance effectiveness less than 94%. For values of performance effectiveness greater than 94%, the optimum distribution is one satellite – that is, no distribution. This is because high values of performance effectiveness correspond to small values maximum acceptable performance degradation. If little or no performance degradation is acceptable, then there is no advantage in reliability for distributing system functionality. Recall (Chapter 4) that a low acceptable performance degradation is unlikely since the performance requirements for most systems are padded (over specified) to ensure a minimum level of performance in uncertain conditions. The performance effectiveness is the original, unpadded performance requirement. Optimal distribution for other configuration probability of detection's is plotted in *Figure 8.5*.

In *Figure 8.3* and *Figure 8.5*, note that the value of performance effectiveness at which the optimum distribution transitions from distribution to no distribution decreases as the configuration probability of detection decreases. The configuration probability of detection is the capability of the system. As system capability decreases, the leverage in reliability due to distribution decreases which decreases the distribution performance effectiveness transition point. This is inherent in the definition of performance effectiveness, which is a function of configuration probability of detection, *eqn 4.3*: $P_e = \dfrac{1-\left(1-P_d\right)^m}{P_c}$. Note that the satellite probability of
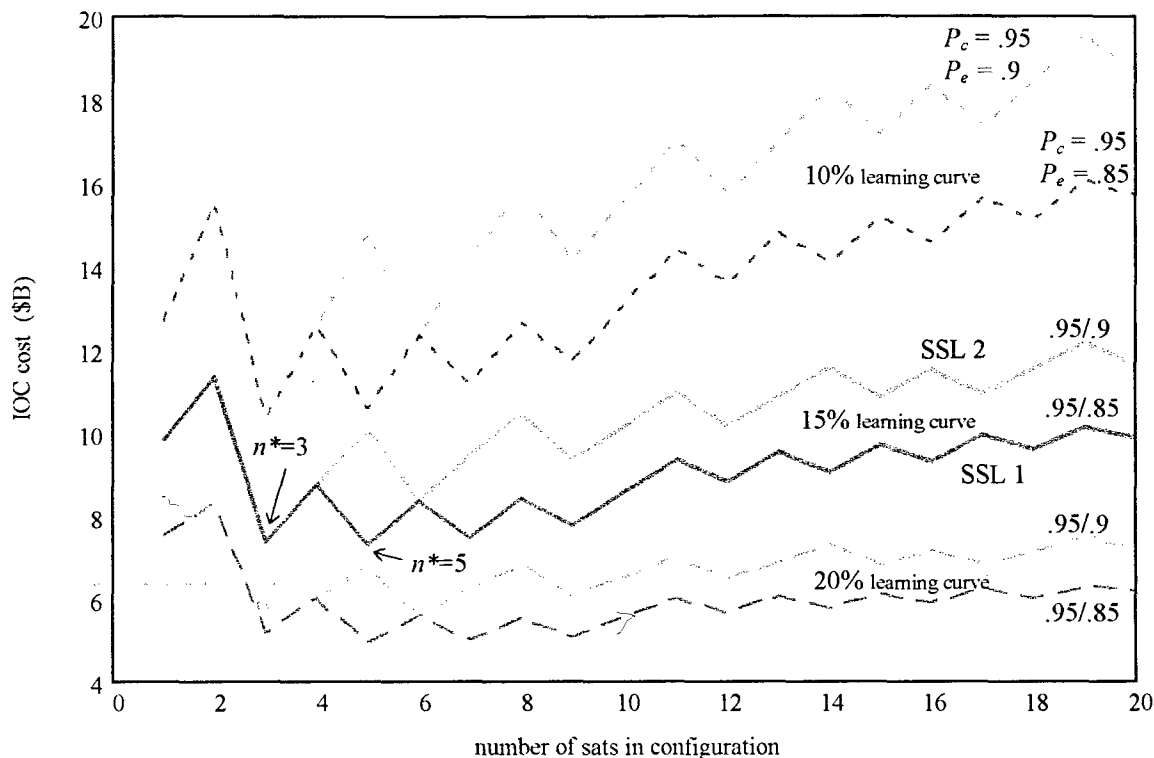
150

Figure 8.4: Cost as a function of distribution for SSL 1 and SSL 2 for a range of learning curves and reliability parameters.

detection, P_d, is a function of the configuration probability of detection, P_c, *eqn 4.2*. For an incremental decrease in configuration probability of detection, the satellite probability of detection decreases by a larger increment. Thus, performance effectiveness decreases with configuration probability of detection by its definition. It follows that the distribution transition point should decrease as configuration probability of detection decreases.

For radar, distribution is most advantageous for a mission with a high required capability, that is, configuration probability of detection. As capability decreases, so does the transition point at which distributed architectures are more cost effective. For less capable mission requirements, a distributed radar system would have to have a large acceptable performance degradation which decreases the expected performance further. This is not necessarily generalizable to other distributed systems. Radar has the unique attribute that performance can be measured as a percentage: probability of detection. The behavior of the performance effectiveness transition point is an aspect of the interrelationship of the system probabilities. Note also that conclusion that distribution is best for highly capable systems does not diminish the utility of distribution for the space based AWACS mission. The probability of detection requirement, either as a single look or cumulative probability, will certainly be large enough to gain benefits from distribution. The probability of detection must be high because of the very critical mission performed by the surveillance platform; see Section 5.3.2.
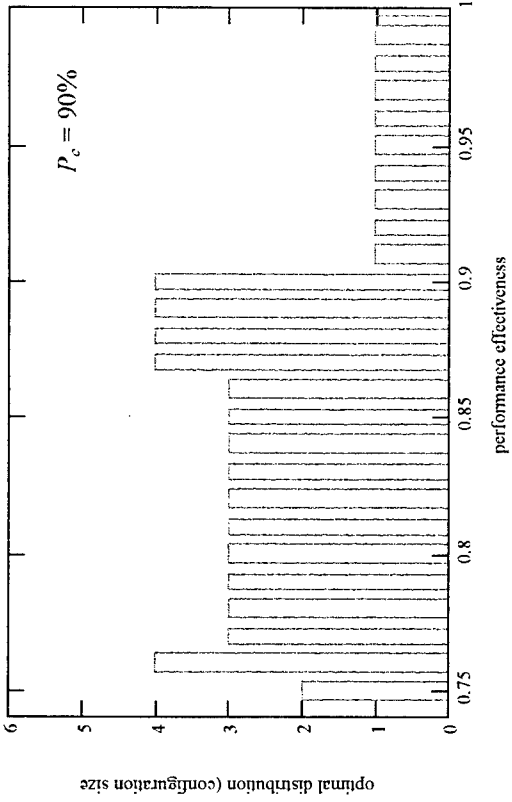
A: *Optimal distribution for 90% configuration probability of detection.*
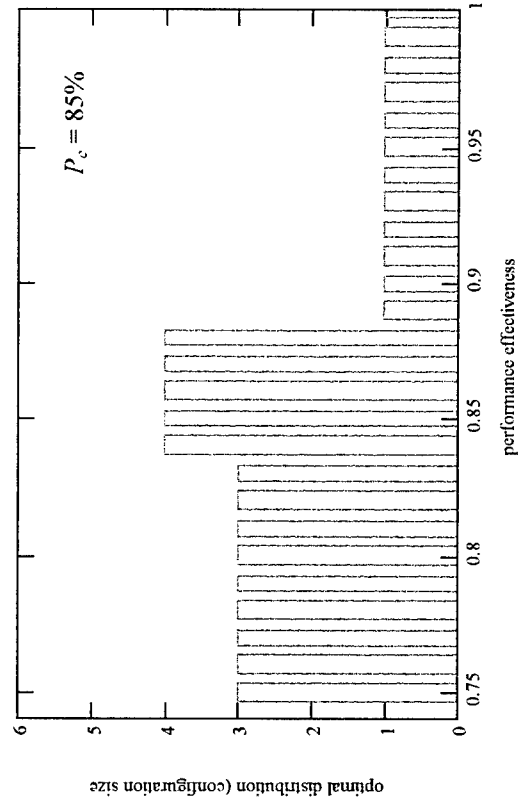
B: *Optimal distribution for 85% configuration probability of detection.*
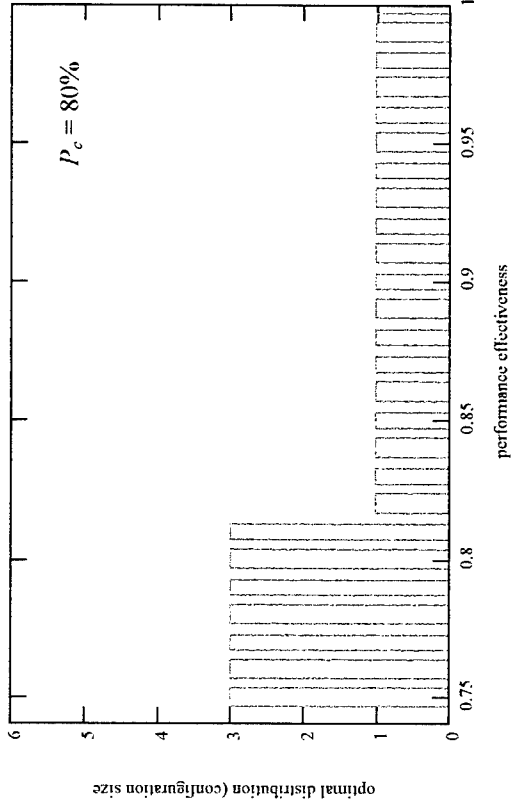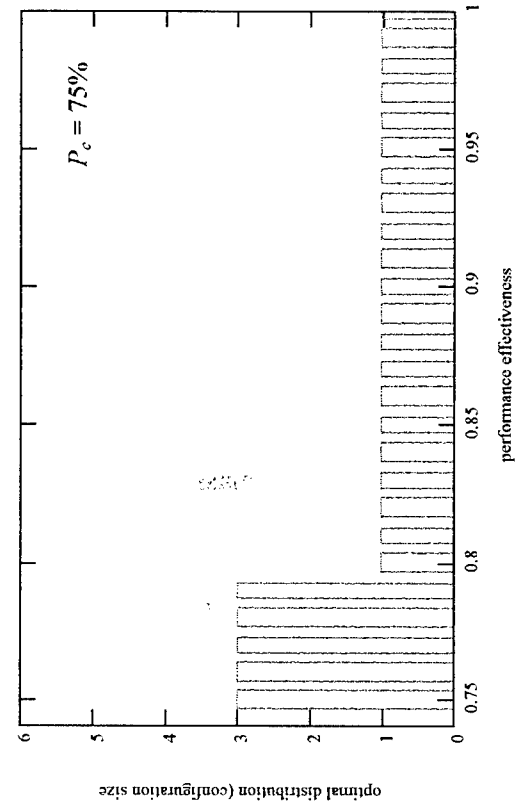
C: *Optimal distribution for 80% configuration probability of detection.*

D: *Optimal distribution for 75% configuration probability of detection.*

Figure 8.5: Optimal configuration size for other configuration probability of detection mission specifications.

Increasing the learning discount curve also drives the optimum point towards higher levels of distribution. For all of the results presented, we employed a 15% learning curve discount rate based on dialogue with Lockheed Martin [27]. *Figure 8.4* shows the cost metric as a function of other learning curve values. As the learning curve increases the mass production cost discount increases which is one of the benefits of a distributed architecture. The learning curve discount rate should increase as orbital altitude decreases since the constellation size grows and discount rates generally increase with the number of production units. Other factors that may increase future discount rates are satellite mass production techniques and automated testing and check-out, Section 3.7.

8.3.4 Cost as a function of frequency

Since the aperture mass density is a function of frequency and is a significant cost driver, system cost is a function of frequency. The relationship is given in *Figure 8.6* assumes a phased array aperture. As frequency increases, the spacing between transceiver elements of the phased array must decrease to avoid grating lobes (see Section 3.3). Thus, increasing the frequency increases the total number of transceiver modules for a given aperture area.

A higher order relationship between frequency and aperture that is not included in the model that gives *Figure 8.6* is that smaller aperture areas may be realizable with higher frequency systems. Although the total received power flux is proportional to aperture area, the received power for a discrete phased array depends on the number of elements and not on the aperture area. Thus, *Figure 8.6* probably overstates the cost at higher
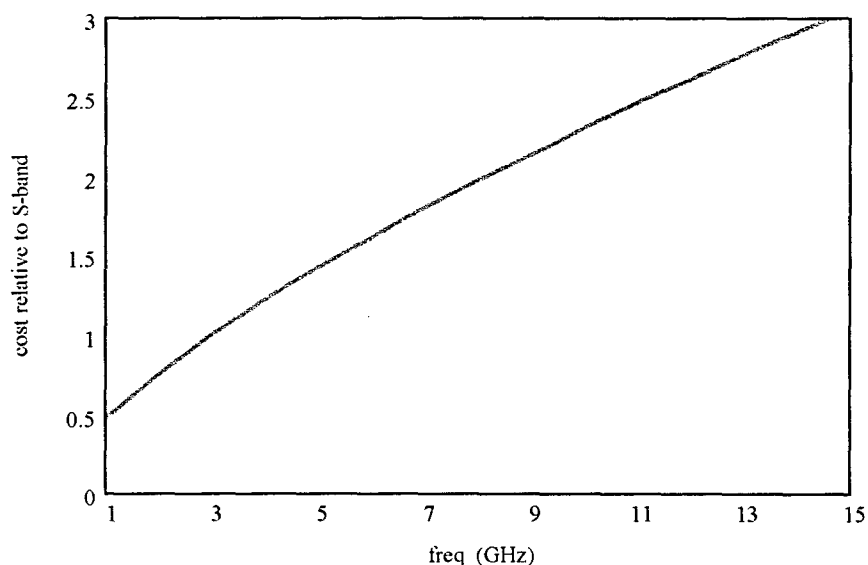
Figure 8.6: Cost as a function of frequency. The increase in cost for higher frequencies may be overstated - see text.

frequencies. Note that power flux is not the only factor that drives aperture. Improved resolution, clutter cancellation, and minimum detectable velocity also required increased aperture area and are not affected by the discrete nature of a phased array. Some of these factors will also benefit from higher frequencies however. The model also does not account for savings that could be realized at higher frequencies by subarray partitioning. Creating a sparse array from filled subarrays is more efficient at higher frequencies, though there is also a corresponding decrease in effective aperture area for a sparse array.

Any or all of these factors could be incorporated into the model and examined in more detail if desired. This again underscores the power and flexibility and of the architecture design process developed here. Nevertheless, *Figure 8.6* can be used as an upper bound for the cost vs. frequency relationship. More detailed analysis, can and should be done, however, to give a more accurate relationship.

Search performance is not a function of frequency. However, performance in clutter improves as frequency increases. Cost also increases with frequency, though the actual relationship is uncertain for the reasons given above. Thus, high frequencies are desired for clutter performance and low frequencies are desired for cost considerations. For clutter analysis, we design SSL 1 and SSL 2 at S-band (3 Ghz) which is the lowest frequency which gives reasonable clutter performance.

8.4 Search Mission Architecture Cost Sensitivity

Another useful application of the architecture design analysis process is the sensitivity of the cost metric to various parameters. The goal of course is too determine which subsystem designs or parameters may offer the potential for significant technological breakthroughs as they affect the total system cost. Analysis in this vein could then be used to determine where to best allocate and partition development resources. Unfortunately, cost sensitivity analysis alone cannot answer this as sensitivity reflects only the derivatives of the cost metric and not where breakthroughs may be likely to occur. Of course, if likely breakthroughs were known, then there would be no need for sensitivity analysis except to predict the overall impact of the breakthrough.

Despite the obvious limitation of sensitivity analysis in recommending how to spend development resources, many potential developments and breakthroughs can be reflected through the various subsystem parameters, particularly the mass density relations. Mass is the ultimate premium in space systems and most technological developments have mass implications. For example, many power subsystem components are currently characterized by low efficiency. More efficient components would results in decreased power subsystem requirements which is reflected by decreased power subsystem mass. A similar argument applies to aperture mass (reductions in transceiver module and support structure mass) and power-aperture payload mass fractions.
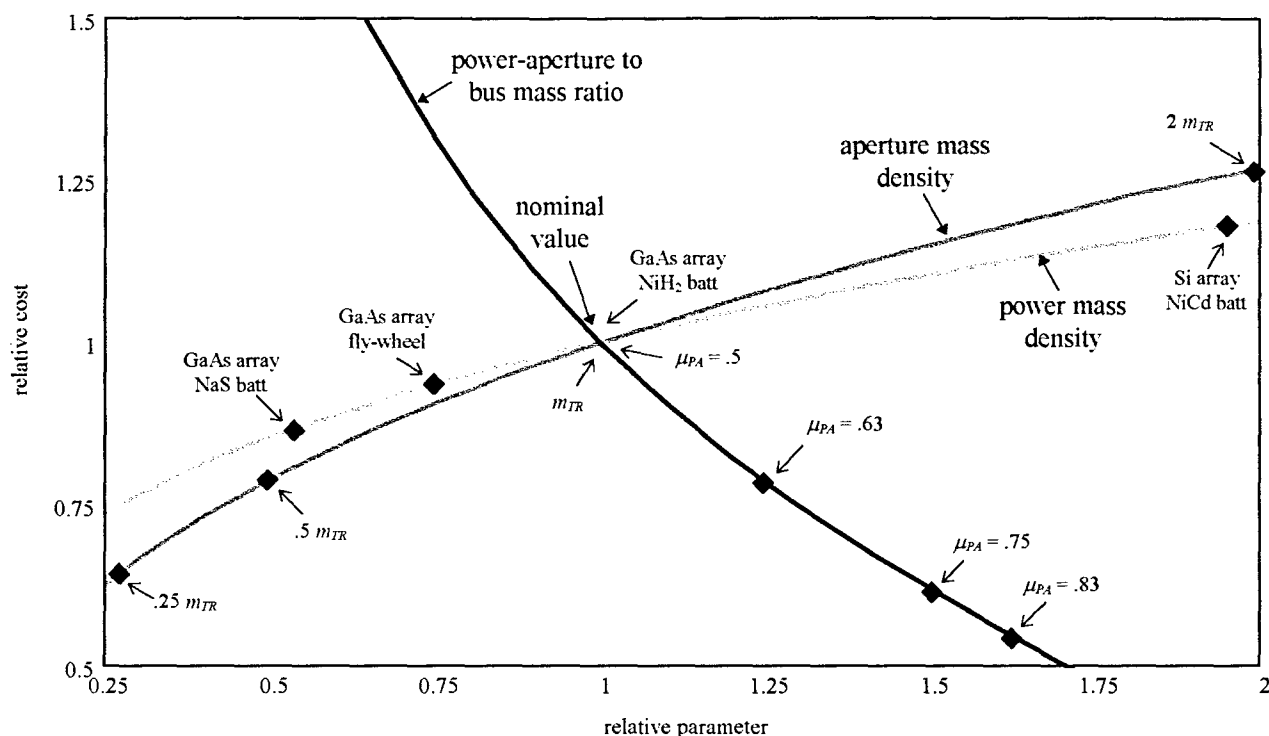
154

Figure 8.7: Cost sensitivity. The ordinate is the relative change in the power, aperture, or payload mass parameter. Points along the curves reflect what technologies represent those relative changes in the parameter value. The abscissa is the relative change in cost.

The cost sensitivity analysis process is illustrated in *Figure 8.7*. The relative change in cost is plotted as a function of the relative change in the mass parameter values from the nominal conditions.

1. System cost is most sensitive to changes in the payload mass fraction. Increasing the payload to bus mass ratio can substantially reduce the system cost. Several possibilities for reducing bus mass include component miniaturization, composite structures, integrated wiring, and integrated structure.

2. System cost is significantly sensitive to the aperture mass density. Reductions in the aperture mass density mass be achievable by reducing the transceiver module mass or by reducing the transceiver structural support mass. Unrelated technologies such as structurally embedded actuators for active surface control may reduce the structure mass necessary to stabilize the array.

3. System cost is only marginally sensitive to the power mass density. Even significant improvements in power storage - such as NaS batteries - can only reduce the system cost by approximately 10%. This is small compared to the improvements likely in other subsystems.

155

8.5 Cost as a function of performance - mean time to detection

For almost every system, system cost increases as performance increases. (Assuming no other differences - e.g. different architectures - between the systems being compared.) This is because the increase in performance usually requires increased system capability, capacity, or coverage rate which in return require more or improved system components and assets. The relationship between cost and performance can generally be categorized into one of three regimes:

1. the marginal increase in performance matches the marginal increase in performance,

2. the marginal increase in performance is much greater than the marginal increase in cost,

3. the marginal increase in performance is much less than the marginal increase in cost.

A marginal increase should be interpreted as a small increment of the variable. The ratio of the margins of change is then the slop of the cost vs. performance curve. For the first regime, the cost vs. performance slope is on the order of one. the slope is much less than one (but greater than zero) for the second regime and much greater than one in the
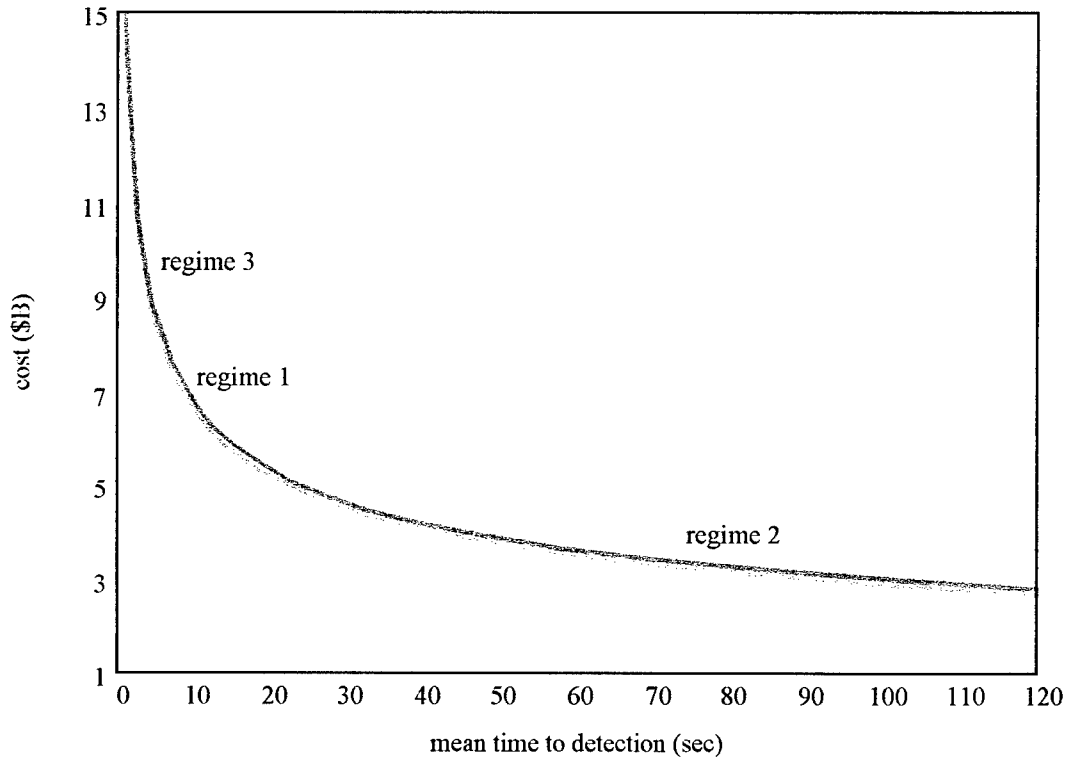

Figure 8.8: Cost as a function of mean time to detection.

156

third regime.

If there are no minimum performance requirements, then the best level of performance falls within the first regime. Many commercial systems operate where the marginal rates of performance and cost are equal since it is here that revenue is maximized. In contrast, most military systems are often driven by performance specifications that strictly define a minimum acceptable performance. Sometimes, incentives are awarded for exceeding the minimum performance specifications by a certain margin. Because of the unique nature of the military mission, combat system requirements should not be determined on the basis of a cost per performance tradeoff. However, it is still important to determine and examine the marginal return of performance relationship if only for the purpose of justifying the high cost associated with high performance. This is the requirements justification step in the system design process.

For SBR, the performance metric is mean time to detection (Section 5.3). The cost as a function of performance curve very clearly shows the three regimes of the cost per performance relationship (*Figure 8.8*). Although assigning boundaries to the regimes is subjective, regime 1 is approximately from 10 to 20 sec mean time to detection, regime 2 is greater than 30 sec, and regime 3 is less than 10 sec mean time to detection. The initial SMC SBR concepts had mean time to detection specifications of 3 to 4 sec which partly accounts for the high cost predicted during SMC's Space Sensors Study [39]. The 10 sec mean time to detection specification used in this analysis lies at the region of transition from regime 1 to regime 3. It would be hard to justify mission specifications with detection capability of much less than 10 sec mean time to detection.

8.6 Performance in Clutter

Clutter performance is a measure of the ability of the SBR system to detect targets in a clutter background. The most common metric of clutter performance is minimum detectable velocity. The cost model was not developed with respect to clutter constraints because mean time to detection is a more important metric and clutter is only a performance limit. Cost can only be optimized with respect one performance metric and mean time to detection was used as that metric. Nevertheless, after the system architecture optimization process is complete, clutter performance is checked to ensure reasonable system performance. Clutter performance is presented and analyzed in this section.

Minimum detectable velocity (MDV) as a function of two of the system architecture optimization variables is given in *Figure 8.9*. The values given in *Figure 8.9* are for absolute MDV which accounts for the cosine loss of the targets relative velocity to the radar. MDV is typically specified by relative velocity, which is always less than or equal to the target's true or absolute velocity. Absolute velocity is used here in order to highlight one of the benefits of distribution: a distributed radar system covers the target

from several different perspectives which increases the probability of a favorable viewing geometry. The relationship for expected target azimuth as a function of distribution developed in Section 7.3.1 is used to convert relative MDV to absolute MDV. Since the values given in *Figure 8.9* are absolute velocity, the clutter performance of SSL 1 is very reasonable for the air target detection mission. Almost all targets of interest fly somewhere from several hundred knots to a well over a thousand knots.

There are two significant drivers of the minimum detectable velocity as a function of altitude relationship:

1. MDV increases at lower altitudes due to the increased orbital velocity and the subsequent increase in azimuth and range cell clutter spread.

2. MDV increases at higher altitudes because of the increased footprint size and the subsequent increase in range cell clutter spread.

These two drivers have opposite effects on MDV and result in a minimum MDV at moderate altitude orbits. The minimum in *Figure 8.9* corresponds to a 3 hr period orbit. The difference in minimum and maximum MDV in absolute terms is only about 40 knots. Thus, although there is a MDV dependance on altitude, altitude does not make an overly significant difference in performance.

There are also two significant effects of distribution on minimum detectable velocity:
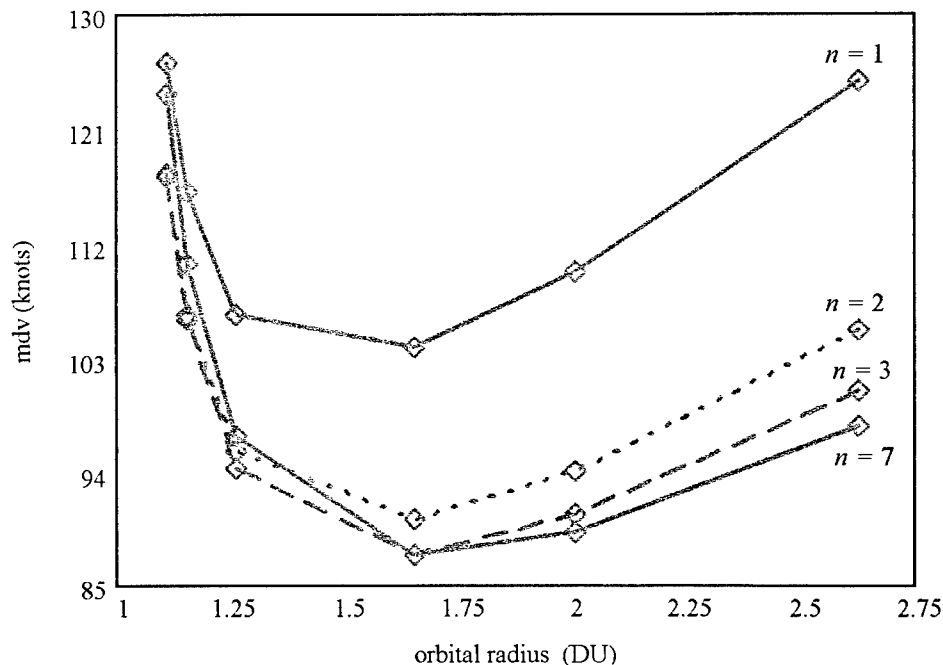
Figure 8.9: Minimum detectable absolute velocity of SSL 1 (operating at S-band) as a function of orbital radius and distribution.

158

1. MDV decreases as distribution increases because of the increased probability for a favorable coverage perspective.

2. MDV increases as distribution increases since the reduced aperture increases the footprint size which subsequently increases the range cell clutter spread.

In the clutter performance as a function of distribution, the first effect is stronger and dominates the relationship. The second effect is somewhat noticeable in *Figure 8.9* as the level curve for a distribution of seven steeply increases at LEO altitudes. As with the clutter altitude performance relationship, although there is an MDV dependance, it is not significant enough to make a difference in the overall performance of SSL 1.

Since a pulse train waveform was employed to improve both range and frequency resolution (Section 2.3 and Section 7.2), the effect of blind velocity should be examined as an additional measure of clutter performance. (Note that the waveform was designed to be unambiguous through the range of the footprint so that range ambiguity is not a problem.) Blind velocity as a function of orbital radius and distribution is given in *Figure 8.10*. The first blind velocity decreases with increasing altitude because the length of the radar footprint increases. The requires a decreased pulse repetition frequency (*prf*) to keep the entire footprint unambiguous in range. Decreasing *prf* decreases the first blind velocity. The first blind velocity decreases with increasing distribution because the reduced aperture area also results in an increased radar footprint length. The clutter frequency spectrum is aliased at integer multiples of the first blind velocity. A higher first
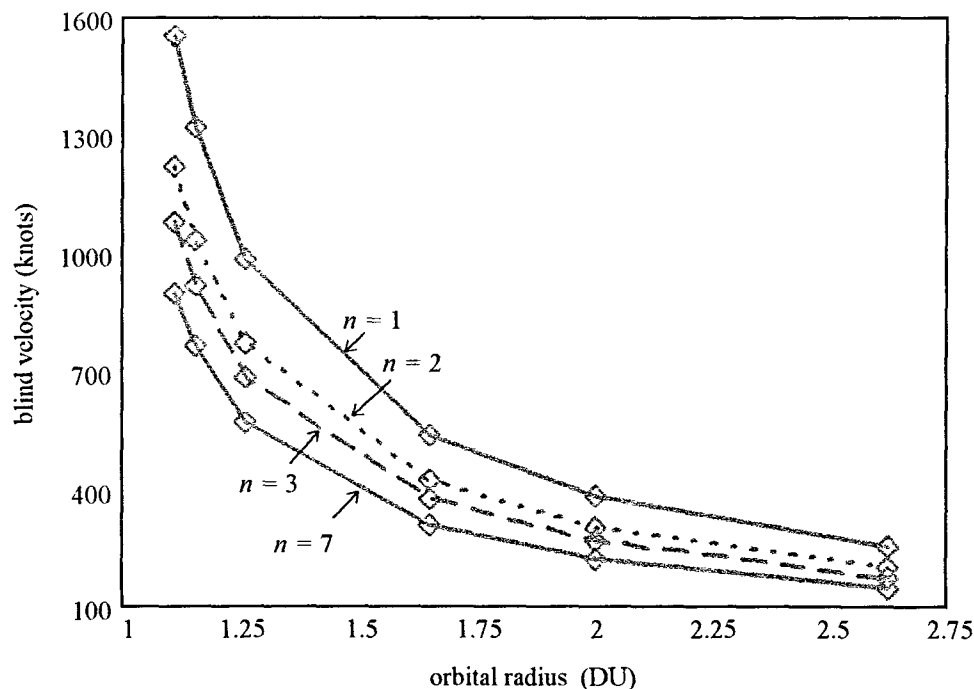

Figure 8.10: The first blind velocity as a function of orbital radius for SSL 1 (operating at S-band).
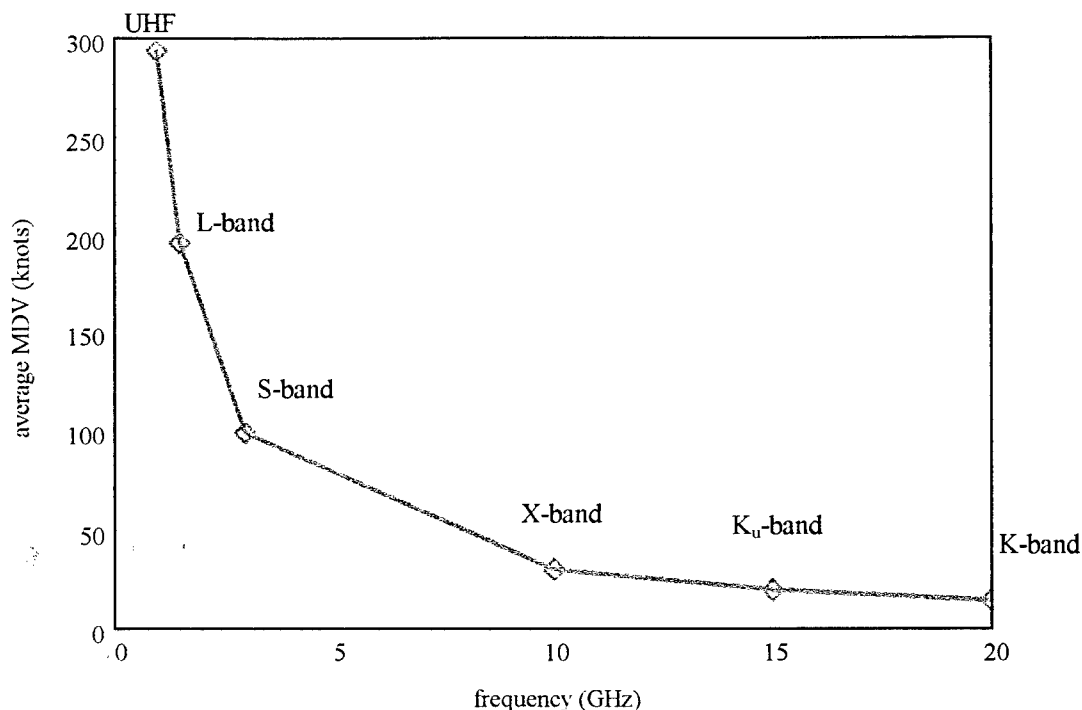
Figure 8.11: Absolute minimum detectable velocity as a function of frequency averaged over beam azimuth angles for a target relative beam azimuth of 45 deg.

blind velocity is a mark of better performance since the frequency aliased clutter spectrum creates Doppler shift holes in which targets are not separable from the clutter.

From *Figure 8.10*, the SSL 1clutter performance due to blind velocity is greatly diminished for middle orbital altitudes compared to LEO systems. In practice, blind velocities can be accounted for by using multiple prf's for the transmitted waveform. However, this still results in an increased detection time for targets with Doppler shifts that fall within the aliased clutter spectrum. Blind velocities are not an insuperable performance limitation, but they do indicate better performance for LEO systems as measured by one clutter performance metric.

The absolute MDV decreases with increasing frequency because the smaller beam patterns associated with higher frequencies results in a narrower clutter spectrum. The relationship is given in *Figure 8.11*.

9. Conclusion

The primary research objective was to assess the feasibility of performing the next generation space surveillance mission from a space based radar (SBR) platform. Since the AWACS surveillance mission has the most challenging update rates of all possible surveillance missions, we particularly sought the feasibility of a non-colossal SBR platform for AWACS. Through a system architecture design and optimization process, we develop an SBR concept for the AWACS mission. The concept employs satellite function distribution to reduce the size of the satellites, increase system reliability, improve system performance, and reduce system cost. The required satellites are well within currently available technology and capability. Although the system acquisition cost is still nontrivial - on the order of 5-10 billion dollars - the concept optimized by the system architecture design process results in up to a 50% reduction in cost from similar concept proposals. The two optimum system concept architectures are summarized in *Table 9.1*.

The distributed operations concept and architecture optimization are the primary reasons for the reduced system cost of SSL 1 and SSL 2 compared to other SBR concepts. Distribution is achieved by distributing the function of aircraft detection from one radar to several radars. This is done by decreasing the probability of detection for a single radar. The overall probability of detection for all the distributed satellites however remains high. Decreasing the probability of detection on a single satellite results in a decrease in the power and aperture required on that satellite. If some level of performance degradation is allowed for the distributed system, then the reliability of the distributed system is higher than a system of only one radar. The level of performance degradation and the reliability cost savings can be quantified by the distributed reliability model developed in Chapter 4. The distributed radar system also has the advantage of an increased number of perspectives of the target area which results in improved target detection and clutter rejection capability. Different perspectives of the same target also improves the spatial and frequency resolution of the target. Distribution also improves constellation coverage by decreasing the variability of coverage over time for a given latitude. In a similar manner,

Table 9.1: Optimal system concept architectures.

	SSL 1	SSL 2
power (kW), P*	6.4	8.2
aperture (m²), A*	50	64
dry mass (kg)	2000	2600
altitude (km)	700	700
distribution	5	3
constellation size, N	176	107
IOC cost (B)	$7.27	$7.35

distribution decreases the variability of target slant range which results in a more efficient design with less excess capability.

The benefits of distribution are:

1. improved reliability/survivability
2. slightly smaller constellations for the same minimum coverage
3. improved design efficiency through decreased range variability
4. decreased power-aperture per individual satellite
5. increased savings due to production learning curves
6. increased target perspectives and detection capability

The second primary reason for the reduced system cost of SSL 1 and SSL 2 relative to similar SBR proposals is the optimized system architecture. A system cost model was developed from a set of subsystem cost estimating relationship models, the distributed reliability model, a constellation coverage model, and an operations model that integrates system performance. The system architecture design process gives optimum results for the four primary system drivers:

1. Power
2. Aperture
3. Distribution
4. Orbital altitude

These are the most important architecture variables as they drive the system cost and all other architecture and system values may be derived from them. Optimal power and aperture were analytically determined from the optimality relation derived in Section 3.6. Optimum distribution was semi-analytically derived - both analytic and numerical components to the solution - as a function of system performance requirements, constellation coverage, and reliability issues. The optimum orbital altitude was numerically determined by iterating over the cost function.

There are several additional conclusions that support many of the secondary research objectives:

1. Performance distribution - distribution of the detection function - can reduce the cost and system scale required to perform the AWACS surveillance mission from a space based radar platform. (Chapter 5, Section 8.3.3)

2. The system architecture design and optimization process is a powerful tool for concept development. The cost model required for the optimization process is useful for comparing the cost and performance of different concepts on the same basis. (Section 8.3)

3. Analytic optimality conditions, e.g. the power aperture optimality relation, are convenient rules of thumb for sizing system concepts. (Section 3.6)

4. Mean time to detection is a useful performance metric for quantifying AWACS

surveillance capability. (Section 5.3)

5. A significant savings in system cost can be realized through an increased payload to bus mass ratio or a decreased transceiver module and support structure mass. Cost reductions due to improved power subsystems are less significant, including improved solar arrays and power storage systems. This tends to imply that reductions in cost are more likely from improvements in bus and aperture technologies. Research and technology development funds may want to be concentrated in these areas. (Section 8.4)

Search performance is not a function of frequency. However, performance in clutter improves as frequency increases. Cost also increases with frequency, though the actual relationship is uncertain. Thus, high frequencies are desired for clutter performance and low frequencies are desired for cost considerations. For clutter performance analysis, we design SSL 1 and SSL 2 at S-band (3 Ghz) which is a reasonably low frequency which gives reasonable clutter performance.

The system architecture design process developed here is a powerful systems engineering tool and should be useful in the optimization and analysis of a wide variety of space systems. The process has led to a distributed Space Based Radar concept that may someday be the primary surveillance component of the next generation Airborne Warning and Control System.

References

[1]   USAF Fact Sheet 96-13, "E-3 Sentry (AWACS)," available from Air Combat Command Public Affairs office, 115 Thompson St, Ste 211; Langley AFB, VA 23665-1987 or on the internet at http://www.af.mil/news/factsheets/E_3_Sentry__AWACS.html

[2]   Canavan, Gregory, and Edward Teller, "Strategic Defense for the 1990s," *Nature*, vol. 344, no. 6268, pp. 699-704, 19 Apr 1990.

[3]   Canavan, Gregory, and Edward Teller, "Low-Level Satellite Expand Distributed Remote Sensing," *Signal*, Aug 1991, pp 99-103.

[4]   Heimiller, R.C., J. E. Belyea, and P. G. Tomlinson, "Distributed Array Radar", *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-19, no. 6, Nov 1983, pp 831-839.

[5]   Steinberg, Bernard D., and E. Yadin, "Distributed Airborne Array Concepts," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-18, no. 2, Mar 1982, pp 219-226.

[6]   NASA briefing, "Distributed Array Radar (DAR)," NASA Study, Feb 1981.

[7]   DeLap, Ronald, and Scott Suhr, "Transition of Airborne Surveillance/Reconnaissance to Space", presented at the 1996 *AIAA Space Technologies Conference*, Huntsville, AL.

[8]   Cohen, Richard E., *The Physics Quick Reference Guide*, Woodbury, New York, American Institute of Physics Press, 1996.

[9]   Personal course notes, Sonar, Radar, and Seismic Signal Processing I, MIT 6.455, Fall, 1996.

[10]   Kirkwood, James R., *An Introduction to Analysis*, PWS-KENT Publishing Co., Boston, 1989, p 158.

[11]   Siebert, William M., *Circuits, Signals, and Systems*, The MIT Press, Cambridge, MA, 1986.

[12]   Van Trees, H. L., *Detection, Estimation, and Modulation Theory*, Wiley, New York, 1971.

[13]   Edde, Byron, *Radar: Principles, Technology, Applications*, Upper Saddle River, NJ, Prentice Hall, 1993.

[14]   Kong, Jin Au, *Electromagnetic Wave Theory*, 2nd ed., New York, John Wiley & Sons, Inc., 1990.

[15]   Skolnik, Merrill I., *Introduction to Radar Systems*, 2nd ed., New York, McGraw Hill, Inc., 1980.

[16]   Cantafio, Leopold J., ed. *SBR Handbook*, Norwood, MA, Artech House Inc., 1989.

[17]   SBR constellation study done at Phillips Lab, available via the internet at http://www.plk.af.mil/SUCCESS/sbrstdy.html

[18]   Knepp, Dennis L. and J. Todd Reinking, "Ionospheric Environment and Effects on Space-Based Radar Detection," *SBR Handbook* (ed. by L. J. Cantafio), Norwood, MA, Artech House Inc., 1989.

[19] *Space Division Unmanned Spacecraft Cost Model.* 5th ed. El Segundo, CA: USAF Space Division, Directorate of Cost Analysis.

[20] *Space Division Unmanned Spacecraft Cost Model.* 6th ed. El Segundo, CA: USAF Space Division, Directorate of Cost Analysis.

[21] Wong, Robert, "Cost Modeling, " *Space Mission Analysis and Design* (ed. W.J. Larson and J.R. Wertz), Ch 20: 715-740, Microcosm, Inc., 1992.

[22] Wright, T.P. "Factors Affecting the Cost of Airplanes," *Journal of Aeronautical Sciences. 3: 122-128.* 1936.

[23] McDermott, Joseph, K. "Power," *Space Mission Analysis and Design* (ed. W.J. Larson and J.R. Wertz), Ch 11.4: 391-408, Microcosm, Inc., 1992.

[24] Lee, C.C., J. Sasonoff, and R. Naster, "Transceiver (T/R) Module Technology for Space-Based Radar," *SBR Handbook* (ed. by L. J. Cantafio), Norwood, MA, Artech House Inc., 1989.

[25] National Space Science Data Center models: http://nssdc.gsfc.nasa.gov/space/model/models_home.html

[26] Seltzer, S.M. SHIELDOSE: A Computer Coc̄ ̄c̄ ̣spac̄ -̄5nielding Radiation Dose Calculations, National Bureau of Standards, NBS Technical Note 1116, US Government Printing Office, Washington, DC, 1980. (The code is available via anonymous FTP from the NSSDC [19].)

[27] McQuiddy, Chuck, briefing and dialogue with Daniel Hastings, Lockheed Martin, Sunnyvale, CA, 4 Oct 96.

[28] Ropelewski, Robert, "Satellite Services Soar," *Aerospace America,* Nov 1996, 26-31.

[29] Personal course notes, Real Time Systems, MIT 16.840, Spring 1997.

[30] Hecht, H. "Figure of Merit for Fault-Tolerant Space Computers," *IEEE Transactions on Computers,* C-22, No 3: 246-251, 1973.

[31] Hecht, H. "Reliability During Space Mission Concept Exploration," *Space Mission Analysis and Design* (ed. W.J. Larson and J.R. Wertz), Ch 19.2: 700-713, Microcosm, Inc., 1992.

[32] Rice, S.O., Mathematical Analysis of Random Noise, Bell System Tech. Journal, vol 23, pp 282-332, 1944; vol 24, pp 46-156, 1945.

[33] Swerling, P., Probability of Detection for Fluctuating Targets, IRE Trans, vol IT-6, pp 269-308, Apr 1960.

[34] Swerling, P., Detection of Fluctuating Pulsed Signals in the Presence of Noise, IRE Trans. vol IT-3, pp 175-178, Sep, 1957.

[35] Schwartz, M., Effects of Signal Fluctuation on the Detection of Pulse Signals in Noise, IRE Trans., vol IT-2, pp66-71, Jun 1956.

[36] Scholtz, RA, et. al., The Detection of Moderately Fluctuating Rayleigh Targets, IEEE Trans, vol AES-12, pp 117-125, Mar 1976.

[37] Fadok, David S., John Boyd, and John Warden, *Air Power's Quest for Strategic Paralysis,* Maxwell AFB, Ala.: Air University Press, February 1995

[38] Whitaker, Randall, *GLOSSARY: The Convoluted Terminology of Information Warfare*, 1997, available via the WWW:

[39] Preston, Robert A, Col. (USAF, commander SMC/XR), personal briefing and dialogue, SMC/XR, LA AFB, 7 Nov 1996.

[40] Walker, J.G., Some Circular Orbit Patterns Providing Continuous Whole Earth Coverage. *Journal of the British Interplanetary Society*. vol 24, 369-384, 1971.

[41] Walker, J.G., Continuous Whole-Earth Coverage by Circular-Orbit Satellite Patterns, Royal Aircraft Establishment Technical Report No. 77044, 1977.

[42] Walker, J.G., Satellite Constellations, *Journal of the British Interplanetary Society*. vol 37, 559-572, 1984.

[43] Draim, J. Three and Four Satellite Continuous Coverage Constellations. *Journal of Guidance, Control and Dynamics*. vol 6, 725-730, 1985.

[44] Draim, J. A Common-Period Four Satellite Continuous Global Coverage Constellation. *Journal of Guidance, Control and Dynamics. vol 10, 492-499, 1987.*

[45] Draim, J. A Six Satellite Continuous Global Double Coverage Constellation. AAS/AIAA Astrodynamics Conference, 1987.

[46] Rider, L., "Nadir Hole-Fill be Adjacent Satellites in a Single Orbit," Technical Note, *J. Astro. Sci.*, vol. 28, no. 3, July - Sep 1980, pp 299-305.

[47] Adams, WS, and L Rider, "Circular Polar Constellation Providing Continuous Single or Multiple coverage Above a Specified Latitude" *J. Astro. Sci.*, vol. 35, no. 2, April-June 1987, pp 15-192.

[48] Rider, L, "Optimized Polar Orbit Constellation for Redundant Earth Coverage," *J. Astro. Sci.*, vol. 33, no.2, April-June 1985, pp. 147-161.

[49] Vitto, Vince, personal briefing and dialogue, MIT Lincoln Labs, Bedford, MA, 26 Jan 1996 and 17 May 1996.

[50] Blackledge, Ronald G. (Space based surveillance program manager), personal briefing and dialogue, Phillips Lab, Kirtland AFB, NM, 25 Mar 1997.

[51] Blackledge, Ronald G. (Space based surveillance program manager), "Spear Concept Briefing," Phillips Lab, Kirtland AFB, NM, 25 Mar 1997.

[52] Preston, Robert A, Col. (commander SMC/XR), "Space Sensors Study," briefing to Gen Viccellio (commander AFMC), 3 Nov 1995. available on the Internet at http://www/afbmd.laafb.af.mil/xrt/sensor/index.html

[53] Larson, Wiley J., and James R. Wertz, ed., *Space Mission Analysis and Design*, Boston, Microcosm, Inc., 1992.