JOURNAL OF SPACE MISSION ARCHITECTURE

JSMA

Issue 3 Fall 2003

A PUBLICATION OF THE JET PROPULSION LABORATORY'S CENTER FOR SPACE MISSION ARCHITECTURE AND DESIGN, A NASA CENTER OF EXCELLENCE

NOTICE TO AUTHORS

The *Journal of Space Mission Architecture* is published by the Jet Propulsion Laboratory's Center for Space Mission Architecture and Design. The journal's goal is to provide a medium for the presentation of space mission architecture concepts, processes, and tools. Authors are invited to submit articles on all areas of space mission architecture, such as case studies of the development of missions including new or unusual architectural challenges, non-NASA mission architectures, the mission architecture process and tools, historical or biographical notes, and space program architecture.

For instructions on the submission of articles, authors are referred to

http://www.jpl.nasa.gov/csmad/journal/

All correspondence should be addressed to the Editor-in-Chief:

DONNA M. WOLFF
Jet Propulsion Laboratory
301-230
4800 Oak Grove Drive
Pasadena CA 91109

This publication was prepared by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

National Aeronautics
and Space Administration

Center for Space Mission Architecture and Design
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

JPL 03-15

Foreword

The word “architecture” in the title of this journal implies a systems point of view, and one of creation, not just analysis. “Mission” is there to keep the papers focused on topics that provide real value to the space community. Space mission architecture cannot be done entirely on paper (or computer), but a significant fraction can be, a fact that allows a journal like this to be closer to the action than might otherwise be the case. This journal is intended to be a forum where both the role of architect and the product, architecture, can be displayed, discussed, and debated openly, limited only by the quality controls of peer review and readership demands.

As always, each issue covers a wide variety of topics. This issue’s historical piece provides examples of current technologies that were predicted by early science fiction writers and extrapolates to possible future technologies based on modern-day science fiction. Technology is also a subject of two other papers: one looking at the architecture of the program to select for flight validation technologies for future missions and a second describing the process for making appropriate technology investments to maximize science return under a constrained budget. Another paper tackles the problems encountered when considering the architecture of a commercial human space venture. The final paper details the steps taken by a team of students and faculty in developing a space mission architecture around an existing instrument.

Donna Wolff
September 2003

NASA's New Millennium Program: Flight Validation of Advanced Technologies for Space Applications

Charles P. Minning and David Crisp

Abstract

NASA's New Millennium Program (NMP) was created to accelerate the insertion of advanced spacecraft and instrument technologies into future science missions by validating these technologies on deep space and Earth-orbiting missions. This paper describes the currently approved NMP flight projects, the technology validation results obtained to date, and briefly describes the processes used to select and validate their associated technologies. Future NMP flight opportunities are also discussed.

I. INTRODUCTION. In 1995 the National Aeronautics and Space Administration (NASA) created the New Millennium Program (NMP). The objective of the NMP is to conduct spaceflight validation of breakthrough technologies that could significantly benefit future space- and Earth-science missions. The breakthrough technologies selected for validation must (1) enable new science capabilities to fulfill NASA's Space and Earth Science Enterprise objectives and/or (2) reduce the costs of future space and Earth science missions. The goal of spaceflight validation of these technologies is to mitigate the risks to the first users and to promote the rapid infusion of these technologies into future science missions. A secondary objective is to return high-priority science data to the extent possible within mission and cost constraints. The Jet Propulsion Laboratory (JPL) was assigned to manage the program for NASA. Additional information on the New Millennium Program is available on the Internet [1].

The first-generation NMP missions include Deep Space 1 (DS1), Deep Space 2 (DS2), and Earth Observing 1 (EO1). These missions were designed to provide a comprehensive, system-level validation of suites of interacting, high-priority spacecraft and measurement technologies.

The second-generation NMP missions include Space Technology 5 (ST5) and Earth Observing 3 (EO3). These missions also focus on system-level validations, but they make greater use of partnerships, and their technologies were selected through a revised process. While the NMP plans to continue flying these system-level technology validation missions where appropriate, this approach is being augmented with more highly focused, component-level validation flights of breakthrough technology subsystems. Brief descriptions of the first- and second-generation NMP flights are given below. We then describe the system and subsystem validation objectives for future NMP flight validation opportunities. Finally, we review the processes used to select technologies for validation on NMP missions.

II. First-Generation Validation Flights.

Deep Space 1 (DS1)

DS1, the first of the New Millennium missions, was launched from the Kennedy Space Center on 24 October 1998. This spacecraft, depicted in Figure 1, carries a complement of 12 technologies that were validated during the ten months following launch. These technologies are

1. ion propulsion system (IPS) with a suite of diagnostic sensors,
2. solar concentrator arrays,
3. autonomous optical navigation (Autonav),
4. miniature integrated camera and spectrometer (MICAS),
5. plasma experiment for planetary exploration (PEPE),
6. small deep space transponder (SDST),

Figure 1. Deep Space 1 contains 12 technologies for spaceflight validation. The spacecraft intercepted Asteroid 1996 Braille in July 1999, and the technology validation mission was completed the following September. Deep Space 1 is now a science mission with the objective of intercepting Comet Borrelly in 2001.

7. Ka-band solid-state power amplifier (KAPA),
8. beacon monitor operations,
9. autonomous remote agent experiment,
10. silicon-on-insulator low-power electronics experiment,
11. multifunctional structure, and
12. power actuation and switching module (PASM).

These technologies and the DS1 mission are described in more detail in [2] and [3]. Detailed validation reports for each of these technologies are available on the World Wide Web [4].

The Ion Propulsion System: The IPS offers significant mass savings for future space missions with high ΔV requirements. The IPS uses xenon as the propellant, and at peak operating power consumes 2.3 kW and produces 92 mN of thrust at a specific impulse of 3100 s. Throttling is achieved by balancing thruster and propellant feed parameters at lower power levels. At the lowest thrust level, 20 mN, the power consumption is 0.5 kW at a specific impulse of 1900 s. The diagnostic sensors were included to quantify the interactions of the IPS with the spacecraft and science instruments, to validate models of those interactions.

Once in space the IPS got off to a shaky start, shutting down automatically after only 4.5 minutes of operation. This shutdown was attributed to a short-circuit caused by a piece of conductive debris trapped between the ion engine's closely spaced (0.6 mm) ion acceleration grids. To dislodge the debris, the grids were thermally cycled, causing them to move relative to each other. After this process was repeated several times, the engine started normally. Since then, it has worked flawlessly throughout the validation flight and well into the extended mission. At the time of this writing, the IPS had logged more than 230 days of operation in space, far longer than any other space propulsion system.

Solar Concentrator Arrays: Because ion propulsion systems require large amounts of electric power, a high-power solar array was required to validate the IPS. The solar array technology adapted for DS1 was the Solar Concentrator Array with Refractive Linear Element Technology (SCARLET), which was sponsored by the Ballistic Missile Defense Organization (BMDO). This solar array uses cylindrical silicone Fresnel lenses to concentrate sunlight onto 3600 dual-junction $GaInP_2/GaAs/Ge$ solar cells arranged in strips. The SCARLET array includes two wings, each of which consists of four (113 cm × 160 cm) panels that are folded for launch. When fully extended, the wings measure 11.8 m from tip to tip. It was designed to produce 2.5 kW at 1 AU.

SCARLET was the first concentrator array used for primary power on a spacecraft. This technology was extensively tested on DS1, validating the performance of the multijunction cells, the Fresnel optics and their innovative deployment approach, and the compatibility of their electrical design with the IPS.

Autonomous Optical Navigation: The Autonav system has piloted the spacecraft from shortly after separation from the launch vehicle through the encounter with Asteroid Braille and is currently being used for navigation to the planned encounter with Comet Borrelly in September 2001 as part of the extended mission. Autonav uses data stored in the flight computer as well as data acquired and processed during the mission. The stored data consists of the spacecraft trajectory (generated and optimized on the ground), the ephemerides of the target bodies, about 250 "beacon" asteroids, and all planets (except Pluto), as well as the positions of about 250,000 stars. During the mission, once or twice each week, the spacecraft is turned to point the MICAS sequentially at 4 to 20 "beacons." Visible images from the MICAS are processed and combined with other information to determine the location of the spacecraft.

Autonav worked flawlessly during most of the validation flight, but miss-targeted the images scheduled for closest approach to Asteroid Braille during the flyby on 29 July 2000. This tracking problem apparently resulted from the MICAS camera's inability to reacquire this dim object after the spacecraft recovered from a safing event that occurred a few hours before the encounter. The Autonav software has since been updated to address these and other challenges faced during the validation flight, producing a much more robust product for future deep space missions.

The Miniature Integrated Camera and Spectrometer: MICAS is an advanced 12-kg instrument that includes two visible imaging channels, an ultraviolet (UV) imaging spectrometer, and a short-wave infrared (SWIR) imaging spectrometer. All sensors share a common 10-cm-diameter telescope. This instrument contains no moving parts, and the structure and optics are fabricated from thermally stable silicon carbide.

The two MICAS visible imaging channels and the SWIR imaging spectrometer were successfully validated in flight, but the quality of their data was seriously compromised by scattered light in the instrument. This scattered light was attributed to a poorly designed solar calibration port and sun shade. The UV channel could not be validated because the detector (a frame-transfer charge-coupled device [CCD]) failed early in the mission.

The Plasma Instrument for Planetary Exploration: PEPE combines several plasma physics instruments in one compact 5.6-kg package to determine 3-dimensional plasma distribution over its 4π steradian field of view. PEPE also provides information about the plasma environment associated with the IPS and its interactions with spacecraft surfaces and instruments and with the solar wind.

PEPE data taken in the vicinity of Earth was validated directly through comparisons with measurements from plasma instruments on other spacecraft (Advanced Composition Explorer [ACE], Wind, and Cassini). PEPE measurements also confirmed that high-quality plasma measurements could be obtained at energies greater than 50 eV while the IPS is operating. Below this energy, PEPE also measured xenon ions and secondary electrons from the IPS and SCARLET arrays.

Small Deep Space Transponder: Three telecommunications technologies were included on DS1 for validation. The SDST combined the receiver, command detector, telemetry modulator, excitor, beacon tone generator (for beacon monitor operations, another technology validated on the mission), and control functions. These capabilities were integrated into one 3-kg package. The SDST allows X-band uplink and both X-band and Ka-band downlink.

All SDST functions for uplink, downlink, and radio ranging were thoroughly validated in flight, including the optional Ka-band downlink capability. These validation activities reduced the risk of this advanced telecommunications technology sufficiently that the SDST has been adopted as the baseline on the Mars ’01 Orbiter and the Space Infrared Telescope Facility (SIRTF).

The Ka-Band Solid State Power Amplifier: The Ka-band (32-GHz) solid-state power amplifier has a potential for providing a 4-fold increase in the data rate when compared to conventional X-band systems. KAPA is the highest-power device of this type ever used for deep space communications. Its key technology is 0.25-μm GaAs Pseudomorphic High Electron Mobility Transistors (PHEMT). KAPA’s mass was 0.66 kg, its RF output power was 2.2 W, and its gain was 36 dB. In flight, KAPA operated nominally, completing 28 power cycles, and accumulated over 1680 hours of operation.

Beacon Monitor Operations Experiment: The SDST generates tones used during beacon monitor operations, an operational concept conceived to reduce the heavy demand expected on the Deep Space Network (DSN) if many missions are flown simultaneously. In this operations concept, an onboard data summarization system determines the overall health of the spacecraft and then transmits one of four tones to indicate to the operations team (on Earth) the urgency of the need for DSN coverage for the spacecraft. Because they lack data modulation, these tones are easily detected with small, low-cost systems, reserving the large, expensive DSN stations for command uplink and data reception when the beacon indicates that such attention is required.

The DS1 flight allowed a complete, end-to-end validation of the Beacon Monitoring Operations Experiment. Validation tests included tone transmission and detection, engineering summary generation and visualization, and tone message handling and reporting, among other capabilities.

The Remote Agent Experiment: The RAX is an onboard artificial intelligence system for planning and executing spacecraft activities.

This technology uses and executes a mission plan expressed as high-level goals. A planning and scheduling engine uses the goals, comprehensive knowledge of the state of the spacecraft, and constraints on spacecraft operations to generate a set of time-based activities that are delivered to the executive. The executive then creates a sequence of commands that are issued directly to the appropriate destinations on the spacecraft. The executive monitors the responses to the commands and reissues or modifies them as required. A mode identification and reconfiguration engine aids in assessing the spacecraft state and in recovering from faults without requiring help from the ground, except in extraordinary cases.

RAX was tested for several days on DS1, in a series of scenarios based on active cruise mode. In these tests, it commanded a subset of the spacecraft subsystems, including the IPS, MICAS, Autonav, attitude control system, and a series of power switches. The goal of these tests was to execute an IPS thrust arc, acquire optical navigation images as requested by the Autonav system, and respond to simulated faults. After a rough start, the RAX satisfied 100% of its flight validation objectives. It won the NASA 1999 Software of the Year Award.

Low-Power Electronics Technologies: The low-power electronics experiment was developed to characterize the effects of the space environment on sub-0.25-μm, fully depleted, silicon-on-insulator complementary metal-oxide semiconductor (CMOS) test devices that operate at supply voltages of less than 2 V. This experiment functioned nominally throughout the DS1 flight.

Multifunctional Structures: The MFS is an experiment to evaluate the concept of folding spacecraft electronics into the walls of the spacecraft, thereby saving weight and space by eliminating chassis, cables, and connectors. The MFS on DS1 was sponsored by the Air Force Research Laboratory Phillips Laboratory (AFRL/PL). It incorporated multichip modules and flex circuit interconnects along with advanced composites and thermal management systems. Once in flight, the MFS experiment was powered up once every two weeks, and two experiment cycles were run during each test. The validation was a complete success.

The Power Actuation and Switching Module: The PASM combines advanced, mixed-signal application-specific integrated circuits (ASICs) and high-density interconnect technologies to enable significant miniaturization of spacecraft electrical load and switching functions by eliminating the bulky relays and fuses that have been used in the past. Each of PASM's four switches could isolate faults, limit in-rush and fault currents, supply voltage and current telemetry, and perform other functions. They could switch from 30 to 40 V at up to 3 A. The PASM switches were successfully exercised several times during the DS1 flight and showed no performance degradation.

Deep Space 2

DS2, the second of the New Millennium missions, was launched from the Kennedy Space Center on 3 January 1999 and arrived at Mars on 3 December 1999. The objective of this mission was to demonstrate (1) key technologies that enable future network science missions that require multiple landers, penetrators, orbiters, or flyby spacecraft; (2) a passive reentry system; (3) highly integrated microelectronics capable of surviving high-g impact and operation at extremely low temperatures; and (4) in situ subsurface data acquisition. The primary science objectives were to determine if water ice is present below the Martian surface and to characterize the thermal properties of the Martian subsurface soil.

This mission consists of two identical, 3-kg microprobes, one of which is shown in Figure 2. They were attached to the cruise stage of the Mars 98 Polar Lander. Approximately 10 minutes prior to landing, the probes were to separate from the cruise stage, descend through the atmosphere without the benefit of either parachutes or airbags, and survive a high-g impact near the northern boundary of the southern Martian polar region. The probes were protected during entry into the Mars atmosphere by an advanced, nonablative heat shield. At impact on the Martian surface, the heat shield was designed to shatter, and the probes were designed to separate into two parts. One part (the aft-body) was to remain on the surface, and the other part (the fore-body) was designed to penetrate approximately one meter into the Martian soil. The fore- and aft-bodies were expected to experience shock loads of about 30,000 g's and 60,000 g's, respectively.

The fore-body included a novel drill mechanism to acquire subsurface samples and place them in a small crucible. The crucible was to then be heated to release water if any were present. A tunable diode laser was included to detect the presence of water vapor in the evolved gases. The fore-body also included temperature sensors to measure the vertical

Figure 2. Deep Space 2 Mars Microprobe. At impact, the aft-body (left) will remain on the Martian surface, and the fore-body (right) will penetrate into the subsurface soil to detect the presence of water. A multilayer flex cable connects the two sections.

temperature gradient in the soil. Data from these instruments was to be transmitted via an advanced multilayer flex cable to a radio beacon in the aft-body. The beacon was to relay the data to the Mars Global Surveyor spacecraft, which, in turn, was to relay the data back to Earth. The aft-body also included the lithium/thionyl chloride primary batteries, which supplied power to the probes.

Microelectronics were to play a key role in the DS2. The microelectronics technologies to be validated were (1) an advanced microcontroller, (2) a power control unit, and (3) the evolved water experiment with its associated electronics. All of these technologies were located in the fore-body. The advanced microcontroller was to control operation of and store data produced by the evolved water experiment and the temperature sensors, then send the data to the radio beacon for transmission to the Mars Global Surveyor. The power control unit was to provide power management, distribution, and voltage conversion for the evolved water experiment, temperature sensors, and the advanced microcontroller. Some of the unique electronic packaging aspects of the electronics in both the fore-body and the aft-body are described in [5].

Contact was never established with the DS2 microprobes after they landed on Mars. The exact cause of this problem has not yet been determined.

Earth Observing 1

EO1, the third of the New Millennium missions, was launched from Vandenberg Air Force Base in November 2000. This validation flight, depicted in Figure 3, includes three advanced imaging instruments and eight advanced spacecraft technologies. The three instruments, the Advanced Land Imager (ALI), the Atmospheric Corrector (AC), and the Hyperion (hyperspectral imager) will enable a new generation of high-performance, low-mass, low-cost instruments for future Landsat-style measurements obtained by NASA's Earth Science Enterprise. The ALI employs novel, wide-angle optics and a highly integrated spectrometer with a panchromatic channel.

ALI flight validation is designed to demonstrate spectral and spatial performance comparable to or better than Landsat 7, with substantial mass, volume, and cost savings. Earth imagery is degraded by atmospheric absorption and scattering. The EO1 Atmospheric Corrector is a compact, low-resolution imaging spectrometer designed to provide the first space-based test of an Atmospheric Corrector for increasing the accuracy of surface reflectance estimates. The Hyperion is a hyperspectral imager capable of resolving 220 spectral bands at wavelengths between 0.4 to 2.5 μm. Its spatial resolution is 30 m over a 100-km swath.

The advanced spacecraft technologies include an X-band phased array antenna, a carbon-carbon composite radiator, a lightweight, flexi-

Figure 3. Earth Observing 1. This spacecraft is validating technologies contributing to the reduction in cost of future Landsat missions.

ble solar array, a pulsed plasma thruster, and enhanced formation flying capability. These technologies will enable smaller spacecraft buses that have lower mass and require less power. A wide-band advanced recorder processor (WARP) receives data from the three instruments at up to 840 Mbits/s, then formats and stores the data in its 40-Gbit solid-state recorder. The WARP includes a lossless data compression chip and a 10-multichannel-interface processor (MIP) capable of processing science data. The data will be sent to the ground via the X-band phased array antenna at 105 Mbits/s and subsequently sent to GSFC for technology validation and science research. Parallel EIA RS-422 interfaces provide the data path between each of the three instruments and the WARP.

To validate the advanced instruments, EO1 flies in formation with Landsat 7, providing at least 200 paired scene comparisons with that satellite's Enhanced Thematic Mapper + (ETM+) instrument.

III. Second-Generation Validation Flights.

Space Technology 5

The ST5 mission will fly three miniature (≈22-kg) spacecraft in a highly elliptical orbit around the Earth. The ST5 Nanosat Constellation Trailblazer Mission is scheduled for launch (as a secondary payload) in 2004. This NMP flight will validate technologies needed for future constellations of spacecraft required for studies of the magnetospheres of the Earth and other planets. ST5 will validate a suite of eight advanced technologies, including

- a formation flying and communications instrument that communicates between spacecraft and determines their positions using the Global Positioning System (GPS),
- autonomous ground station software for scheduling and orbit determination of constellations of spacecraft,

- an X-band transponder that requires ¼ the voltage and half the power, weighs 12 times less, and is nine times smaller than proven technology,
- advanced multifunctional structures that provide electrical interconnects and reduce cable mass,
- an ultra-low-power electronics experiment that uses a field programmable gate array (FPGA) that is more reliable and uses 1/20 the power of proven technology,
- variable emissivity coatings that are electrically tunable, such that their optical properties can be changed to increase absorption of solar IR radiation when the spacecraft is cool or to increase surface emissivity to reject internally generated heat to space to cool the spacecraft,
- a miniature microelectromechanical system (MEMS) chip that provides fine attitude adjustments on the spacecraft using 8.5 times less power and weighing less than half as much as proven technology,
- a lithium-ion power system for small satellites that stores two to four times more energy and has a longer life than proven technology.

This mission will also validate manufacturing methods needed to produce large numbers of spacecraft. Additional information on ST5 can be found in [6].

Earth Observing 3

The EO3 mission will fly the Geosynchronous Imaging Fourier Transform Spectrometer (GIFTS) and six other advanced technologies to enable improved remote sensing of clouds, moisture, and winds in the Earth's atmosphere. These capabilities are needed for improved weather forecasting and to provide additional constraints on atmospheric trace gases. GIFTS will be carried to geosynchronous orbit in late 2004 as a secondary payload on a satellite provided by the US Navy Office of Naval Research. The EO3 mission will provide a system-level validation of seven advanced technologies, including

- a high-spectral-resolution, imaging Fourier transform interferometer,
- high-speed, onboard signal processing,
- advanced cryogenic cooling,
- data compression,
- autonomous pointing and control,
- low-power, radiation-tolerant microelectronics, and
- lightweight structures and optics.

As a by-product of this technology validation flight, GIFTS will return valuable scientific data that will enable the development and validation of improved strategies for monitoring atmospheric temperatures, water vapor content, trace gas amounts, and winds from geostationary orbit. For example, while existing geostationary instruments can provide data needed to infer winds by tracking clouds, the high-resolution, spatially resolved GIFTS spectra should also reveal water vapor variations in clear skies that can be tracked to yield information on winds as well.

IV. Future NMP Flight Opportunities. The first- and second-generation NMP flights described above were designed to provide a comprehensive, system-level validation of suites of interacting technologies. This technology validation approach is essential in some circumstances, but it is not necessarily the most efficient approach for other technologies. For example, the combination of the ion propulsion system, the SCARLET concentrator arrays, and the Autonav system was a particularly expedient approach for validating the DS1 solar electric propulsion system. However, other DS1 technologies, such as the low-power electronics or the multifunctional structures experiment, as well as a broad range of other technologies currently in development, could be successfully validated as individual components or subsystems on a broad range of platforms.

These considerations suggest that it would be possible to accelerate the rate of technology infusion into future missions by augmenting NMP's existing system-level validation flights with a low-cost, quick-turnaround "subsystem mode" that would include stand-alone validations of a range of payloads, from components to complete subsystems. These flights would focus specifically on technologies that

- require a validation in space to mitigate risks to first science users (e.g., environmental effects, incorporate a major implementation shift, etc.),
- enable critical measurements or spacecraft capabilities,
- yield broad benefits to multiple users, and
- can be tested as stand-alone components without extensive interactions with other payload elements.

By focusing on the specific components of an advanced spacecraft subsystem or instrument that requires a flight validation, this approach should

- enhance the validation rate by allowing components to be flown on the first available flight, thus precluding the need to wait for the development of a range of other technologies, and
- be more cost effective, because it minimizes the investment in low-tech components or technologies that do not need to be validated in space.

To achieve the greatest benefit from this approach, the NMP is currently working with other NASA programs and with other government agencies to identify flights of opportunity that could be exploited for component-level flight validations. The program is also studying the feasibility of a general-purpose technology validation carrier, or space truck that could be used to validate technologies for NASA's Space and

Earth Science programs, as well as technologies contributed by our partners from other government agencies.

In spite of its potential advantages, this subsystem mode cannot satisfy all of NASA's needs for technology validation. The NMP therefore plans to continue to conduct system-level validation flights. These flights are of particular value for testing advanced technologies that represent a system-level paradigm shift in implementation or operations approach or measurement concept. For example, a system-level flight might be needed to validate the use of a coordinated network of spacecraft, rather than a single platform to make a particular measurement (e.g., Magnetospheric Constellation, Terrestrial Planet Finder, or Mars surface weather or seismic networks). Also, a system-level validation of an advanced instrument might be needed to minimize the risk and ensure the continuity of a critical measurement (e.g., Landsat, operational weather satellites).

To address these needs, and to ensure the highest possible rate of technology infusion within the current budget, the NMP is sharpening its criteria for technology validations, to yield a balanced mix of subsystem and system-level validation flights. The selection process for the first subsystem validation flight for the NASA Office of Space Science is currently under way, and will constitute Space Technology 6 (ST6). In the future, we anticipate that technologies for subsystem validation flights will be solicited about once a year. System-level flights will be conducted at intervals of 18 months to two years.

V. Technology Selection Processes for NMP Validation Flights.

Integrated Product Development Teams and Technology Selection for First-Generation NMP Missions

For the first three and a half years of the NMP, technology selection for flight validation was focused in six technology thrust areas: Autonomy, Telecommunications, Modular and Multifunctional Systems, Microelectronics, In Situ Instrument and Microelectromechanical Systems, and Instrument Technologies and Architectures. For each thrust area, teams consisting of representatives from government, academia, federally funded research and development centers, and industry were formed. These teams, referred to as Integrated Product Development Teams (IPDTs), operated as consortia to identify breakthrough technologies, prepare technology roadmaps, and develop flight hardware and software to validate these new enabling technologies in a cooperative and collaborative fashion. Non-NASA members offered specific technologies of interest to the NMP and were selected through a formal source selection process. The organizational membership of these IPDTs is described in more detail in [7].

The IPDTs proposed technologies to be incorporated into the first generation of deep space (DS1 and DS2) and Earth-observing (EO1) validation flights described above. The objective was to validate funded technologies early enough in the NMP schedule to mitigate their cost and schedule risks to the flights. The proposed technologies were evaluated for their potential benefit as well as their impact on cost, schedule, and overall risk at the end of the concept development phase for each project. The selected technologies were then incorporated into the baseline architectures for these three flight projects. Those high-risk technologies that encountered unforeseen development problems during project implementation were deleted from the project to reduce cost and schedule risk.

For those technologies included in the final hardware configuration of a flight project, technology validation agreements were negotiated between the technology providers and the flight project office. These agreements defined the success criteria and quantitative performance goals to be achieved to validate a technology successfully. In addition, data obtained from these technologies were to be analyzed and disseminated to interested organizations/parties by means of appropriate workshops, NMP technology validation symposia, formal technology validation reports, and peer-reviewed journal papers.

VI. Technology Selection for Second-Generation Missions and for Future NMP Flight Opportunities.

After the establishment of the New Millennium Program in 1995, the NASA Strategic Plan [8] was published. This plan defines the Agency vision, mission, and fundamental questions of science and research that are the foundation of Agency goals to be accomplished over the 25 years spanning 1998 to 2023. This plan also describes the four Strategic Enterprises that manage programs and activities to implement the Agency mission. The Strategic Enterprises are Space Science, Earth Science, Human Exploration and Development of Space, and Aeronautics and Space Transportation Technology. These enterprises have published their respective strategic plans that include comprehensive science and focused technology roadmaps for proposed future missions.

NASA also created the Cross-Enterprise Technology Development Program (CETDP) to support technology development for multiple-Enterprise customers. Typically, CETDP acts to develop critical space technologies that enable innovative and less costly missions and enable new mission opportunities through revolutionary, long-term, high-risk, high-payoff technology advances. Many of these technologies are at the very early stages of development and may be viewed as technologies of opportunity (“technology push”) rather than as required technologies identified in the Enterprise-focused technology roadmaps.

The NASA Strategic Enterprises and the CETDP are now responsible for developing technology roadmaps that were previously a key function of the NMP IPDTs. In addition, the technology acquisition process for future NMP flight projects was simplified by using mission-specific technology solicitations. As a result, the IPDTs have been disbanded. NMP has subsequently developed a new process for selecting technologies for spaceflight validation and for formulating technology validation missions that will support the goals of the Space Science and Earth Science Enterprises [7].

Flight Validation Domain

The number of systems, subsystems, or components that might be flight-validated is very large. The reasons for flight validation range from cannot be tested on the ground to lack of flight heritage due to an advance in the technology or to procedural change in hardware assembly or mission operations. Thus, a rational and equitable selection process is required to allow an orderly and open selection of technologies for flight validation on NMP missions.

As depicted in Figure 4, the technology selection process begins with aligning emerging technologies being developed by NASA, other government agencies, universities, and industry with the science capability needs of the Space and Earth Science Enterprises. Emphasis is placed on identification of emerging high-risk, high-payoff breakthrough technologies. Using flight validation justification factors, the candidate breakthrough technologies for flight validation are identified [7]. Due to resource limitations, NMP can flight-validate only a small portion of the candidate technologies.

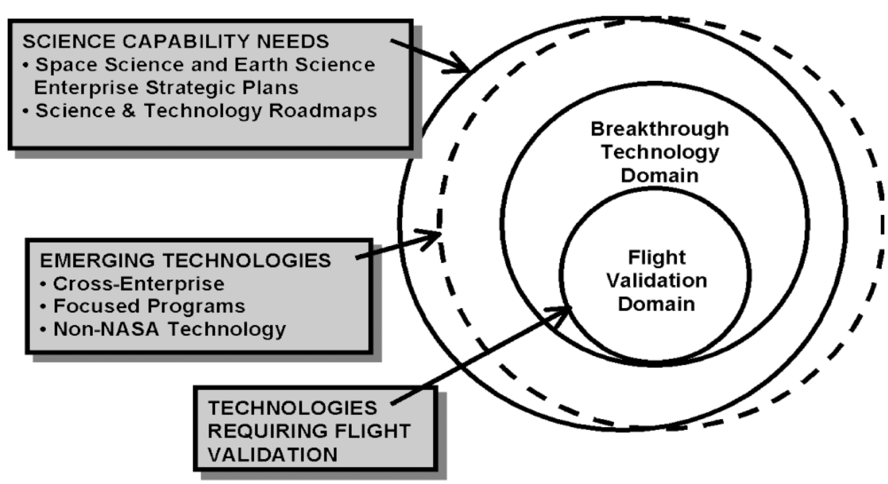


Figure 4. The relationship between the technology development domain and the identification of candidate technologies for spaceflight validation on NASA NMP missions

Technology Selection Process

The NMP process for planning and implementing technology validation flights is summarized in Figure 5. The process consists of four major activities: (1) a preproject planning activity for identifying and capturing candidate concepts, (2) establishing teams to study candidate concepts, (3) studying the concepts in detail, and (4) selecting one concept for continuation into project formulation, implementation, flight, and dissemination of flight test results.

The process for identifying flight validation technologies and assimilating them into candidate flight validation missions is initiated by the NASA Enterprise Theme technologists, who review the technology and capability needs identified in the Strategic Enterprise (science and technology pull) roadmaps, compile a capability needs inventory for each theme.

In parallel, the NMP staff compiles a list of candidate technologies for flight validation from information in the NASA Technology Inventory. This compilation step is constrained and guided by several factors, including (1) breakthrough nature of the technology, (2) its breadth of applications, (3) flight validation justification factors, (4) risk identification, and (5) Technology Readiness Level (TRL). Technology breakthroughs are defined in terms of performance and cost, as compared to the state of the art. The breadth of future applications is determined from the support shown by the Enterprise Theme technologists. The risk identification factors are customer-focused and are meant to determine the degree to which the technology will be utilized. The maturity of the technology is indicated by the TRL (Table 1). The justification factors are a key requirement in the technology selection process. These factors (environmental, paradigm shift, and interdependency and/or complexity) are discussed in detail [7]. The capability needs inventory compiled by the Theme Technologists is then combined with the list of candidate technologies compiled by the NMP staff. The results are then assimilated into a list of candidate flight validation concepts with the concurrence of the Theme Technologists, the CETDP thrust area managers, and the NMP staff. The list of candidate flight validation concepts is also

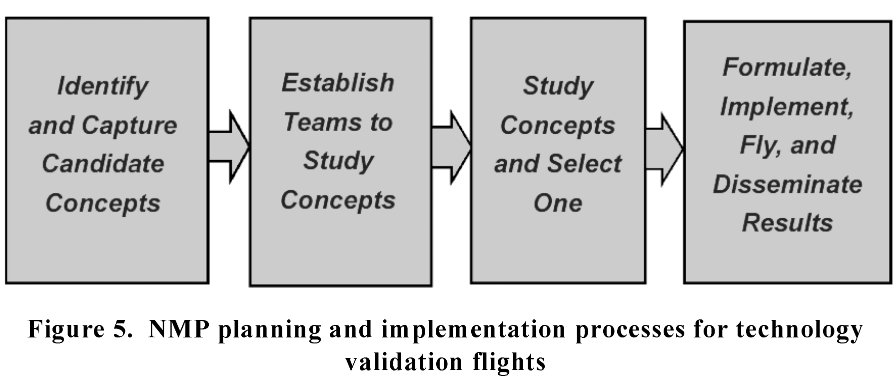


Figure 5. NMP planning and implementation processes for technology validation flights

Table 1: Technology Readiness Levels

Level	Description
TRL 9: Actual system mission-proven through successful mission operations (ground or space)	Fully integrated with operational hardware/software systems. Actual system has been thoroughly demonstrated and tested in its operational environment. All documentation completed. Successful operational experience. Sustaining engineering support in place.
TRL 8: Actual system completed and mission-qualified through test and demonstration in an operational environment (ground or space)	End of system development. Fully integrated with operational hardware and software systems. Most user documentation, training documentation, and maintenance documentation completed. All functionality tested in simulated and operational scenarios. Verification and validation completed.
TRL 7: System prototype demonstrated in operational environment (ground or space)	System prototype demonstration in operational environment. System is at or near scale of the operational system, with most functions available for demonstration and test. Well integrated with collateral and ancillary systems. Limited documentation available.
TRL 6: System, subsystem model or prototype demonstrated in a relevant end-to-end environment	Prototype implementations on full-scale realistic problems. Partially integrated with existing systems. Limited documentation available. Engineering feasibility fully demonstrated in actual system application.
TRL 5: System, subsystem, component validated in relevant environment	Thorough testing of prototype in representative environment. Basic technology elements integrated with reasonably realistic supporting elements. Prototype implementations conform to target environment and interfaces.
TRL 4: Component, subsystem validated in laboratory environment	Stand-alone prototype implementation and test. Integration of technology elements. Experiments with full-scale problems or data sets.
TRL 3: Analytical and experimental critical function and/or characteristic proof-of-concept validated	Proof-of-concept validation. Active R&D is initiated with analytical and laboratory studies. Demonstration of technical feasibility using breadboard/brassboard implementations that are exercised with representative data.
TRL 2: Technology concept and/or application formulated	Applied research. Theory and scientific principles are focused on specific application area to define the concept. Characteristics of the application are described. Analytical tools are developed for simulation or analysis of the application.
TRL 1: Basic principles observed and reported	Transition from scientific research to applied research. Essential characteristics and behaviors of systems and architectures. Descriptive tools are mathematical formulations or algorithms.

made available to the non-NASA technical community for informal comment and feedback on relevant technology developments taking place outside NASA.

Flight Project Formulation and Implementation

The candidate flight validation mission concepts are further refined using feedback received from the non-NASA technology community and programmatic priorities and constraints established by NASA Headquarters. Several of these concepts are then selected and a report describing them is prepared by the NMP staff. This report outlines the approach for each proposed mission, the technologies bundled in each concept, and the risk reduction approach for each concept. It is submitted to NASA Headquarters for review. Two or more of these concepts are then selected for the project formulation phase.

The NMP staff then uses a competitive solicitation process to form concept study teams. Membership in these study teams is open to US industry and academia, NASA centers, other US government agencies, nonprofit organizations, and Federally Funded Research and Development Centers (FFRDCs). These organizations are encouraged to propose technologies that meet the needs of the mission concepts described in the technology announcement. The proposed technologies should be at TRL 3 or 4 and have a realistic plan to reach level 7 in time to support launch of the mission. The proposals are peer-reviewed, and NASA Headquarters makes recommendations for membership on the concept study teams. Formal membership selection is made by the NMP. NASA Headquarters also assigns leadership responsibility to a NASA center for each of the concept study teams.

Each of the concept study teams work to refine their respective concepts and develop a detailed concept proposal. During this study phase it may be found that all of the technology validation goals cannot be achieved due to either funding or technology readiness constraints. Thus it is possible that some of the technologies selected will not be included in the final concept proposal. The suppliers of those technologies that are included in the final concept proposal will be funded to supply the flight articles if the concept is selected for detailed project formulation.

Once a flight validation concept is selected, a solicitation for a spacecraft bus provider will be conducted if this is required for the mission. A detailed project plan is prepared. This plan includes detailed schedules, cost estimates, a technology validation plan including technology validation agreements with the technology suppliers, a technology infusion plan, and a risk management plan. At this point, if there is sufficient justification, science instruments may be included in the mission. The science instruments will be acquired through the standard NASA AO (Announcement of Opportunity) process. These plans are submitted to NASA Headquarters for approval, and implementation of detailed design, fabrication, and software development activities take place. If science measurements are included in the mission, the science team is selected through the NASA AO process. After the mission is completed, the technology validation results are disseminated via workshops, NMP symposia, technology validation reports, and journal papers.

VII. SUMMARY. Technology validation for future NASA science missions is a complex process that requires careful planning, coordination, and execution. NASA created the New Millennium Program in 1995 to perform the technology validation needs for the NASA Office of Space Science and Office of Earth Science. The first- and second-generation NMP missions and their associated suites of technologies and technology validation results to date have been summarized. The scope of NMP validation flights has been increased to include more frequent validation flights for high-risk, high-payoff subsystem and component technologies. We reviewed the processes for selecting technologies for NMP validation flights.

Acknowledgments

The work described in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration and at the National Aeronautics and Space Administration's Goddard Spaceflight Center and Langley Research Center.

References

[1] NMP Web page (http://nmp.jpl.nasa.gov)

[2] Rayman, M. D. and Lehman, D. H., "Deep Space One: NASA's First Deep-Space Technology Validation Mission," *Acta Astronautica* 41, pp. 289–299, 1997.

[3] Rayman, M. D., Varghese, P., Lehman, D. H. and Livesay, L., "Results from the Deep Space 1 Technology Validation Mission," *50th International Astronautical Conference Proceedings,* 1999.

[4] "Deep Space 1 Technology Validation Reports," JPL Publication 00-10, October 2000, available at http://eis.jpl.nasa.gov/nmp_techval_rep.

[5] Arakaki, G. and D'Agostino, S., "New Millennium DS-2 Electronic Packaging: Smaller, Faster with 'Managed' Risk," *Proc. IEEE Aerospace Conference,* vol. 1, pp. 385, 1998.

[6] Ticker, R. L., "The New Millennium Space Technology 5 (ST5) Project," presented at Space 2000 Conference, Albuquerque, NM, 28 February–2 March 2000.

[7] Minning, C. P., Buehler, M. G., Fujita, T., Lansing, F., Man, G., Aljabri, A. and Stevens, C., "Technology Selection and Validation: New Millennium Flight Projects," *Proc. IEEE Aerospace Conference,* vol. 7, pp. 583, 2000.

[8] *NASA Strategic Plan,* NASA Policy Directive (NPD)-1000.1, 1998.

Authors

Dr. Charles P. Minning is a microelectronics technologist and previously served as co-lead of the Microelectronics Integrated Product Development Team for NASA's New Millennium Program. Prior to joining JPL in 1997 he worked for 25 years at the Hughes Aircraft Company, where he served in both line management and project management positions. He has broad experience in the fields of thermal management and electronic packaging for commercial communications satellites and NASA spacecraft, airborne radar signal processors, and infrared sensor electronics for spaceborne and tactical weapons systems. He has authored over 20 technical papers, holds four patents and has one patent application pending. He received his BS, MS, and PhD in Mechanical Engineering from the University of California at Berkeley. He also received an ME in the Engineering Executive Program from the University of California at Los Angeles.

Dr. David Crisp is a senior research scientist in the Earth and Space Sciences Division at JPL and the Chief Scientist of NASA's New Millennium Program. He received his PhD in Geophysical Fluid Dynamics from Princeton University in 1984. There he specialized in atmospheric physics and studied the thermal balance of the middle atmosphere of Venus. Dr. Crisp has been an instrument supplier and science team member on several missions. These missions included the Soviet/French/US Venus Vega Balloon Mission, the Hubble Space Telescope Wide Field and Planetary Science 2 Project, and the Mars Pathfinder ASI/Met Science Team.

Management Architecture: Problems Facing Lunar-Based Entrepreneurial Ventures

Mike H. Ryan and Michael R. Luthy

Abstract

Managers in remote locations have always faced numerous complexities while attempting to conduct business. Firms wishing to operate on the moon will encounter not only the typical problems associated with doing business at a distance but also some relatively unique ones. Managers of commercial operations, compared to military commanders, have less formal power than they might prefer and therefore are more dependent on leadership skills for successful outcomes. Between the time of a project's start-up phase, with relatively few employees, and commencement of large-scale operations, involving an entire community, lunar-based managers will find themselves responsible for a broad range of activities not encountered in their business experiences on Earth. Acknowledging the need for an expanded set of leadership and operational skills, in addition to those traditionally expected of managers, is a requisite condition for successful lunar-based ventures.

I. INTRODUCTION. Business plans for developing successful lunar enterprises rest on two fundamental assumptions: (1) being able to reach to the moon with sufficient numbers of people and materials to sustain the operation, and (2) being able to work there over some minimum duration. The first is constrained by issues related to profitability, more so than by those related to technology. Humans have had the technology to reach the moon and operate in its environment for more than 30 years. The problem, however, has been that there are few proposed ventures that would generate sufficient revenue to attract the needed capital from nongovernmental sources—both to build the outbound and return transportation system and to support construction of lunar-based facilities.

The second assumption, working on the moon, is clearly a function of the first, that is, getting there. Previous discussions related to the viability of lunar enterprises focus on issues related to getting there, virtually to the exclusion of how operations will be conducted upon arrival. The implicit bias in these discussions is that lunar-based operations would not be significantly different from their terrestrial-based counterparts. By extension, the problems that good lunar managers will face are likely to be very similar to those they might face in a field office back on Earth. As experienced managers of complex projects or businesses know, however, making general assumptions that assume success transference across undertakings with diverse circumstances is highly dangerous.

The evolution of management theory generally points to a progression from more restrictive forms, or as some have termed, "industrial feudalism" (for example, feudal societies of Europe, the Hershey experiment, certain robber barons, Carnegie), with a capitalist or other individual having nearly total social control of a community, to less restrictive forms. It has been speculated that lunar organizations would start out more in the management style of early industrial-age organizations, with strong central control, and then evolve towards the more freewheeling organizations that are common on Earth today. This is linked to an initial situation where there would be central control of transport and vital resources such as power, water, and air. Unlike earlier periods however, today there is an evermore prevalent march toward individualism and the rights of the individual that come into conflict with such a freewheeling approach. The patterns of the past likely will not be applicable in the future due to the relatively unique needs and demands of lunar communities and operations. The initial group of individuals chosen to work in a lunar environment will undoubtedly be selected in part because they would be responsive to a clear chain of command and control. This might give the appearance of creating a lunar operating environment with "feudal" characteristics. It is highly unlikely, however, that the United States or Western Europe would permit, much less tolerate, any of the less desirable corporate forms common to the 19th and 20th centuries. The appropriateness of specific models can be debated and should be studied further. Lunar management practices, however, may be better framed by recognizing that even if the historical precedents are considered, the current pace of change and/or evolution common to modern business would likely come into play. There is reason to hope that the positive lessons of running large technologically-based organizations might be applied to lunar operations.

While there are many issues common to all types of businesses, successful lunar-based business activities in some instances will require adaptation of current business theories and practices, while in others the development of new managerial responses to unique problems will be required. Advance identification of areas in which common Earth-based business practices would prove unsatisfactory (or downright dangerous) might make the difference between a given venture being successful or the subject of an expose on the *60 Minutes* television program.

In the isolated, inherently dangerous environment of the moon there is an extreme downside risk to poor decisions. To date, all space missions with crews (whether American or Russian) have utilized a vertical, military command structure that has generally been satisfactory in terms of operation and completion of mission objectives. Given the process through which the majority of American astronauts and their foreign counterparts were selected, this makes sense. When combined with the lengthy and homogeneous programs used for astronaut training and the relatively well-defined goals for missions, this system has served fairly well.

There have been occasional noteworthy conflicts, including disagreements between distant ground-based commanders and onboard flight personnel (for example, Apollo, Skylab, Mir, and the Space Shuttle). There are arguments favoring some rigidity in command structures for safety and operational reasons. The best cites the inadvisability of polling the passengers in situations where the pilot has to make critical decisions. Alternatively, while there are business ventures that operate with command structures not unlike those of the military, that is, offshore drilling platforms, saturation diving operations, and demolition activities, they are the exceptions.

Yet even within these operations there are some interesting insights for future space-based activities. For example, on offshore oil platforms, meal service is available 24 hours a day. This is done not only to accommodate the rigorous 24-hour schedule prevalent in that working environment but also because it provides the employees a measure of control over an important element of their personal environment (Stuster 1996). Food is also an important element in keeping saturation divers, and other individuals working prolonged hours in remote locations, moderately content. The pragmatic lesson is that the more restrictions placed on individuals, even for good cause, the more critical it can become for them to control some aspect(s) of their environment, no matter how small.

Most business ventures operate with far less formality and significantly less structure than what is currently present in space activities. As a more diverse group of people move into space, and as actual business operations are conducted on the moon, problems with this type of rigid command structure is increasingly probable. Centralized command may be acceptable in areas related to safety and/or emergency situations, but it unlikely to be acceptable for day-to-day living or for optimal business performance. Greater employee autonomy, a characteristic common to business, will be necessary if lunar facilities are to attract and keep the best qualified personnel (Vanscoy 2000).The range of potential management problems likely to be encountered is directly related to the organizational structure imposed and the nature of commercial skills needed for lunar-based business operations as well as the number of employees actually involved in commercial operations. As a facility moves away from its likely scientific/military origins toward a true commercial enterprise, one can expect an increasing number of potential friction points to emerge. Over time, it might be expected that the perceived habitability of a now commercialized facility could became a serious issue unless the obvious friction areas created by different organizational structure expectations were addressed. Some of the more obvious possible points of conflict appear in Table 1 and Table 2.

Table 1: Military v. Civilian Command Structures

Military or Military-like Org Structures	Civilian or Civilian-like Org Structures
Often rigid rules as to • What will be done • When it will be done • How it will be done	Often flexible rules derived by • Consensus • Priority shifts allowed • Individual decides how task to be done
Force of authority frequently derived from the position itself	Participative

Table 2: Authority Structures: Examples Where Conflict Could Arise Among Personnel

	Military	Scientific	Commercial
Formal authority	Very high	Moderate	Moderate
Value of expertise	Variable	High	Moderate
Perceived flexibility	Very low	Moderate to low	Very high

II. WHAT MANAGERS ARE TRAINED TO DO. For organizations, the primary focus for managers is the implementation, coordination, control, and evaluation of previously developed goals and objectives. It is unlikely that these considerations would be markedly different within any business enterprise operating on the moon. What would be different, however, is that the manager's responsibility for the overall success of the venture would extend to areas not generally viewed as appropriate for their Earth-based counterparts. For example, most firms view scrutiny of employee behavior outside the firm as "off limits" so long as it is not illegal, does not embarrass the firm publicly, and does not translate into reduced performance. This would, by necessity, change in a lunar environment, where the distinction between being "at work" and "not at work" would be irrelevant within the facility.

Safety, operational, and logistics considerations associated with lunar-based businesses would by necessity blur the lines that managers have used in the past to separate work-related issues from non work-related ones. Undoubtedly, the extension of supervision into areas generally regarded as the prerogative of the individual and/or his/her family will be controversial. There are precedents for such actions, but they generally fall into the realm of military or quasi-military organizations in which individuals give up some individual control for the benefit of the group. Rigid command structures govern who will perform a particular task, when the task will be done, and even how it should be accomplished. Civilian command structures are far less rigid, and the power to control every aspect of an activity is less reliable. In a lunar environment, managerial responsibilities could easily evolve into a combination of camp counselor, cop-on-the-beat, facilitator, and old-fashioned schoolmaster. These are not the typical skills one might expect to acquire during the pursuit of an MBA.

Viewing the employee as a whole person, with family, friends, interests, etc. outside the firm is not a uniformly shared attitude among American businesses. The present shortage of qualified employees in many industries has forced some firms to add benefits recognizing their employees' external needs. Elder care, day care, and more liberal and flexible family leave policies are all intended to bolster employee productivity and retain valuable workers. Some firms have gone so far as to establish corporate concierges to assist their employees with tasks once relegated to spouses or other family members. Competing for the best employees means creating an environment in which they are willing to remain. Few employees would regard a rigid hierarchical structure as one in which they would be willing to invest their "valuable" time and energy.

In the global business community there are also cultural variations to this theme that place more emphasis on factors such as family, free time, recreation, and even vacations. Europeans frequently comment about the American approach to business being driven without regard to other factors they view as equally important, such as reasonable time off. Assuming even a minimum level of diversity for a lunar operation, the divergence of opinion as to appropriate expectations could be considerable and problematic. Barring military-operated businesses, there is little chance that employee expectations will lessen in the foreseeable future. As a consequence, a lunar posting might very well require a greater level of benefits with fewer restrictions to offset the relative difficulty of going home, communications, and the perceived deprivations of working on the moon.

III. What They Don't Teach in Business Schools.

Business schools specialize in all manner of business topics. Some topics, such as leadership, involve multiple sets of courses. All business school classes are intended to equip the soon-to-be manager with all the necessary knowledge, if not skills, requisite for managing resources—human, physical, and financial. There are a few areas however, where a lunar-bound MBA might wish for just one more case study, such as what to do when you lose the coercive power that most managers enjoy. Numerous studies have been undertaken examining the challenges involved with Antarctic communities, exploration ventures, and long-duration voyages (Vaernes et al. 1988, Cornelius 1991, Stuster 1996). People who work on the moon will be atypical by the very nature of their activities: the challenges are sufficiently different in degree to make them virtually different in kind. Yet, even though people working on the moon will not be typical, the gulf between them and the average citizen will be less than that between the average citizen and early astronauts.

Operating in a remote location is always a managerial challenge. Operating in a remote location with your employees 24 hours a day, 7 days a week, is a greater challenge. Operating in a very remote, semi-isolated community while managing a group of smart, independent-minded, self-sufficient employees who may tend to resist imposed authority is a challenge that most managers would be ill-equipped to meet. Part of the difficulty involves the inherent differences between leading and managing.

Leading versus Managing

Zaleznik (1977) has argued that there are significant differences between leadership and management. In academic terms, the difference reflects the view that leadership is reserved for individuals who determine the major objectives and the strategic courses of organizations (see Table 3). Leaders are those who introduce major change rather than those who transmit and enforce rules and policies or implement goals

Table 3: Leadership v. Management Comparison Tasks

Leadership	Management
• Promotes ideological values • Motivates and encourages positive self-perception as part of special group • Creates willingness for individual to forgo self-interest	• Promotes rational analytic behavior • Organizes • Coordinates • Implements • Attempts to promote behavior that supports team efforts

and changes initiated by others at higher organizational levels. Leadership appeals to things such as ideological values, motives, and self-perceptions of followers so as to produce effort beyond that expected of their position or willingness to forego self-interest, and to willingly make personal sacrifices in the interest of group vision. Managers use their formal position to apply rational-analytic behavior to organize, to coordinate, and to implement whatever organizational strategies, tactics, or policies have been deemed appropriate. Leadership behavior appeals to follower motives and is interpersonally oriented while manager behavior tends to be rational-analytic and impersonal. Leaders set the direction while managers provide the intellectual content needed for operational efficiency (McAuliffe 1998).

The boundary points between leadership and management are subjects of ongoing debate within academia and organizations. Lamenting the lack of good results a corporate president may exclaim, "where are my leaders?!" Reflecting on the confidence demonstrated by leaders in combat, one might describe a leader as one whom "troops would follow down the barrel of a cannon." In either case, leadership is likely to become a more critical component for successful lunar enterprises than merely good management. Although there is a considerable amount of empirical evidence and theory relevant to the practice of leadership, this knowledge still needs to be adapted to commercial operations in remote locations (Bass 1990, Yukl 1994). Leadership embodies characteristics and skills that when applied adroitly provide a clear, unambiguous, self-supporting vision of what an organization needs to accomplish (or where it should be headed) as opposed to merely directing day-to-day activities (see Figure 1). Viewed somewhat simplistically, good leadership provides the elements critical to motivating higher levels of individual performance. In difficult situations, such as those that might be anticipated

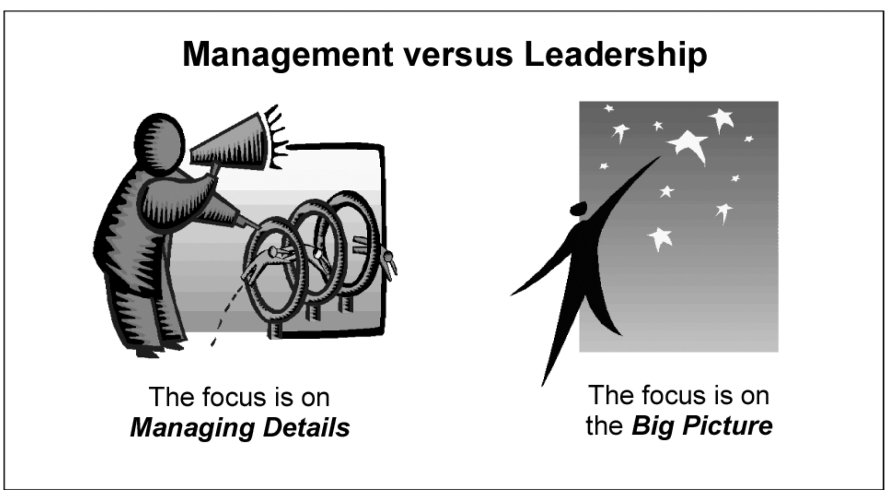


Figure 1. Management v. Leadership Comparison at the Broadest Level

in a geographically isolated and inherently dangerous environment, including a lunar one, leadership is absolutely imperative if people are to perform at their best. It is important to remember, satisfactory completion of management objectives—business or otherwise—is necessary for a successful lunar operation, but will likely prove inadequate without a corresponding vision capable of keeping everyone motivated for the long term.

North American Bias

Perhaps more troubling for the long-term operation of a commercial lunar facility is that most prevailing theories of leadership have a definite North American cultural orientation (see Table 4). Consequently, the primary emphasis is individualistic rather than collective, more concerned with self-interest than duty, oriented more toward rules than norms, emphasizing rationality rather than aesthetics, religion, or superstition, and identifying work as the central focus, imbued with democratic values (Deresky 1994). A substantial body of cross cultural social psychological, sociological, and anthropological research clearly demonstrate that there are numerous cultures that do not share the underlying assumptions of North American based leadership theories (Bowie 1990, Elashhmawi and Harris 1993, Glover 1990). Consequently, while the general challenge of providing quality leadership for lunar operations may be difficult, it might be even more troublesome finding individuals capable of leading a culturally-mixed operation

Table 4: The North American Cultural Divide

North American View of Leadership May Be Problematic for Lunar Operations	
• North American • Individualistic • Self-interested • Governed by rules • Rationality assumed • Democratic focus • Capitalist society	• Other Cultures • Collective orientation • Duty bound • Governed by norms • Aesthetics, religion, and superstition have a place • Not necessarily democratic or capitalist

Regardless of the perspective, a fundamental goal for good management is improved performance, for both the organization and its members. Leadership, with its demands, will be linked to good management and a goal of performance. Performance is the result of a complex interaction of factors. It is vital that lunar-based managers have a clear understanding of employee performance so they can maximize the performance potential of the employees in the facility and avoid the severe downside of mistakes. Virtually everything a manager does influences performance; the selection of personnel and the kind of training they receive can affect aptitude and skill levels; design of the compensation system and the way it is administered can influence motivation levels and overall performance; and the organization and deployment of workers can affect how they perceive their job. One model (see Figure 2) developed as a tool for describing salesperson performance, is illustrative for discussion purposes (Churchill et al. 2000).

Industrial and organizational psychology literature suggests that a worker's job performance is a function of a number of different factors including aptitude, skill level, and motivation (Walker et al. 1977, Brown et al. 1979, Plank and Reid 1994). The histories of both the U.S. and Russian space programs reflect the practice of selecting space crews with very high qualifications in these areas. Entrepreneurial ventures in a lunar environment will, by necessity, have to continue this practice. Because of the changing "familiarity" of space exploration, however, two additional factors from the above cited literature may prove more problematic for managers chosen to lead in a lunar environment, namely role perceptions, and personal, organizational, and environmental variables.

Organizations must present accurate portrayals of the job to recruits, who in turn must possess skills and aptitudes compatible with the needs and offerings of the organization. If these objectives of recruiting and selection are met, personnel socialization is enhanced, and ultimately, performance, satisfaction, job involvement, and commitment are improved (Ingram, LaForge, and Schwepker 1997). Moreover, the role a

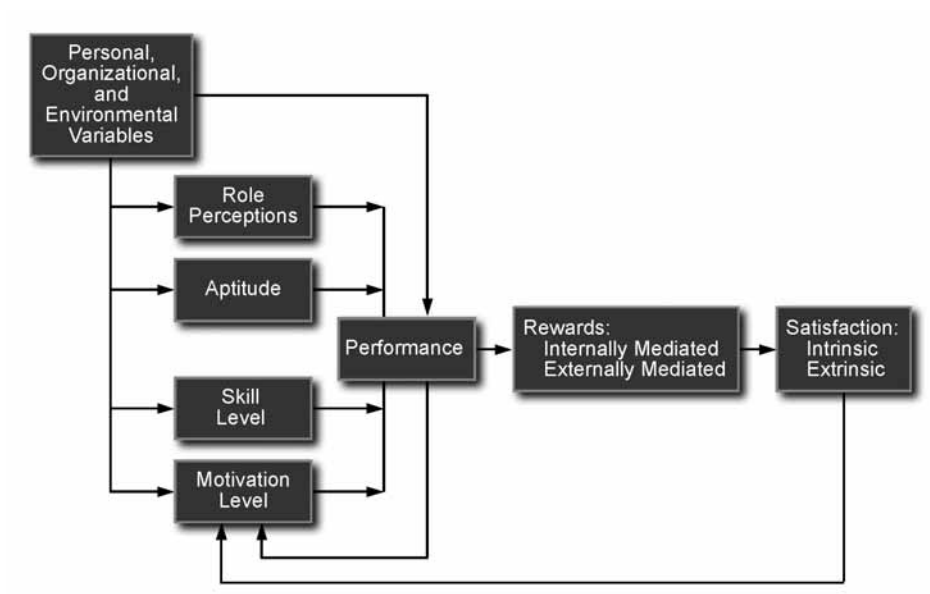


Figure 2. Employee Performance Model

job candidate is expected to perform in a lunar environment represents a set of activities or behaviors. Therefore, miscommunications or misunderstandings concerning roles in a lunar environment, defined largely through the expectations, demands, and pressures communicated to the worker by his or her role partners, may give rise to poor (and potentially disastrous) management outcomes. The worker's perception of these expectations strongly influences the individual's definition of his or her role in the team and behavior on the job (and beyond).

These perceptions can be influenced by a variety of factors including both the nature of the leadership structure and the ability of the on-site management team to address any issue likely to affect individual performance. In a lunar environment, performance is subject to conditions with few parallels on earth. Commercial operations in a hostile environment are not unknown. However, with the exception of offshore drilling platforms, which although numerous are limited in overall scale, general managerial experience with some relatively unique performance issues is quite limited. In a large-scale lunar enterprise with multiple commercial operations involving hundreds of personnel, some issues have the capability of not only degrading performance but also creating an entirely new set of management problems.

IV. UNIQUE PROBLEMS FOR BUSINESS MANAGERS. Business operations at a lunar compound are going to be different just by virtue of their location. Managers selected to lead their firms' lunar activities will face many issues that are relatively uncommon to business. Situations related to hygiene, privacy and personal space, and relationships will take on entirely new dimensions within the closed world of a lunar facility. The fact is that many of these relatively unique problems have been experienced elsewhere (Stuster 1996). The problem for most business school trained managers is that these circumstances still fall well outside their education and experience. Furthermore, they represent areas that (as with managers elsewhere) they probably would have been actively encouraged to avoid. This is because the dilemma facing even an experienced manager is that he or she would be functioning much like a head of a household with all of a family's concomitant social-management problems to deal with. Managers in remote, isolated, technically complex locations such as the moon may find that their managerial model is much like that of a parent of a bunch of very bright children. This is not to suggest that people would be behaving childishly, but that the variety and interaction of potential issues they will face is more complex than just running the operation. Much like a parent, the manager would have to become concerned about his/her kids bathing regularly, playing well with others, and monitoring personal relationships. The analogy is somewhat strained considering that a fully developed management architecture would need to address business and city management, resource, transportation, and health issues as well. The private sector has faced similar issues in a variety of locations throughout the world while building dams, rail lines, factories, and even entire cities. The extension of the full range of these experiences to a lunar environment, while beyond the focus of the current paper, is clearly important for future study. And while not all-inclusive, the social management issues of hygiene, personal relationships, and employee isolation underscore significant differences in the operational domain of lunar managers, compared to their terrestrial counterparts. Managing on the moon might be as much about adopting a different perspective on *what* to manage as *how*.

Personal Hygiene

A troubling arena for the erstwhile manager would be his or her foray into the heightened sensitivities of diverse individuals. Whereas military authority provides a ready solution to an individual failing to meet the minimum standards of hygiene, no such absolute authority typically operates for civilian support or scientific personnel. There are many examples of extended duty operations being complicated by having individual members fail to meet minimum standards for personal hygiene (Fraser 1968). Conflicts were often avoided only because others within the group did not wish to provoke a situation that might seriously affect group performance. There is a definite cultural component to opinions and perceptions of cleanliness. These views can even vary within a society depending on whether one is describing military or civilian behavior. Culturally speaking many Americans, compared with people in other societies, appear extreme in their concern for personal hygiene. Most Americans believe that they are entitled to at least one shower per day, whereas daily bathing is considered unusual in many countries (Stuster 1996), but Americans are not the only society in which standards for hygiene might appear extreme.

What matters from a managerial perspective is that these standards are learned. Consequently, there is the possibility to train people to accept lower standards to a point. Beyond that point there will be an increasing array of problems and issues with the potential for confrontations. In any environment having a mixed military and civilian staff, employees from various cultural backgrounds, as well as men and women (as a moon facility might), differences in the expected level of hygiene would cause friction that a manager would by necessity need to address. Confrontations over personal hygiene always have the potential to escalate into major problems with serious consequences for operational performance and employee morale. Complicate the situation with insufficient water for bathing, special conditions related to low gravity, and a closed environmental system, and a small thing like failure to take personal hygiene seriously quickly becomes a manager's problem.

Privacy and Personal Space

Who-does-what-and-where issues will also become critical management concerns in a lunar facility. In the initial stages of a lunar facility's development, work and living space is probably going to be scarce. Even among the most congenial and calm individuals there is a need for privacy and the desire to be left alone from time to time. Without a place to be left alone in, people have been known to become increasingly difficult to manage. Faced with similar situations, leaders of polar expeditions and long-duration military operations have tried to select crew members according to exacting criteria that would minimize the potential for friction. There are suggestions that all going to a lunar operation be subjected to extensive testing to evaluate their ability to tolerate isolation and to get along with others in a close environment. For small groups, extensive testing makes some sense; however, as the groups get larger and the time constraints become tighter, the relative cost in terms of time and preparation may not represent as good a tradeoff. Selection of employees for their ability to get along might be preferred by everyone; however, firms with tight budgets may feel compelled to choose skills over personality. Regardless, lunar managers will be expected to deal with a wide range of habitability issues. Roommates, snacking, clothing, after-hours parties, gambling, alcohol consumption, etc., are all areas that managers and police in a closed community will have to establish policies for.

Habitability issues clearly range from the sublime (long pants versus shorts) to the serious (availability of alcohol). Managers may wish to remain detached from such factors in preference of the more important matters of business. However, experience on submarines, with commercial saturation divers, with Antarctic winter-over personnel, and other remote operations have demonstrated time and again that these factors are of some consequence. Operational efficiency is still about people and how they respond to their environment. The harsher, the more difficult, the more remote that environment is the more important seemingly small issues of habitability become.

An example of one such issue that has the potential to bedevil mangers for some time to come is the availability and use of alcohol. There is a large gulf between the European and American approaches toward alcohol consumption. Among members of the European Space Agency the availability of alcohol represents a significant issue. After all, what would a meal be if there were not appropriate glasses of wine to accompany it? Should not European employees be able to enjoy the simple pleasures they have every reason to expect at mealtime? Alternatively, the American approach, viewing alcohol as only a problem waiting to happen, might be to eliminate it all together. In American facilities alcohol-free dining may be the norm. However, in international facilities, particularly those owned or operated by European firms, the American policy would be expected to give way to the European tradition. In mixed-use facilities, those with many firms or nationalities, the compromise may be limited availability. As a larger and more diverse group of employees find themselves working on the moon, a policy of alcohol exclusion may give way to one of limited availability. The management teams of different firms from different nations will be expected to create and implement successful solutions for all manner of "people-related" problems that might threaten the efficiency of lunar businesses.

Personal Relationships

Human interactions of all sorts involving employees and managers will become much more significant in a lunar environment. What might pass for poor manners on the job back at company headquarters on Earth could easily take on added meaning in the close confines of a lunar facility. Politeness and tolerance would by necessity become the lubrication for successful business operations. However, people are people wherever they work and live.

Perhaps no task would become more difficult and fraught with more managerial peril than that of overseeing the interpersonal lives of employees. Few managers would wish to be responsible for their employees' sexual activities. Yet someone may very well find himself or herself assigned to that task in the initial phases of a lunar operation. Of all the potential pitfalls for an erstwhile lunar manager, none is more problematic than becoming the arbitrator of sexual mores. No matter what course of action is decided, someone, some group, or some set of groups, will be offended. Businesses, which hope to avoid controversy, will be thrust headlong into an ongoing, unending series of public discussions.

It is a forgone conclusion that men and women will be working together on the moon. Some would suggest that benign neglect might represent the best official policy. However, from a business perspective, unintended pregnancies could be extremely costly on a variety of dimensions. Experience with the U.S. military's mixed operations suggests that sexual activity is both common and frequently produces the same results as in the civilian world. The data are not available that would allow for a complete evaluation of the risks inherent in a lunar pregnancy. But a reasonable policy would be to view pregnancy as something to be avoided until more information is collected on the long-term health effects of living and working in a lunar environment.

The managerial problem is along the lines of how to forestall the inevitable. Celibacy is a good option. However, it has not worked even in the strict environment of the U.S. Navy. People seem to manage to find time and opportunities, even on very busy and crowded ships, to engage in personal relationships. Mandatory birth control is another option. U.S. firms would run headlong into judicial and legislative precedent that forbids firms linking jobs to avoiding pregnancy. In their treatment of pregnancy, U.S. firms are also bound by Title VII of the Civil Rights Act of 1964, as amended with the inclusion of the Pregnancy Discrimination Act (PDA) of 1978 (Section 701 (k)). The amendment prevents employers from treating pregnancy, childbirth, or other related medical conditions in a manner different from other disabilities. Women "disabled" due to pregnancy, childbirth, or other related medical conditions must be provided with the same benefits as other disabled workers.

What this might mean in a lunar environment remains to be seen. An employer who does not provide disability benefits or paid sick leave to other employees is not required to do so for pregnant workers. It is unlikely, however, that the ideal solution would be to eliminate all benefits to prevent awarding others. More unlikely still would be the proposition that businesses would be able to entice men and women to commercial centers on the moon without benefits commensurate with the dangers and risks. Even the issue of whether it should be the men, women, or both who practice birth control would be controversial. This doesn't even get into the problems associated with religious freedom, cultural differences, or the myriad of other complicating factors. Perhaps the best option is to allow only those who volunteer to be temporarily sterilized to apply for lunar positions. At best, even an all-volunteer approach would be problematic. At least this is one issue that can be thought about in advance. Others may not be so obvious.

Ultimately, facilities on the moon would take on most, if not all, of the attributes one expects of any large, diverse community. Until that time, whenever that might be, lunar managers are going to have their hands full of problems they might have avoided by staying on Earth. However, avoiding complex situations by ignoring them is not what leadership or good management is supposed to be about. Critical and problematic areas exist in every business activity regardless of its location. Lunar operations will be no different. However, the critical areas facing lunar-based managers will require uncommon leadership because of the sensitive, politically charged, inherently personal nature of the issues themselves. Undoubtedly, solutions to even the most vexing operational and performance-related issues would be found. Unfortunately, solutions may not be implemented before one or more issues become significant, well-publicized, and potentially career-limiting. It would be far better to identify potential problem areas and to develop appropriate solutions in advance wherever and whenever possible. Some of these problems, and their relative severity (based on where they are encountered), are depicted in Figure 3.

Managers of non-terrestrial operations can expect to deal with problematic issues similar to those their terrestrial counterparts encounter, the key difference being the severity of their downside impact. For

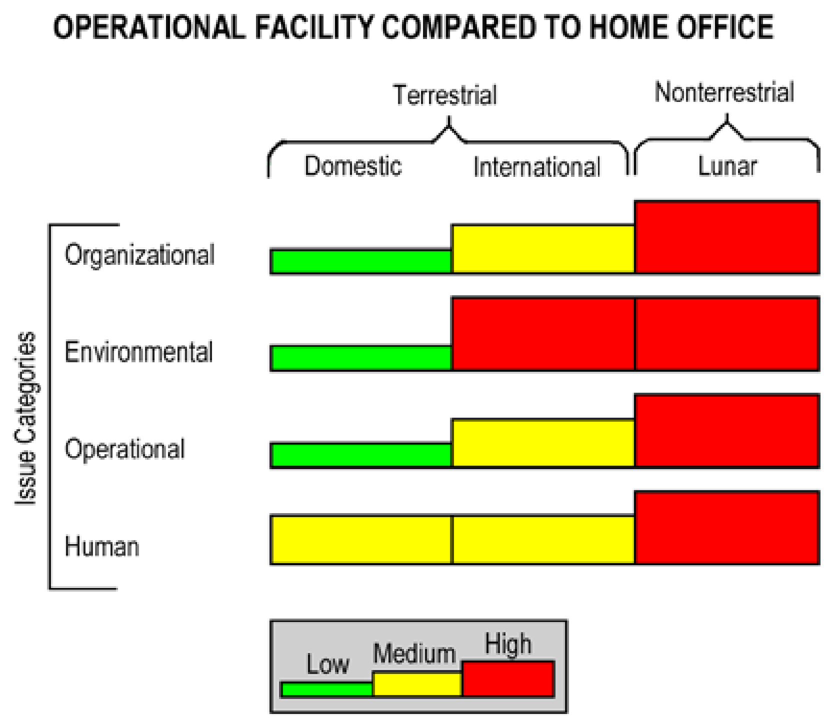


Figure 3. Criticality/Downside Severity of Various Management Issues

example, inventories of critical parts could fall below anticipated needs, resulting in an out-of-stock situation. Minimizing inventories is typically viewed as a prudent business practice, unless it is done in order to inflate profits or to hide real costs. When such activities lead to safety or operational efficiency concerns, the behavior moves from merely disreputable to criminal (and perhaps lethal). Assessing the severity of such potential problems is rooted in the need for almost total self-sufficiency in the lunar environment due to its unforgiving nature. It is easy to imagine any number of scenarios in which a short-sighted decision to keep inventory costs down could produce a tragic situation on the moon (e.g., emergency oxygen canisters; replacement parts for heating, radiation protection, and water reclamation equipment; surgical supplies).

The issue of self-sufficiency while operating in a hostile environment has its parallels on Earth. Severe storms have destroyed innumerable commercial facilities and taken many lives. The damage done to offshore oil drilling platforms in severe conditions and the periodic loss of life serve as reminders that even the most robust structures can fail under the right conditions. The recently well publicized rescue of a physician stationed in Antarctica but in need of medial treatment herself illustrates the logistics and weather problems associated with supporting operations in remote areas on the surface of our planet. Translating them to another celestial body magnifies the time and effort, and perhaps even precludes the possibility of rescue operations. Imagining the failure of a lunar facility on the scale of an off-shore drilling platform raises the issue of whether abandoning the damaged facility and waiting for rescue in some sort of lunar lifeboat would even prove to be workable.

Operational issues and human factors have always been concerns for plant managers. Sloppy handling of even seemingly innocuous equipment has been known to cause extreme damage and result in the loss of life. Punctilious attention to proper procedures (even approaching the level of religious zeal) in an environment where failure to pay attention can get you or someone else killed must be mandatory. Yet history is replete with examples of organizations and individuals engaging in behavior that produces the very result they hoped to prevent. The Chernobyl nuclear power plant disaster may be the classic example. The very people training to prevent a nuclear accident caused it to occur, in part, because they became careless. Even well-trained individuals occasionally do dumb things.

The possibilities of extreme negative consequences for all areas related to the operation of commercial lunar facilities compared to their terrestrial counterparts are magnified because of

- the lunar environment's inability to support human life without extensive life support (e.g., oxygen, water, heat, radiation protection);
- the relative distance to the moon, and hence the time it would take to provide relief in the event of a serious life-threatening event or situation (e.g., lack of more than a short-duration abandonment strategy);
- the potential serious consequences that even seemingly simple business oversights could have for safety and/or operational effectiveness (e.g., "You wanted O_2; I thought you said H_2O");
- the potential danger that might result from human error caused by boredom and/or inattention to details ("I thought Frank was right behind me, so I didn't secure the hatch"); and
- situations that result from isolating groups of individuals together for prolonged periods of time (e.g., lack of visual stimulation, individual idiosyncrasies becoming the basis for feuds, irritations).

Summary

The key components of business success are likely to remain constant regardless of its location, whether on the Earth or on the moon (Ryan 1999). Paradoxically, the factors critical for managerial success in a lunar environment will likely require some radical departure from managers' current expectations and experience. Those selected to manage lunar operations will need to recognize and respond to the differences requisite to leading a work group on the moon. They will, by necessity, find themselves acting in capacities seldom dealt with in business schools or within their experience, outside of the military. Preparing to operate a lunar facility will require managers to exhibit greater leadership ability than would most business operations. Furthermore, lunar managers will need to acquire a diverse set of skills more common to colonial governors of the 18th and 19th centuries than to 21st-century technological wonder kids. Fortunately, emerging technologies and developing technological shifts may assist lunar business managers performing some of these tasks. Firms must recognize that the environment in which employees will work will likely require unique managerial talent, new approaches to common business issues, and a distinct departure from present managerial skill sets.

The authors are currently developing a follow-up article, furthering the discussion of the management architecture needed to successfully operate nonterrestrial, entrepreneurial facilities. This new paper will examine evolving, nonsocial management issues. We are actively soliciting readers' views on any aspect of managerial architecture they view as pertinent to future commercial activities. Please send any and all comments, musings, thinly veiled threats, personal examples, and/or good stories that might be helpful to Dr. Mike H. Ryan at mryan@bellarmine.edu.

References

Bass, B. M. (1990). *Bass & Stogdill's Handbook of Leadership* (3rd edn). New York: The Free Press.

Bowie, N. E. (1990). "Business Ethics and Cultural Relativism" in P. Madsen and J. M. Shafritz (eds.), *Essentials of Business Ethic*s. New York: Meridian.

Brown, S. P., Cron, W., and Slocum, J. W., Jr., (1977). "Effects of Goal Directed Emotions on Salesperson Volitions Behavior, and Performance: A Longitudinal Study," *Journal of Marketing*, January.

Churchill, G. A., Jr., Ford, N. M., Walker, O. C., Jr., Johnston, M. W., and Tanner, J. F., Jr. (2000). *Sales Force Management* (6th edn). Boston: Irwin/McGraw-Hill.

Cornelius, P.E. (1991). "Life in Antarctica" in A. A. Harrison, Y. A. Clearwater, and C. P. McKay (eds.). *From Antarctica to Outer Space: Life in Isolation and Confinement.* New York: Springer-Verlag, 9–14.

Deresky, H. (1994). *International Management* (1st edn). New York: HarperCollins.

Elashmawi, F., and Harris, P. R. (1993). *Multicultural Management: New Skills for Global Success*. Houston, Texas: Gulf Publishing Company.

Fraser, T. M. (1968). *The Intangibles of Habitability During Long Duration Space Missions*, Albuquerque, New Mexico: Lovelace Foundation (NASA CR 1084).

Glover, M. K. (1990). "Do's and Taboos: Cultural Aspects of International Business," *Business America*, August, 2–6.

Ingram, T. N., LaForge, R. W., and Schwepker, C. H., Jr. (1997). *Sales Management: Analysis and Decision Making: Third Edition*. Fort Worth, Texas: The Dryden Press.

McAuliffe, Robert E. (1998). "Leadership" in Gary L. Cooper and Chris Argyris (eds.), *The Concise Blackwell Encyclopedia of Management.* Blackwell Publishers, Ltd., 355–358.

Plank, R. E. and Reid, D. A. (1994). "The Mediating Role of Sales Behaviors: An Alternative Perspective of Sales Performance and Effectiveness," *Journal of Personal Selling and Sales Management*, 14 (Summer).

Ryan, M. H. (1999). "Business Scenarios for Space Development," *Space Manufacturing 12: 1999 Space Studies Institute Space Manufacturing Conference*. Princeton University, 101–110.

Stuster, J. (1996). *Bold Endeavors: Lessons from Polar and Space Exploration*. Annapolis, Maryland: Naval Institute Press, 139.

Vaernes, R., Warncke, M., Bergan, T., and Urshin, H. (1988). "Selecting the Right Crew for Future Space Stations: An Analysis of Selection Research on Offshore Divers, Aviation Pilots and Other High Risk Groups in Scandinavia" in *Proceedings of the Colloquium on Space and Sea*. European Space Agency.

Vanscoy, K. (2000). "The Hiring Crisis," *Smart Business for the New Economy*, July, 85–97.

Walker, O. C., Jr., Churchill, G. A., Jr., and Ford, N. M. (1977). "Motivation and Performance in Industrial Selling: Present Knowledge and Needed Research," *Journal of Marketing Research*, 14 (May).

Yukl, G. (1994). *Leadership in Organizations* (3rd edn). Englewood Cliffs, NJ: Prentice-Hall.

Zaleznik, A. (1977). "Managers and Leaders: Are They Different?," *Harvard Business Review*, May–June, 47–60.

Authors

Mike H. Ryan is Associate Professor of Management at the W. Fielding Rubel School of Business and former Director of Programs for the Bellarmine Center for eWorld Education at Bellarmine University in Louisville, Kentucky. Prior to his arrival at Bellarmine, Dr. Ryan established and operated a small multimedia company, Prometheus Press, Inc. He has been a faculty member at four major universities including Texas Christian University, the University of North Texas, Southern Methodist University, and the University of Texas at Dallas. He has served as a consultant for a variety of firms and government agencies on issues ranging from technology and innovation to strategy and public policy. He is the author or editor of several books along with numerous articles and presentations. He has written and lectured extensively on the issues related to doing business in space with an emphasis on strategy and practical operational considerations. Dr. Ryan has presented his work in a variety of venues including the Space Studies Institute's Conference on Space Manufacturing, The Case for Mars Conferences, International Space Development Conferences, and Space Frontier Conferences, as well as providing testimony on space and technology issues to the Congress of the United

States and the State of Texas. Dr. Ryan holds a PhD in Management Science and an MS in Management and Administrative Science from the University of Texas at Dallas and a BA in History from the University of Dallas.

Michael R. Luthy is Associate Professor of Marketing at Bellarmine University's W. Fielding Rubel School of Business in Louisville, Kentucky. He received his PhD and BS degrees in Business Administration from the University of Illinois at Urbana-Champaign and his MBA degree from the University of Iowa. Dr. Luthy's professional work experience has been with the U.S. Treasury Department, Oscar Mayer Foods Corporation, and consulting with organizations in the service sector. His areas of research interest include services marketing and customer satisfaction, sales force training and management, international negotiation behavior, and business education pedagogy. His teaching interests include services marketing, marketing management, and sales management. Dr. Luthy's research has appeared in a variety of outlets, including *Industrial Marketing Management, Theory & Psychology,* and *International Research in the Business Disciplines,* as well as numerous conference proceedings. He is Professional Certified Marketer and a member of the Society for Marketing Advances, Marketing Management Association, and American Marketing Association, who presented him with their Teaching Innovator of the Year award in 2000.

Science Fiction as an Engine of Prediction

William I. McLaughlin

Abstract

Science fiction is distinguished by its focus on the future, and the future is stocked with new things. The genre's greatest predictive success, begun by Jules Verne in 1865, was in portraying spaceflight as possible, well in advance of Konstantin Tsiolkovsky's pioneering efforts in space engineering. On occasion science fiction goes beyond being an agent of prediction and participates in the coming-to-be of something: inspiring space-flight planning in the first half of the twentieth century and, at present, the human settlement and terraforming of Mars. A "band of prediction" can be identified, its lower end populated with specialized conceptions such as gravity assists (partially anticipated by Lester del Rey), and its upper end containing far-out ideas such as the transporter of *Star Trek*. (The transporter also serves to illustrate an issue relevant to extraterrestrial intelligence.) The band can be outlined using two "laws" due to Arthur C. Clarke, plus one supplementary condition. Finds at the lower end of the band might be increased through active mining of science fiction using the technologies of optical scanning and natural-language processing.

I. Science Fiction. The primary characteristic marking the province of science fiction (SF) is its use of the future. Some branches of SF share with the mystery story a reliance on ratiocination; concern for social change exists within SF and throughout literature; adventure has been a basic ingredient since tale-telling began; nonhuman creatures are not restricted to SF, appearing in horror stories and fantasies. However, it is the future that can be claimed by SF as almost its private fictional domain. Since the future is well stocked with things that do not exist in the present, it is not surprising that SF provides a spawning ground for predictions, a foretelling of things to come.

Of course, there are exceptions to SF as future fiction. The concept of parallel or alternate worlds presents an "other" without, necessarily, any temporal relation to our present. The idea of science fiction set in the past, even apart from tales of time travel, is possible: the popular television series *The Wild Wild West* (1965–1970) was a blend of western and SF themes, the technological innovations being relative to the norms of the time. Another example of temporally-displaced SF is *Story for Icarus*, a novel by Ernst Schnabel (1961) that is based on the fabulous inventions of Daedalus, the first engineer.

There is no general agreement on who was the first science-fiction writer. Claims go back to ancient times and become more numerous (and convincing) after the close of the Middle Ages. Candidates include Jonathan Swift (1667–1745), Mary W. Shelley (1797–1851), and Edgar Allan Poe (1809–1849), who, most agree, invented the detective story. See Disch (2000) for a historical discussion (and advocacy for Poe). Jules Verne (1828–1905) serves as the founder of science fiction for the purposes of the present paper, and the exact starting point is his novel *De la Terre à la Lune (From the Earth to the Moon)* of 1865. As Peter Costello (1978) puts it, "... what he writes is truly the beginnings of *science* fiction" (p. 18).

One should not try too hard to define SF. The only objects definable without fuzziness or ambiguity are those that are finite and rule-based, e.g., chess or the theory of finite groups (in mathematics). Allow complexity to enter, and the ability to wrap something in a definition is lost. Witness the host of unsuccessful attempts to define life or consciousness. Even mathematics, long held up as the subject that lives by definition, was found in the last century to itself elude definition. Kurt Gödel (1906–1978), through his incompleteness theorem, achieved the most spectacular result, but Thoralf Skolem (1887–1963) and others showed that even structures we think we know intuitively, like arithmetic, can have unintended, nonstandard interpretations.

However, it is quite possible to perform analyses that illuminate the structure of SF. One such is a factor analysis conducted by Bainbridge (1986). Factor analysis is an established technique of statistics that allows the analyst to discover what underlying factors, and in what proportions, might be used to explain data. In this case, the data are taken from questionnaires completed by hundreds of SF writers, readers, and critics.

Bainbridge identifies four factors, the major categories of SF, which he labels "hard science," "new wave," "fantasy," and "classic."

The labels are reasonably self-explanatory. Classic is SF written by early authors in the field. As Bainbridge says (p. 39), "... [this] factor expresses a residue from the historical infancy of the field." He makes the judgment that fantasy is not true SF, rather a companion field of writing, but useful for the contrasts it provides. Terms like "new wave" are often employed in literary analyses (or with regard to other forms of art): "modern" is a popular alternative.

Two axes emerge as a coarse architecture for SF as revealed by these four factors. One is time-based—new wave is recent work and classic is old work—the other is world-view based; hard science is rational while fantasy is romantic and mystical.

Another way to sort the factors is by (statistically) correlating them with the physical sciences and the social sciences (p. 44). It is no surprise that the hard-science factor strongly correlates with physical science and is neutral with regard to social science. Fantasy has a slightly negative correlation with physical science and is in neutral balance with respect to social science (zero correlation). The two temporal factors, classic and new wave, go separate ways: classic correlates strongly with physical science, new wave with social science. This difference is not difficult to understand when one notes that the social ferment of the 1960s fell between the works included within these two factors.

We would expect SF to yield predictions within three of the four factors Bainbridge finds, with classic and hard science supplying material in the physical sciences and new wave in the social sciences. His analysis is somewhat dated, in a field whose structure exhibits strong temporal dependence. Nevertheless, we can enter into an examination of SF and prediction with the anticipation that the genre is not monolithic and will deliver items across a wide spectrum of human activity, a spectrum whose primary colors are given next.

II. PREDICTION. Prediction is a fundamental activity not only of humans but of any conscious creature. Recognizing this breadth, we will not attempt to treat the subject in a comprehensive manner, resting content with listing and illustrating four categories relevant to the present study—philosophy, physical science, social science, and engineering—with the intention of gaining insights into modes of prediction. More thorough expositions are available in Morgan (1980), and Morgan and Langford (1981).

Philosophy

Philosophical systems outline the state of the world or a piece thereof, and, on occasion, can prove to be astonishingly ahead of their time. The biophysicist Max Delbrück (1906–1981), himself a Nobel Laureate, says:

> ... if that committee in Stockholm, which has the unenviable task each year of pointing out the most creative scientists, had the liberty of giving awards posthumously, I think they should consider Aristotle for the discovery of the principle implied in DNA. (Delbrück 1971; Lowenstein 1999, p. 337)

He is referring to Aristotle's (384–322 PCE) theory of essence or form, better known as a constituent of his metaphysics.

Perhaps the most celebrated encounter of philosophy with prediction is in David Hume's (1711–1776) "sceptical philosophy," which denies that we have any reason other than habit to believe in predictions. In *A Treatise of Human Nature,* he says:

> ... our experience in the past can be a proof of nothing for the future, but upon a supposition, that there is a resemblance betwixt them. This therefore is a point, which can admit of no proof at all, and which we take for granted without any proof.

A lot of philosophical work from the eighteenth century to the present has been devoted to building on Hume or attempting to refute him.

Physical Science

If Hume formulated the classical position of doubt concerning the basis for predictions, Pierre-Simon Laplace (1749–1827) spoke for the opposition, using the language of physical science:

> Given for one instant an intelligence which could comprehend all the forces by which nature is animated and the respective situation of the beings who compose it—an intelligence sufficiently vast to submit these data to analysis—it would embrace in the same formula the movements of the greatest bodies of the universe and those of the lightest atom; for it, nothing would be uncertain and the future, as the past, would be present to its eyes. The human mind offers, in the perfection which it has been able to give to astronomy, a feeble idea of this intelligence. Its discoveries in mechanics and geometry, added to that of universal gravity, have enabled it to comprehend in the same analytical expressions the past and future states of the system of the world.

In the twentieth century, quantum theory and chaotic systems conspired to roll back Laplacian optimism, but prediction in the grand manner is alive and well with scientists forecasting the long-term prospect for the continuance of life in the universe, even after the Sun dies (Krauss and Starkman 1999). Currently, the most visible project involving prediction is estimating the extent of climate change and understanding the mechanisms that drive it.

Social Science

Within the social sciences, no credible theorist has emerged to produce anything resembling Laplace's manifesto, but from time to time an overarching theory of economics or history is proposed, supplying a framework for predictions.

High-profile work in the foundations for capitalism and socialism was done by Adam Smith (1723–1790) and Karl Marx (1818–1883), respectively, and their visions competed fiercely during the Cold War.

There is usually perceived to be little at stake when theories of history are debated, but public attention fastens on formulations with a grand sweep or those that seek to explain events of particular interest. Francis Fukuyama's thesis about the "end of history" is a recent example of popular history, and conspiracy theories of the assassination of President Kennedy never seem to be out of date. A general theory of history was developed by the English historian Arnold Toynbee (1889–1975), who claims to have identified 29 civilizations within the historical period. With "a civilization" the unit of historical study (as opposed to, say, "a nation"), Toynbee traces the relationships between civilizations and a theory of their rise, flowering, and decline in his multivolume *A Study of History*.

A proposal to assemble and make widely available a universal reference resource, functionally anticipating the World Wide Web by 50 years, was described by H.G. Wells (1866–1946) in *World Brain*, a collection of essays written in the 1930s (Wells 1938, McLaughlin 1996). His inspiration ultimately comes from the French Encyclopédistes of the eighteenth century. Wells failed to foresee modern electronic means of data distribution, envisaging radio and fast post for dissemination, but we must not be too harsh with him: who could anticipate that the capability to shuffle symbols with superhuman speed would change the world?

The social-science predictions we most commonly consult are economic forecasts, always of questionable value, and preelection polling, pretty solid if samples are unbiased.

Engineering

Engineering predictions weigh less in the larger society than in SF, but they are a well-established tradition.

There is a subcategory based on the "Seven Wonders of the World" school of design: large projects such as described in *Engineer's Dreams* by Willy Ley (1954). Ley did a good job of selection for his nine projects: one treats an underwater tunnel between England and France; three deal with generation of power from Sun, waves, and wind. A legion of lesser proposals exists, from Dick Tracy's wrist radio ("cell phone") to various kinds of would-be perpetual-motion machines.

An important set of engineering predictions is that which relates to space travel itself. This set is intertwined with SF: witness the science-fact columns that have often graced SF magazines. One of the great SF writers, Arthur C. Clarke (b. 1917), is also the originator of the communications satellite (1945), and throughout his career has probed the future through fiction and nonfiction. His *Interplanetary Flight* (1950), along with Ley's *Rockets, Missiles, and Space Travel*, educated the generation of engineers who carried us into space, and he has been kind enough to publish a chronology of the twenty-first century. For example, the entry for 2021 reads: "The first humans land on Mars and have some unpleasant surprises" (Clarke 1999a).

Not all predictions are verbal or mathematical: Chesley Bonestell's (1888–1986) paintings inspired engineers and scientists with imagined scenes on the Moon and planets when only Earth-based data were available (Bonestell and Ley 1949, Clarke 1999b, Clarke 1999c). Moreover, not all predictions are predictions. The above summary within four categories, brief though it is, illustrates several kinds of activities which, for the purposes of this paper, will be included under the "prediction" label because they grant glimpses of future states, even if it is not their primary intent to forecast. A syntax of prediction follows, with examples.

Predictions per Se

Global warming is underway; 2021 is not an unreasonable date for humans to land on Mars.

Proposals

These are closely related to predictions per se and may be difficult to distinguish, in practice, from them. Space enthusiasts, in 1925 or today, tend to advocate an action, such as landing on Mars, while also claiming the action is inevitable. But intellectually the distinction is obvious: a prediction is a statement that something will happen; a proposal says we should undertake something because it would be good for us.

Worldviews: Philosophical, Scientific, Social, or Engineering

A worldview presents a state of affairs that frequently implies the existence of new things, i.e., has predictions implicit (or explicit) within its compass. Such predictions are conditioned by the likelihood of the worldview, but all predictions are conditioned in some way. Worldviews that contain advocacy overlap "proposals." Modern cosmology is a worldview that can be used to predict how long life might survive in the universe. Smith and Marx had views of economics that made certain kinds of economic activities more likely to occur in the future (relative to them) because they were made explicit and attractive.

Denial of Possibility

Denying that something can be done may serve as a goad to action: a negative prediction has its uses. One of the most celebrated denials of possibility, undoubtedly an inspiration to early researchers in aeronautics, is due to Simon Newcomb (1835–1909), prominent astronomer and popularizer of science:

> The demonstration that no possible combination of known substances, known forms of machinery, and known forms of force can be united in a practicable machine by which men shall fly long distances through the air, seems to the

> writer as complete as it is possible for the demonstration of any physical fact to be. (Newcomb 1906)

Such denials not only inspire, they may have predictive force in themselves. "Clarke's first law" says:

> When a distinguished but elderly scientist states that something is possible, he is almost certainly right. When he states that something is impossible, he is very probably wrong. (Clarke 1999d)

The next four sections traverse again the categories of philosophy, physical science, social science, and engineering in order to see how science fiction deals with them regarding prediction.

In the nineteenth century, advances in astronomy and romantic currents in culture influenced the visual arts and colored scientific romances like those of Jules Verne and H. G. Wells (Figures 1 and 2).

III. PHILOSOPHY. The portrayal of galactic empires is a staple of SF (Clute and Nicholls 1993, pp. 461–462). From the fictional rise and fall of institutions and cultures Toynbeean in scope, one might hope to extract lessons that pertain to social science. However, the postulation of a galactic empire—Isaac Asimov's (1920–1992) *Foundation* and subsequent novels in the series are the canonical expression—is more interesting than its structure.

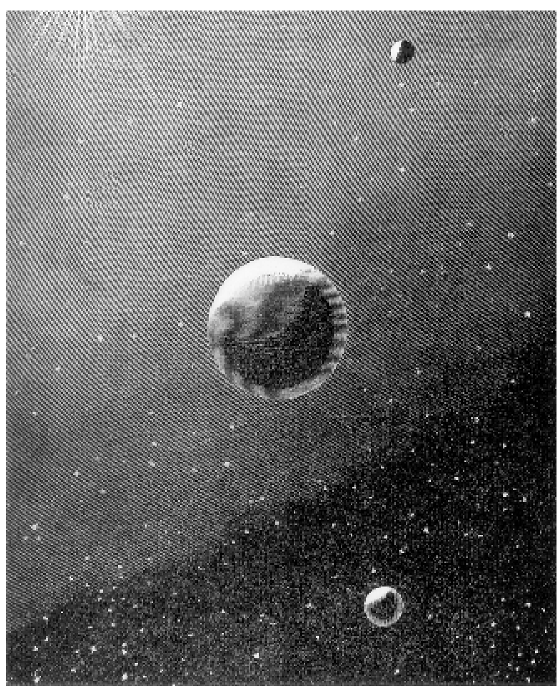


Figure 1. "The Earth in Space" from *New Descriptive Astronomy* (1884) by Joel Dorman Steele

Figure 2. From *The Breaking Waves Dashed High* (1880) by Felicia Hemans

A very likely consequence of a galactic empire (unless it is in a galaxy "far, far away" or it spread from Earth) would be the arrival of citizens of the empire at Earth. In 1950, physicist Enrico Fermi (1901–1954) asked during informal conversation with colleagues, "If there are extraterrestrials, where are they?" (Dick 1996). Fermi's question lay fallow for over a quarter century before its insight was appreciated. The universe is old, and even allowing time for life to arise and evolve, if there are intelligent extraterrestrials they would have by now swept over Earth in waves of colonization, or so the new Fermians calculate. The phrase "Where are they?" was picked up by Michael Hart and others who have an answer: "Intelligent extraterrestrials do not exist" (Hart and Zuckerman 1982). Carl Sagan (1934–1996) and his allies form the opposing school of thought, arguing that interstellar migration probabilities and/or rates are less than assumed by Hart et al., or other explanations of absence that do not foreclose on the possibility of the existence of intelligent extraterrestrials.

The importance of SF for resolution of the Fermi paradox, as it has come to be known, is not in its portrayal of galactic empires. To see where SF makes its most important contribution to the debate, we must turn from Isaac Asimov to Olaf Stapledon (1886–1950).

William Olaf Stapledon was born in Wallasey, England and received a BA from Balliol College, Oxford in 1909 and a PhD in philosophy from Liverpool University in 1924 (Crossley 1994). Stapledon was unsuccessful in establishing himself as an academic philosopher and also failed in an attempt to launch a career in poetry. He set out in an original direction, writing a set of philosophical novels cosmic in scope and unprecedented in their portrayal of long-term possibilities for life in the universe. His *Last and First Men* (1930) covers a period of two billion years, in which 18 varieties of humans come and go (*Homo sapiens* is at the first level). *Star Maker* (1937) extends the time span to 100 billion years. In these works and others, Stapledon's imagination ranges through a myriad of mind forms, body types, and social organizations in building a picture of a complex biological universe. (In fairness to Asimov, it must be noted that his galactic empires evolve to advanced states, but Stapledon's writings are saturated with a feeling for the "other.")

Stapledon's achievement is remarkable in two ways. The imaginative and literary qualities of his work are both high: it is no mean feat to produce a readable, meaningful, philosophical novel. (*Sophie's World* [1996] by Jostein Gaarder is a worthwhile philosophical novel and shows the potential of pure philosophy for producing SF-like effects, although its primary thrust is exposition of the history of philosophy.) Stapledon did not think of himself as a science-fiction writer and was surprised to be claimed by the genre (Clute and Nicholls 1993, pp. 1152–1153). The nature of his fiction (and of his political inclination) connected him to H.G. Wells, with whom he exchanged views. Also, Stapledon is cited by Arthur C. Clarke as a major influence on his own work (McAleer 1992), linking the twentieth century's three foremost futurists.

It is time to return to Fermi's paradox. The problem with the standard galactic-empire scenario is that it is too static. In reality, the rate of human evolution is increasing with the accumulation of technology (Russell 1983). Consider, for example, the large, recent leap in average life span and the prospects for accelerated change through genetic engineering. Our descendants would not tool around the galaxy with the same goals and in the same somatic forms of present-day *Homo sapiens*. There are two main roads to the future: 1) we will become extinct as a species (or stagnate on Earth), or 2) we will colonize the galaxy, but in increasingly altered form.

Gott (1993) has made a convincing argument for the first, the disaster scenario, basing his reasoning on the "Copernican principle" that we should not assume improbable states of affairs: don't base your financial planning on winning the lottery. From this principle, and some simple mathematics, he is able to draw far-reaching conclusions. Of course, Gott's arguments extend to "they" as well as "we" for the purpose of addressing Fermi's paradox. Read Gott's paper: it is a classic.

Stapledon's work comes into play with regard to the second possible future: galactic colonization. Even if space-faring species do not destroy themselves or retreat to isolation on their own planet, they would not arrive at Earth uttering the phrase, "take me to your leader." They, in the lapse of time while expanding through the galaxy, would have evolved well beyond planetary huggers, both in outlook and in form. Either we wouldn't recognize them as life forms or they wouldn't be interested in coming to our neighborhood. This kind of argument, an epistemological solution, has been made in the technical literature, e.g., McLaughlin (1983), but it is harder to conceive, imaginatively, than is the stagnation or extinction of our species. Reading *Star Maker* helps to avoid failures in imagination. It is not claimed that Stapledon's novels are consistent with the epistemological resolution of Fermi's paradox—issues of evolutionary timing and interstellar-migration rates remain—but he does make it easier to believe that there could be natural intelligences far beyond our ken. (However, Stapledon in his nonfiction does not subscribe to total epistemological disjunction between life forms: "Even if they possessed senses outside our range ... we would still be able to comprehend them" [Stapledon 1948, p. 233].)

Stapledon has enlarged our conceptions of how vast and strange the universe might be. A service in itself, it is one that also makes us more clear-eyed in predicting what the future might be like. It also keeps us from simple-minded assumptions when answering questions like Fermi's. Stapledon's literary achievement is rare and a hard thing to do. Some 20th-century abstract painters have undertaken work parallel to his, such as John Golding's *Paths to the Absolute* (2000), while pursuit of distant visions is not uncommon in music, e.g., the Canon in D of Johann Pachelbel (c. 1653–1706).

However, no matter how well Stapledon and other artists do their work, they can only open our eyes to possibilities. Whether intelligent species are prevalent and are more likely to follow Gott's scenarios to oblivion or Stapledon's growth curves, we do not now know. Research in philosophy and exobiology, and continuation of Search for Extraterrestrial Intelligence (SETI) programs are required in order to discover the facts about life and mind in the universe.

Interactions between philosophy and science fiction are extensive, and Miller and Smith (1989) provide an introduction to the relationship.

IV. PHYSICAL SCIENCE. Considering its name, "science fiction," and its distinctive tilt toward the future, one might expect the genre to be a cornucopia for prediction in the domain of the physical sciences. To a certain extent this expectation has been met, with some arresting realizations of fictional accounts. One of the best known concerns Jonathan Swift's *Gulliver's Travels,* where the author purports to describe the high level of scientific knowledge of the Laputans and speaks of their discovery of two satellites of Mars. Swift also gives the distance of each from the center of the planet: three Martian diameters for the inner satellite, five for the outer. At the time, no satellites of the fourth planet were known. In 1877, about 150 years after Swift's novel, Asaph Hall (1829–1907) discovered two in orbit about Mars using the 26-inch refractor at the U.S. Naval Observatory in Washington, D.C. The inner satellite, Phobos, was found to orbit at 1.4 Martian diameters from the center of the planet, while Deimos orbits at 3.5 diameters.

Swift's account was probably based upon a tradition, going back to Kepler, that Mars has two satellites (O'Meara 2001). The novelist's awareness of Kepler's work is demonstrated in *Gulliver's Travels* when he, in effect, recites Kepler's third law of planetary motion. The relatively good agreement of fiction and fact with respect to the scale of the two orbits must be marked as accidental.

Clute and Nicholls (1993, pp. 957–958), in their article on prediction, give a number of examples within SF, but state, "For every correct prediction a dozen were wrong ... ." That seems to be a reasonable assessment. These kinds of predictions are entertaining to contemplate, either as possibilities for our future or, "hits," like Swift's, but there is another class of fictional account that has importance beyond entertainment: the stretch account.

The film *Star Wars* (1977) and its three successors represent a popular success for science fiction that is matched only by the *Star Trek* constellation. The films emphasize entertainment, not venturing deeply into matters of science or characterization. Nonetheless, the *Star Wars* films contain advanced conceptions that stretch our scientific credibility to the breaking point. A stretch account, like a denial of possibility, can generate creative tension and induce people to speculate: how could such things be? In order for a stretch account to yield fruitful speculation, it should be attractively presented (as art or as entertainment), so that it gains entry to the mind, and it should reach a sufficient number of people that active minds are likely to be met and sparked. *Star Wars* meets both conditions.

Jeanne Cavelos (1999) says, "As I went through college studying astrophysics, though [she had first seen *Star Wars* while a junior in high school and was captivated by it], I was taught again and again the scientific truths that made *Star Wars* impossible. We cannot travel faster than the speed of light ... . And the Force? Pure fantasy." In *The Science of Star Wars*, Cavelos endeavors to construct a framework.Within this framework, which the Force—an all-pervading substance that adepts can draw upon to allow them to levitate objects, see the future, engage in telepathy, etc.—might possibly be seen, if only dimly, as scientifically comprehensible, rather than as an object of pure fantasy. (See also Krauss 1997.)

The path Cavelos takes runs through quantum-theoretic models of modern physics. It employs both particle and wave representations in an attempt to relate our scientific worldview to the kinds of mental faculties required by those who employ the Force. This is a sound approach. More than one student of the mind has appealed to quantum theory for help, particularly in circumventing strict determinism with regard to human actions, e.g., Eccles and Robinson (1984). Moreover, many interpretations of quantum theory itself are famously tied to the workings of consciousness. In an advanced society, Cavelos notes, one might expect the human mind to be assisted by impressed technology. The vacuum of physics, long known to be stuffed with energy and a source of virtual particles, is identified by Cavelos as an analog for a physical medium with the pervasive nature of the Force.

A stretch account, by its nature, is a top-down approach to a subject. Instead of synthesizing a structure by means of logic, it starts with a (fictional) whole, challenging us to explain by analysis how it might possibly exist. Stretch accounts have an analogy in astrophysics when an exotic, new object, e.g., a pulsar in 1967, is observed and an explanation sought. Another analogy exists in two contrasted methods of proof used for demonstrating theorems in elementary geometry—analysis and synthesis. Synthesis goes from the known to the unknown: connecting from axioms to the purported theorem. Geometrical analysis reverses things and moves downward, seeking to make contact with something known: from the purported theorem to axioms (Heath 1981). Analysis, then, resembles attempts to make sense of a stretch account.

In the course of her discussion of the Force, Cavelos states, "As Arthur C. Clarke said, 'Any sufficiently advanced technology is indistinguishable from magic.' The power of the Force certainly seems magical" (p.182). A stretch account, to qualify as SF, must be at least partially soluble in the acid of logic. Potential benefits from stretch accounts are several: they entertain and inspire; puzzles, like games, have proven their value in the service of discovery; a new perspective on the world may open up otherwise neglected lines of thought.

Bonestell's paintings of imagined scenes on planets (or Swift's fictional portrayal of two Martian satellites) and the Force of *Star Wars* illustrate two ends of a band of SF prediction. The band is not definable with precision, but below it lies triviality and above it lies pure magic.

This SF prediction band ranges from modest below to mind-boggling, stretch accounts above. The probability of it coming true may decrease as one goes "up," but not necessarily. It can also be characterized by two "laws" formulated by Arthur C. Clarke. Clarke's first law can be understood by examining Simon Newcomb's opinions on airplanes and defines the lower boundary of the prediction band. Cavelos cites above what is usually called Clarke's third law, which defines the upper boundary of the prediction band. (Clarke's second law, by the way, says that the only way to find the limits of the possible is by going beyond them to the impossible, which might be taken as a credo of science fiction.)

In the region above the prediction band, there are the fantastical and the purely magical, some of which might, in time, migrate inside the band. Up there also resides the invisible, an example of which was given in the discussion of Fermi's paradox—the extravagantly advanced extraterrestrial. While a computer would capture Aristotle's attention and seem magical to him, it would only be recorded as inedible by an ant, its primary function invisible to the insect. Hence, the observation that “extravagantly advanced technology is invisible” serves to supplement Clarke’s laws in charting the prediction band for SF.

V. SOCIAL SCIENCE. Are pajamas inevitable? In SF film and TV, undistinguished, pajama-like apparel is often the costume of choice. However, the convention need not be interpreted as a prediction of social customs: the temporal neutrality of this kind of garb avoids distracting the viewer with anachronistic costumes drawn from an identifiable historical period. As such, the convention belongs to a set of theatrical practices and is not an omen. Similar convention is the “555” telephone number, forming a sterile prefix and recited as part of dialog.

In the past, “take me to your leader” was dismissed as a likely form of address, but even without the prospect of ubiquitous pajamas or imperious aliens, SF has created a rich collection of social behaviors and organizational principles. For the lower end of the prediction band, a partial inventory is given in Clute and Nicholls (1993, pp. 957–958). There are societies driven by rampant consumerism; sporting events that resemble bloody Roman games; cities enclosed protectively, like medieval walled towns; societies made comfortable or miserable with advanced technology; and much more.

Striking as these representations are, “social science” is dominated by items at the top end of the band. This dominance issues from one concept—utopia, the complete makeover of society, creating an ideal world. With utopia in the future, SF reverses the prevailing attitude of classical antiquity, which places a Golden Age in the past and descends to a present Iron Age. Ovid’s (43 BCE–c. 17 CE) *Metamorphoses* describes this devolution from best to worst through intermediary Silver and Bronze Ages. In order to balance the ledger while still retaining the utopian option, the concept of dystopia (the dysfunctional place) flourishes along with utopia (Aldridge 1978). A less abstract reason for its existence is the greater opportunities dystopia gives for dramatic developments: compare Dante’s *Inferno* with his *Paradiso* or Milton’s *Paradise Lost* with his *Paradise Regained*.

“Utopia” is a word assembled from Greek parts (*ou*, not, *topos*, a place) by Thomas More (1478–1535) for the imaginary island in his political romance *Utopia* (1516). Written in Latin, the work describes problems in contemporary England and then describes Utopia, a country where life unfolds rationally and happily. Dystopia flourished with special vigor in the disastrously-governed twentieth century with novels including *Brave New World* (1932) by Aldous Huxley, *1984* (1949) by George Orwell, and *Lord of the Flies* (1954) by William Golding.

Many utopian/dystopian stories hover between what one would clearly classify as SF and what is closer to a thinly fictionalized form of social commentary. Nevertheless, those that depend on space travel devices fall unambiguously in the former. The interstellar ark, a closed ecosystem with residents traveling through the galaxy for long periods of time, lies within SF. The first well-known treatment of the theme is Robert Heinlein's (1907–1988) short story "Universe" (1941) in *Astounding Science Fiction*. Brian Aldiss (b. 1925) matured the idea in a series of works starting with the novel *Non-Stop* (1958).

The concept of the interstellar ark has never been popular outside SF. Those who advocate interstellar flight favor faster modes of travel than the lumbering ark. Providing a place to live has also proved unappealing to futurists. They have looked more favorably upon artificial habitats that remain local and integrated with the larger solar-system society (O'Neill 1977) or revision of planetary environments (especially Mars or Venus) to make them suitable for human habitation (see below).

Nevertheless, the negative judgments may be premature. When the technology to create sustainable ecosystems matures—concerns about global warming are hastening that day—the interstellar ark may become the utopian venue of choice. Consider some advantages:

- Nearly complete independence from the larger society can be achieved, a goal for certain religious and secular agendas.
- The architecture of "the universe" can be chosen by its inhabitants (putting a better cast on "universe" than did Heinlein in his story: those inhabitants had forgotten their origins and believed their ark was the universe, in every sense of the word).
- If desired, a degree of contact can be maintained with the larger society without compromising independence.

With regard to the last point, one might envisage funding obtained for an ark with the provision that it send back periodic scientific reports or other news to the solar system. Since World War II, the number of independent nations has increased dramatically, and it is not implausible to foresee the splintering to continue to the ultimate point of hiving off interstellar arks for the disaffected or the adventurous. If pursued in a noncolonizing mode, the idea does not collide with Fermi's paradox.

After the ending of the Apollo lunar-landing program, it required a few decades for advocates of human spaceflight to develop credible plans for exploration beyond Earth orbit. Going back to the moon appeals primarily to a small group, nostalgic for the heady days of Apollo missions. That is not to say a lunar site lacks advantages: astrophysical facilities and certain schemes for energy generation may find a home there. Recent advances in mission design for human trips to Mars have put that planet on the table for discussion, removing the enormous price tags that had been estimated for going there (Zubrin 1996). In addition to design improvements, biology has become a focus for the rationale for a trip, a more effective driver than nostalgia or geology. Though not "enormous," price tags are still "very large," to exercise the language of superlatives, and a mission to Mars is unlikely to be made as an act of will or geopolitics, as was Apollo: more is needed for rationale than "not enormous in cost," and biological science may meet that need.

The turn toward interest in Martian biology has taken two paths. The first leads to investigation of possible indigenous life on the planet, past or present, and is symbolized by discussions of the Martian meteorite, ALH 84001 (Beatty 1999). Only a gleam in the eyes of a few, the second looks at ways to create a flourishing Martian biosphere, perhaps one that could even support human life outside of enclosed habitats.

The word "terraforming" was coined by SF writer Jack Williamson (b. 1908), writing under the pseudonym "Will Stewart," in the story "Collision Orbit" in the July 1942 issue of *Astounding Science Fiction*. Transforming other worlds to match a terrestrial template, terraforming was part of Olaf Stapledon's master plan for the universe. The most comprehensive scientific and engineering synthesis of the subject is contained in the works of British engineer Martyn Fogg (1995, Hiscox and Fogg 2001). Terraforming of Mars is likely to be expensive and require centuries to complete, but there is some hope that shortcuts exist and that "nudging" will suffice to reach a useful state rapidly and cheaply.

Nearly a century ago, when spaceflight seemed impossible to most, SF provided an imaginative setting that encouraged astronautical pioneers to persevere (the mind requires more than data and theories to carry on with its work). (See the next section, "engineering," for the role of SF in early spaceflight.) Today, SF is performing a service, parallel to its work of a century past, in making people more comfortable with the idea of terraforming Mars, inducing them to envisage a set of acts that will yield a new, habitable world. With terraforming we are at the upper end of the band of prediction, but its elevated status is, unlike the Force, mostly owed to social issues, not those of physics. Why would we want to do such a thing? Would the end result justify the cost and the loss of old Mars? Is it worth persevering? These kinds of questions are why the topic is included in "social science."

Kim Stanley Robinson (b. 1952) has written, in epic style, of the terraforming of Mars. His trilogy, spanning 200 years, begins with *Red Mars* (1993) when 100 settlers land on the planet in 2027 (the first human landing was in 2019) and begin the process of producing an atmosphere and introducing plants and animals. The tale, continued in Green Mars (1994) and Blue Mars (1996), is rife with human conflict. It erupts into revolution and sees the planet terraformed. Canals and shallow seas form a part of the new world and political independence from Earth is achieved. Neither planet can prosper alone, neglecting the other: Mars has a role in regenerating Earth.

An extensive exercise in story telling such as Robinson's accustoms the mind to a difficult-to-accept idea—here, terraforming a planet—so that automatic-rejection mechanisms are disabled. At the same time it invites the reader to exercise his or her imagination in creating other possibilities. SF has a legitimate role, complementing the compression of science and the glossiness of advocacy, in creating a balanced picture of the future, particularly when large, complex undertakings are being considered.

VI. Engineering. Lester del Rey (1915–1993) is not in the front rank of SF writers, but over a long career he produced some memorable work, much of it illustrative of engineering aspects of SF. He attended George Washington University for two years before dropping out, but extended his education through a variety of enthusiasms and jobs: photography, electronics, helping assemble DC-3 airplanes, managing a fast-food restaurant, all while writing science fiction stories. Elements of his life prior to making the decision to become a full-time writer are contained in a series of biographical inserts in a collection of his early fiction (del Rey 1975).

At least two of del Rey's topics assumed importance decades after his treatment of them: accidents in nuclear powerplants and the use of gravity assists for interplanetary spacecraft.

"Nerves," published in *Astounding Science Fiction* in 1942, is one of del Rey's best works and lies on the intellectual path to Chernobyl. It is part of a tradition in SF, going back to the early twentieth century, of exploring military and industrial ramifications of atomic energy. (See the article "Nuclear Power" in Clute and Nicholls [1993, pp. 881–882].)

The versatility of del Rey is exhibited in his handling of the gravity-assist technique: using a close approach to a celestial body to change the momentum of a spacecraft. Most notably, a series of gravity assists by Jupiter, Saturn, and Uranus made possible the exploration of the outer solar system by the two Voyager spacecraft (McLaughlin 1989), but this method of trajectory shaping has found application elsewhere as well. In 1939, del Rey's short story "Habit," published in *Astounding Science Fiction*, used Jupiter to redirect a spacecraft that was taking part in an interplanetary race. The story has been republished in del Rey (1975, pp. 46–57). The author was evidently taken by the idea (and the interplanetary-race story line), using the Sun for a gravity assist in his 1952 novel *Rocket Jockey* (republished in del Rey 1978). From his autobiographical note (del Rey 1975, pp. 57–59) following "Habit," it seems the author sometimes calculated dynamical properties of spacecraft trajectories, but laments that for this story he failed to do so in a proper manner. In addition, del Rey only partially anticipated the gravity assist. He was aware of the capability to redirect the velocity vector, but not that the speed of the spacecraft could also be changed.

The influence of del Rey's speculations on atomic policy would be difficult to trace, but one might guess that he added to the general stew of opinions. There is no evidence that his employment of gravity assists influenced the actual development of the subject (Cutting 1974, Dowling et al. 1990).

The single greatest predictive success of science fiction came from circulating the idea that it would be possible to travel in space. Such journeys had been addressed through the ages in myth, fiction, and non-fictional speculation (Wright et al. 1968), but it was not until Jules Verne's *De la Terre à la Lune* (1865) that a science fiction account of substance could be said to exist. Substantive engineering designs for spaceflight were first published in 1903 by Konstantin Tsiolkovsky (1857–1935) (McLaughlin 1999).

Science fiction precedes space engineering not only in chronology, but also causal links. Tsiolkovsky, the first of the great space pioneers, was influenced by the science fiction of Jules Verne and himself wrote SF, though not at the level of his technical work (Clute and Nicholls 1993, p. 1242). While Verne's influence was worldwide, a less well known writer, Kurd Lasswitz (1848–1910), exerted a strong influence on German space research early in the twentieth century, an influence that was amplified by the trail-blazing nature of this research. Willy Ley (1906–1969) participated in these activities and has written extensively on space matters. He says, with regard to Lasswitz's thoughtful brand of science fiction, "... German scientists [were] preconditioned to taking space-travel seriously ..." (Ley 1957, p. 114; see also pp. 45–48 on Lasswitz). Verne also influenced the German community, e.g., Hermann Oberth (1894–1989), one of the most prominent of the pioneers, read the French author as well as Lasswitz (Ley 1969).

Star Trek has two signature technologies: warp drive and the transporter, both at the upper edge of the band of prediction. In order for either to be realized, enormous problems of physics would have to be solved, problems that lie beyond the grasp of current theory. Both technologies involve philosophical problems, too. They are nonetheless motivated by transportation needs—warp drive for long hauls and the transporter for local hops between planetary surface and spacecraft—and, hence, are properly listed in the present section.

Of the two, the transporter is the more interesting stretch account because it raises a host of questions (Krauss 1995, Hanley 1997, Gresh and Weinberg 1999). One is whether, during a "beam me up" operation, the actual atoms of the affected person are transported or just the information, the bits, required to assemble the person at the point of reception. Is it a technology of reassembly or assembly? (The writers of *Star Trek* generally endorse the former technique, but not consistently.) Either option is troubling, technically and philosophically. To what level in the hierarchy of matter does one have to go to capture a person: atoms? protons, neutrons, and electrons? quarks? Does Heisenberg's uncertainty principle prevent knowing enough to do the job? ("Heisenberg compensators" are alluded to in *Star Trek*.) How can such humongous collections of information be manipulated and transmitted? Are persons nothing but their material constituents?

These are interesting questions about the operation of the device but do not touch the core issue it presents. The transporter concept is a freak of fiction—an important, cross-cutting find—and provides a measure of the potential span of knowledge of *Homo sapiens*, an indicator of how much we could possibly achieve as a species. Most far-distant technological aspirations are chimerical: we didn't need a constantly-improving steam engine because that technology was replaced. However, a transporter would be perfect, even in the far future, as long as we are a body-based species. Furthermore, it is a broadly-based concept, crossing many technical and philosophical lines. (As *Star Trek* indicates on occasion, the uses of the transporter extend beyond transportation because of its body-manipulation capabilities.)

The transporter marks the edge of our species, straddling a taxonomic line. It's not that we'll discover enough about science and technology to build a transporter and then go on intellectual "hold" as a species. Rather, having this knowledge would cause us to lose our biological identity as *Homo sapiens*. Consider. With the ability to disassemble the body, it would be easy to decide to cast the essence of humans into another physical form: the sentient cloud of science fiction is an option (Hoyle 1957, McLaughlin 1965).

The transporter of *Star Trek* would be the final technological milestone for our species and would represent a step in the evolution by technology to a state of invisibility, resolving the Fermi paradox.

VII. Evaluation of the Engine. Science fiction does not have a predictive role to play in the direct manner of technical enterprises such as trajectory propagation, weather forecasting, or business-cycle extrapolation. The reason is that fiction does not carry within itself the means for evaluating the probability that its creations foreshadow future events. The physical sciences, and to a lesser extent, the social sciences do: there is evidence to inspect and methodologies to assess. At best, SF is an incomplete engine of prediction, and the reader must look at each scenario, using external information in judging its implications for the future.

The reasons for the extraliterary successes that SF has enjoyed in dealing with the world of the future fall into two categories: domesticating the unfamiliar and cross cutting through new territories. An example of the first is Robinson's trilogy on the human settlement and terraforming of Mars. The second category, "crosscutting," threads plots through imagined futures and, on occasion, such as del Rey with the gravity assist, encounters something new and real. Crosscutting is chosen to describe the process because a story line often cuts through various social, physical, technical, and cultural systems in its evolution, looking at the components of the world in fresh ways, from new angles of vision. Analogies can be made with "end-to-end-information systems" in space engineering or books that cut across layers of culture by the device of relating the story of a ubiquitous item, e.g., a chemical element: *The 13th Element: The Sordid Tale of Murder, Fire, and Phosphorus* by John Emsley.

Domestication is most effective at the upper end of the band of prediction, while crosscutting works best at the lower end of the band. Domestication, of course, is not predictively oriented but renders a topic—spaceflight in the early twentieth century or terraforming today—emotionally credible so that the intellect can do its work of evaluation and planning. Crosscutting is responsible for "finds," but it is not the only mechanism of discovery. Sometimes pure invention enters the scene, creating something new, such as Verne's realistic 1865 depiction of spaceflight.

It is not clear that SF can repeat its massive predictive success with regard to spaceflight (and, to a lesser degree, its anticipations concerning applications of atomic energy). Although aliens of various stripes have featured in the genre, speculation on life in the cosmos is not nearly so concentrated within SF as was the possibility of spaceflight: philosophy and science have long participated in the various arguments (Dick 1996). The technology of terraforming is also in other hands, but if interstellar arks ever become a reality, SF can legitimately claim them as its own. Using the past as a guide, and noting the process of crosscutting (and pure invention), it seems likely the genre will continue to spin off previews of the future.

Despite Swift's relation of two Martian satellites or del Rey's incorporation of gravitational swingbys in stories, the overall impact at the lower and middle parts of the band has been slight. It could become significant if inspection of science fiction for ideas were to be done in a systematic manner. One can envisage a pilot project in idea mining where several readers of various backgrounds would parse the stories in *Early del Rey* and inventory their contents for ideas, allowing a judgment to be made of the quality and amount of ore in the mine. Of course, exercises in hindsight are not easy for the reason that they are too easy, which is why more than one miner is suggested. This process is too labor intensive to be of practical value when the large body of SF is contemplated. Thus, if a pilot project showed promise, labor-saving techniques would be required in order to be able to extract value. Natural-language processors have matured in the last decade (Katz 1997), and employing them to organize a large amount of optically scanned science fiction writings would greatly increase the scope of the operation.

References

Aldridge, A. (1978), *The Scientific World View of Dystopia*, UMI Research Press, Ann Arbor, Michigan.

Bainbridge, W.S. (1986), *Dimensions of Science Fiction*, Harvard University Press, Cambridge.

Beatty, J.K. (1999), Life at the Limit, 40–42, *Sky & Telescope* 98(3), September

Bonestell, C. and W. Ley (1949), *The Conquest of Space*, The Viking Press, New York.

Cavelos, J. (1999), *The Science of Star Wars*, St. Martin's Press, New York, p.xiii.

Clarke, A. C. (1999a), The Twenty-First Century: A Very Brief History, 534–539, in *Greetings Carbon-Based Bipeds!: Collected Essays, 1934–1998*, St. Martin's Press, New York.

Clarke, A. C. (1999b), The Conquest of Space, 75–77, in *Greetings Carbon-Based Bipeds!: Collected Essays, 1934–1998*, St. Martin's Press, New York.

Clarke, A. C. (1999c), Willy and Chesley, 305–309, in *Greetings Carbon-Based Bipeds!: Collected Essays, 1934–1998*, St. Martin's Press, New York.

Clarke, A. C. (1999d), Failures of Nerve and Imagination, 196–203, in *Greetings Carbon-Based Bipeds!: Collected Essays, 1934–1998*, St. Martin's Press, New York.

Clute, J. and P. Nicholls, editors (1993), *The Encyclopedia of Science Fiction*, St. Martin's Press, New York.

Costello, P. (1978), *Jules Verne, Inventor of Science Fiction*, Charles Scribner's, New York.

Crossley, R. (1994), *Olaf Stapledon: Speaking for the Future*, Syracuse University Press, Syracuse, New York.

Cutting, E. (1974), Missions to Venus and Mercury (letter), 79–80, *Spaceflight* 16(2), February.

Delbrück, M. (1971), Aristotle—totle—totle, 50–55, in *Of Minds and Life*, Jacques Monod and Ernest Borek, editors, Columbia University Press, New York and London.

del Rey, L. (1975), *Early del Rey*, Doubleday & Company, Inc., Garden City, New York.

del Rey, L. (1978), *Rocket Jockey*, Ballantine Books, New York, p.118.

Dick, S. J. (1996), *The Biological Universe*, Cambridge University Press, p.443.

Disch, T. M. (2000), *The Dreams our Stuff is Made Of: How Science fiction Conquered the World*, Simon & Schuster, New York, pp.32–56.

Dowling, R. L., W. J. Kosmann, M. A. Minovitch, and R. W. Ridenoure (1990), The Origin of Gravity-Propelled Interplanetary Space Travel, IAA-90-630, 41st Congress of the International Astronautical Federation, October 6–12, 1990, Dresden.

Eccles, J. C. and D. Robinson (1984), *The Wonder of Being Human: Our Brain and Our Mind*, Free Press.

Fogg, M. J. (1995), *Terraforming: Engineering Planetary Environments*, Society of Automotive Engineers, Inc., Warrendale, Pennsylvania.

Golding, J. (2000), *Paths to the Absolute*, Princeton University Press, Princeton.

Gott, III, J. R. (1993), Implications of the Copernican principle for our future prospects, 315–319, *Nature* 363, 27 May.

Gresh, L. H. and R. Weinberg (1999), *The Computers of Star Trek*, Basic Books.

Hanley, R. (1997), *Is Data Human? The Metaphysics of Star Trek*, Basic Books, New York, pp.123–162.

Hart, M. J. and B. Zuckerman (1982), *Extraterrestrials: Where are They?*, Pergamon Press, New York.

Heath, T. (1981), *A History of Greek Mathematics*, volume I, Dover Publications, Inc., New York.

Heinlein, Robert (1941), Universe, in *Astounding Science Fiction.*

Hiscox, J. and M. J. Fogg (2001), Terraforming Mars, 153–155, *Spaceflight* 43(4), April.

Hoyle, F. (1957), *The Black Cloud*, William Heinemann Ltd., London.

Katz, B. (1997), Annotating the World Wide Web using Natural Language, Proceedings of the 5th RIAO Conference on Computer Assisted Information Searching on the Internet (RIAO '97), Montreal, Canada.

Krauss, L.M. (1995), *The Physics of Star Trek*, Basic Books, New York, pp.65–83.

Krauss, L. M. (1997), *Beyond Star Trek*, Basic Books, New York, pp. 87–100.

Krauss, L. M. and G. D. Starkman (1999), The Fate of Life in the Universe, 58-65, *Scientific American* 281(5), November.

Ley, W. (1954), *Engineer's Dreams*, The Viking Press, New York.

Ley, W. (1957), *Rockets, Missiles, and Space Travel* (revised and enlarged edition), Viking Press, New York.

Ley, W. (1969), *Events in Space*, Popular Library, New York, p.24.

Lowenstein, W. R. (1999), *The Touchstone of Life*, Oxford University Press, Oxford and New York.

McAleer, N. (1992), *Odyssey: The Authorized Biography of Arthur C. Clarke*, Victor Gollancz Ltd., London, p.21.

McLaughlin, W. I. (1965), The Lonely Hours, 96–101, *Worlds of If* 15(1), November.

McLaughlin, W. I. (1983), Human Evolution in the Age of the Intelligent Machine, 307–319, *Interdisciplinary Science Reviews* 8(4), December.

McLaughlin, W. I. (1989), Putting Voyager Together, 4–7, *The Planetary Report* IX(4), July / August.

McLaughlin, W. I. (1996), The World Brain of H.G. Wells, 27, *Spaceflight* 38(1), January.

McLaughlin, W. I. (1999), The First Space Mission Architect: Tsiolkovsky, *Journal of Space Mission Architecture* 1:43–48.

Miller, Jr., F. D. and N. D. Smith, editors (1989), *Thought Probes: Philosophy through Science Fiction Literature*, second edition, Prentice Hall, Englewood Cliffs, New Jersey.

Morgan, C. (1980), *The Shape of Futures Past: The Story of Prediction*, Webb & Bower, Ltd., Exeter.

Morgan, C. and D. Langford (1981), *Facts and Fallacies: A Book of Definitive Mistakes and Misguided Predictions*, St. Martin's Press, New York.

Newcomb, S. (1906), The Outlook for the Flying-Machine, 330–345, in *Side-Lights on Astronomy*, Harper & Brothers, New York and London.

O'Meara, S. J. (2001), The Demon Sprites of Mars, 102–104, *Sky & Telescope* 101(6), June.

O'Neill, G. K. (1977), *The High Frontier: Human Colonies in Space*, William Morrow and Company, New York.

Russell, D. A. (1983), Exponential Evolution: Implications for Intelligent Extraterrestrial Life, 95–103, *Adv. Space Res*. 3(9).

Schnabel, E. (1961), *Story for Icarus* (tr. from the German by J.J. Dunbar), Harcourt, Brace and Company, New York.

Stapledon, O. (1948), Interplanetary Man?, 213–233, *Journal of the British Interplanetary Society* 7(6), November.

Wells, H.G. (1938), *World Brain*, Methuen & Co. Limited, London.

Wright, Hamilton, Helen Hamilton, and S. Rapport, editors (1968), *To the Moon!*, Meredith Press, New York.

Zubrin, R. (with R. Wagner) (1996), *The Case for Mars*, Free Press, New York (Simon and Schuster, 1997).

Author

William I. McLaughlin worked in NASA programs for 30 years and is now writing a nonfiction work on space and civilization as well as a mystery novel.

MicroMaps Space Mission Analysis and Design

Ossama Abdelkhalik, Bassem Nairouz, Timothy Weaver,
Brett Newman

Abstract

Knowledge of airborne CO concentrations is critical for accurate scientific prediction of global-scale atmospheric behavior. MicroMaps is an existing NASA-owned gas filter radiometer instrument designed for space-based measurement of atmospheric CO vertical profiles. Due to programmatic changes, the instrument does not have access to the space environment and is in storage. MicroMaps hardware has significant potential for filling a critical scientific need, thus motivating concept studies for new and innovative scientific spaceflight missions that would leverage the MicroMaps heritage and investment, and contribute to new CO-distribution data. This paper describes engineering feasibility and trade studies for the NASA and Virginia Space Grant Consortium (VSGC) MicroMaps space mission. Conceptual studies encompass (1) overall mission analysis and synthesis methodology, (2) major subsystem studies and detailed requirements development for an orbital platform option consisting of a small, single-purpose spacecraft, (3) assessment of an orbital platform option consisting of the International Space Station, and (4) survey of potential launch opportunities for gaining access to orbit. Investigations are of a preliminary, first-order nature. Results and recommendations from these activities are envisioned to support future MicroMaps mission design decisions regarding program down-select options leading to more advanced and mature phases.

I. Introduction. This work describes the activities and accomplishments conducted under a contract with the Virginia Space Grant Consortium (VSGC), and indirectly with the National Aeronautics and Space Administration, Langley Research Center (NASA LaRC). The subject matter comprises engineering feasibility and trade studies for the NASA/VSGC MicroMaps Space Mission. MicroMaps is an existing NASA-owned gas filter radiometer instrument with 3° field of view designed for space-based nadir measurement of atmospheric carbon monoxide (CO) vertical profiles in the 4.67-μm wavelength [1]. The MicroMaps instrument was part of an overall scientific mission to be flown on the latter of the two Lewis and Clark spacecraft. Unfortunately, this mission was canceled, leaving the completed instrument without access to the space environment [2,3]. Currently, the instrument is in storage.

Atmospheric CO is a byproduct of natural and human surface activities, such as biomass burning or industrial processing. Trace CO gases can be transported by natural phenomena over great distances and altitudes, and can undergo mixing and chemical reaction with other natural atmospheric species such as oxygen-hydrogen radicals (OH). Reduction of upper-atmospheric OH content may adversely affect the natural removal of undesirable greenhouse gases such as methane (CH_4). Furthermore, CH_4 is tightly coupled to the dynamic life cycle of atmosphere ozone (O_3). These mechanisms may significantly influence the Earth's greenhouse effect and other global climate trends [4,5]. These large-scale dynamic processes are not yet well understood. Furthermore, scientific data, such as CO spatial and temporal distributions, to be used as inputs for global atmospheric and climate prediction models is severely lacking. A critical need exists for expanded atmospheric CO databases so that accurate scientific predictions can be undertaken and reported to appropriate governing political bodies making large-scale environmental policy and regulation.

MicroMaps hardware has great potential for filling this critical scientific need. This potential motivates concept studies for new and innovative scientific spaceflight missions that would leverage the MicroMaps heritage and investment, and would contribute to new CO distribution data for use in global-scale atmosphere and climate modeling and prediction. Conceptual studies encompass a broad spectrum of topics from launch options to platform design requirements in various subsystems. Results and recommendations from these studies will aid future MicroMaps Mission design decisions regarding policy and program down-select options leading to more advanced and mature phases. These studies include quantifying the merits and/or deficiencies of the options, in terms of facilitating scientific objectives, cost and complexity, reliability and robustness, and sizing and requirements.

Section II describes analysis and synthesis methodology for the MicroMaps Space Mission. Emphasis is given to development of the requirement flow-down relationships where science objectives, instrument specifications, environment factors, and resource reserves are used to formulate requirements on such aspects as orbit design, platform selection, and subsystem sizing and definition. Such relationships can be used to expose critical factors that impact the overall system design. Section III describes subsystem studies and detailed requirements development for the MicroMaps orbital platform option consisting of a small dedicated spacecraft with a single-purpose mission. Section IV describes key issues associated with the MicroMaps orbital platform option consisting of the International Space Station. Section V describes potential launch opportunities for gaining assess to orbit for the MicroMaps instrument, regardless of the orbital platform option chosen.

II. MISSION ANALYSIS AND SYNTHESIS.

Requirement Flow-Down Relationships

Development of the requirement flow-down relationships for the MicroMaps Space Mission are addressed to the extent that resources allow, and to the extent that available information allows, in this early stage of mission analysis and synthesis. In this process, objectives and constraints (such as science goals, instrument specifications, environment constraints, and resource reserves) are used to formulate requirements on such mission aspects as orbit design, platform selection, and subsystem sizing and definition. Formulation of the most significant mechanisms and mappings of objectives and constraints into requirements related to orbit design and selected subsystem definitions (for a small dedicated spacecraft platform) is emphasized here. Details of these relationships are presented in Section III. Such relationships can be used to expose critical factors that impact the overall system design. Associated insight may be more valuable for program decision making than specific subsystem definition and sizing studies.

Figure 1 illustrates the basic components involved in the flow-down relationships. The top level shows objectives and constraints from the factors of Science, Instrument, Environment, Launch, Resources, and Technology. These factors represent known objective and constraint information, and they serve as input for the formulation process of flow-down relationships. Other input factors can be incorporated into Figure 1 as they become known. The bottom level shows requirements on spacecraft subsystems related to Control, Propulsion, Electrical, Telemetry, and Camera. These components represent unknown requirement information that serves as output from the formulation process for flow-down relationships. Other output factors could be incorporated into Figure 1, if desired. An intermediate level associated with Orbit Geometry is also shown in Figure 1. Requirements for Orbit Geometry are influenced by

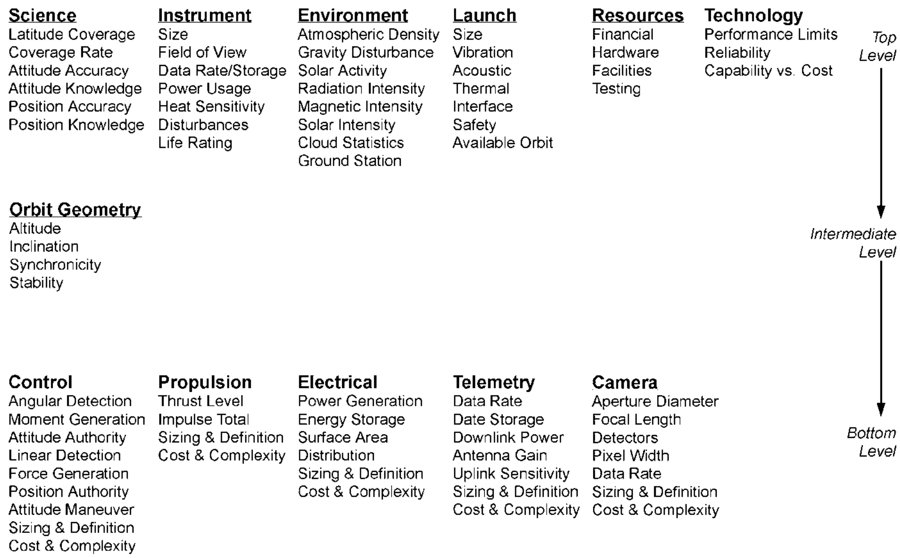


Figure 1. Flow-Down Relationship Components

many objectives and constraints. In turn, the orbit characteristics influence many spacecraft subsystem requirements. Because Orbit Geometry receives and transmits many key flow-down relationships, it is given special consideration.

As a starting point, all known information relating to mission objectives and constraints is collected. At this early stage of mission analysis and synthesis, the following partial list of information was collected. Science and Instrument data originates primarily from Reference 1 and discussions with Dr. Vickie Connors (NASA LaRC) and Dr. Henry Reichle (NASA LaRC Retired). Environment data originates from known facts documented in many texts such as [6,7]. With no specific launch opportunity identified at this time, the Small Spacecraft Technology Initiative (SSTI) design requirements for ascent conditions are interpreted as actual ascent conditions [1]. Note, access to the NASA Spaceflight Tracking and Data Network (STDN) is assumed here. This information is tentative and could evolve as the mission design proceeds.

Science

- Coverage of major CO sources and sinks: latitudes from 0 to beyond 75º
- Temporal resolution in CO: complete coverage every 30 days
- Spatial resolution in CO: 5º longitude by 5º latitude
- Pointing knowledge for data fidelity: ± 0.5º
- Positional knowledge for data fidelity: ± 25 km
- Pointing accuracy for data fidelity: ± 5º nadir (Ref. 4 lists ± 2.5º nadir)
- Pointing/positional update: 0.1 Hz

Instrument

- Life rating: 3 years
- Dimensions: 6 in. (15.2 cm) high, 8.25 in. (21 cm) wide, 13.75 in. (34.9 cm) deep
- Mass: 6.4 kg
- Inertias: I_{xx} = 0.049, I_{yy} = 0.047, I_{zz} = 0.030, I_{xy} I_{yz} I_{zx} 0 kg m^2
- Power consumption: 24 W
- Input voltages: +15, –15, +5 V
- Communication interface: RS 422 with XMODEM
- Data sampling microprocessors: Hitachi 6303
- Data processing microprocessor: RHC 3000
- Data rate: 288.7 bit/s uncompressed, 40 bit/s = 0.432 Mbyte/day compressed
- Data storage buffer: first-in-first-out (FIFO) circular 0.432 Mbyte (1 downlink per day)
- Field of view: ± 1.5º cone
- Circular footprint from low earth orbit: 25 km diameter
- Sensitive wavelength: 4.67 µm
- Detector temperature: 0 to 25ºC
- Chopper max momentum disturbance: 0.05 lbf ft s (0.068 Nms)
- Chopper inertia imbalance: ± 18 mg at 2 in. radius (5.1 cm)
- Chopper frequency: 2,000 rpm
- Calibration assembly max torque disturbance: 0.004 Nm every 2.5 s
- Calibration assembly frequency: 30 min cycle per day
- Radiation exposure:10 krads total, 30 MeV upset free, 100 MeV latchup free
- Magnetic dipole: 0.2 A-m^2 induced from 21 A-m^2 exposure, 0.01 A-m^2 residual

Environment

- Gravitational disturbances: J_2 oblate Earth model
- Atmospheric density:

Solar Max	**Solar Min**	**Orbit**
3.39×10^{-10} kg/m^3	1.69×10^{-10} kg/m^3	200 km
2.56×10^{-11} kg/m^3	1.28×10^{-11} kg/m^3	300 km
7.93×10^{-12} kg/m^3	2.36×10^{-12} kg/m^3	400 km
2.44×10^{-12} kg/m^3	3.26×10^{-13} kg/m^3	500 km
8.62×10^{-13} kg/m^3	5.81×10^{-14} kg/m^3	600 km
3.67×10^{-13} kg/m^3	1.61×10^{-14} kg/m^3	700 km

- Radiation intensity:

≈ 0 krads	every 10 years at 600 km, common shielding
3 krads	every 10 years at 800 km, common shielding
27 krads	every 10 years at 1,000 km, common shielding

- Magnetic intensity: 3×10^{-5} tesla for 200 to 1,000 km at magnetic equator
  6×10^{-5} tesla for 200 to 1,000 km at magnetic poles
- Solar intensity: 1,371 W/m^2 Earth orbit
- Cloud statistics: 30% of measurements randomly compromised
- Ground stations: NASA STDN S-band facilities (longitude, latitude)

Ascension Island (ACN)	345º 40' 22.57"	–7º 57' 17.37"
Bermuda (BDA)	295º 20' 31.94"	32º 21' 05.00"
Guam (GWM)	144º 44' 12.53"	13º 18' 38.25"
Kauai (HAW)	200º 20' 05.43"	22º 07' 34.46"
Merritt Island (MIL)	279º 18' 23.85"	28º 30' 29.79"
Ponce de Leon (PDL)	279º 05' 13.12"	29º 03' 59.93"
Santiago (AGO)	289º 20' 01.08"	–33º 09' 03.58"
Wallops Island (WAP)	284º 31' 25.90"	37º 55' 24.71"

Launch

- Dimensions: TBD
- Resonant frequencies: TBD
- Shock: SSTI design requirement
- Mass and inertias: TBD
- Vibration: SSTI design requirement
- Acoustic: SSTI design requirement
- Thermal: 10 to 24ºC prelaunch (long term), max 125ºC ascent (short term), max rarefied heating 400 BTU/hr ft^2 (1,260 W/m^2) (SSTI DR)
- Pressurization: sea level ambient to vacuum at rate of 0.35 psi/s (2.4 kPa/s) (SSTI DR)

Resources

- Financial: $2 to $4M (estimated)
- Facilities: TBD
- Hardware/software: TBD
- Testing: TBD

Technology

- Attitude/position sensing: TBD
- Impulse/momentum generation: TBD
- Computational capability: TBD
- Communication capability: TBD
- Moment/force generation: TBD
- Energy conversion efficiency: TBD
- Energy storage: TBD

First consider development of requirement flow-down relationships for Orbit Geometry. Only orbital altitude, inclination, synchronicity, and stability are considered in this analysis; and circular orbits are assumed exclusively. Figure 2 shows the most significant mechanisms affecting requirements for these orbital geometry characteristics. Objectives and constraints from Science, Instrument, Environment, and Launch are the most significant factors here. Science objectives associated with high-latitude coverage require orbit inclination angles greater than 75º. Environment constraints associated with drag from atmospheric density and complexities/expenses associated with shielding for Van Allen radiation

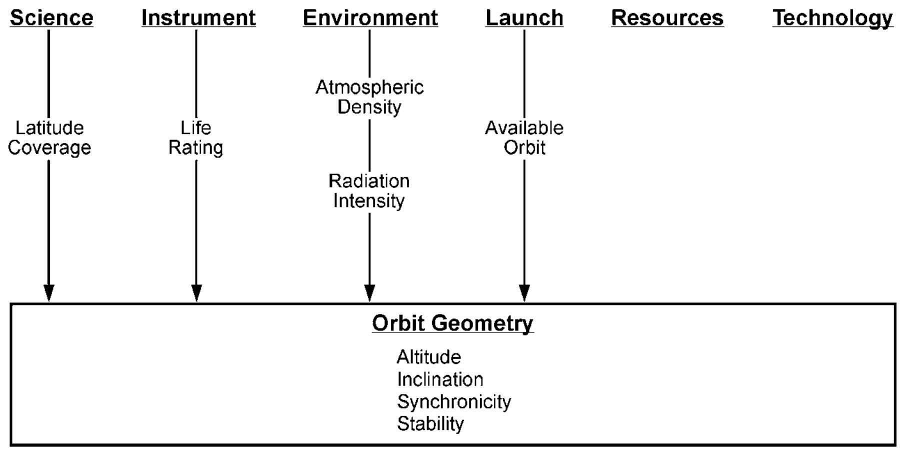


Figure 2. Orbit Geometry Flow-Down Relationships

require the orbit altitude to lie somewhere between approximately 200 and 1,000 km. No requirement seems to exist for temporal-spatial synchronous CO measurements. However, if one were imposed, specific inclination-altitude interdependency would be required. Launch constraints for each opportunity will also impose requirements on orbital inclination and altitude, which are left unspecified at this time. To maximize Science data collection, the Instrument life rating imposes an additional mild requirement for orbital stability to maintain minimum acceptable altitude (200 km) and inclination (75º) conditions for at least 3 years. For a given orbit initialization, inherent natural stability will most likely be sufficient, but could be supplemented with a propulsion system. Figure 2 illustrates these flow-down relationships. Resulting requirements are summarized below.

Orbit Geometry

- Inclination: greater than 75º
- Altitude: greater than 200 km, less than 1,000 km
- Stability: 200 km or higher altitude for 3 years; 75º or higher inclination for 3 years.
- Synchronicity: none or optional

Now consider development of requirement flow-down relationships for Control. Within the control subsystem, only requirements for angular detection, moment generation, attitude authority, linear detection, and position authority are considered here. Force generation requirements are addressed under Propulsion. Figure 3 shows the most significant mechanisms affecting requirements for these control system characteristics. Objectives and constraints from Science, Environment, Technology, Orbit Geometry, and several Subsystems (Telemetry and Other) are the most significant factors.

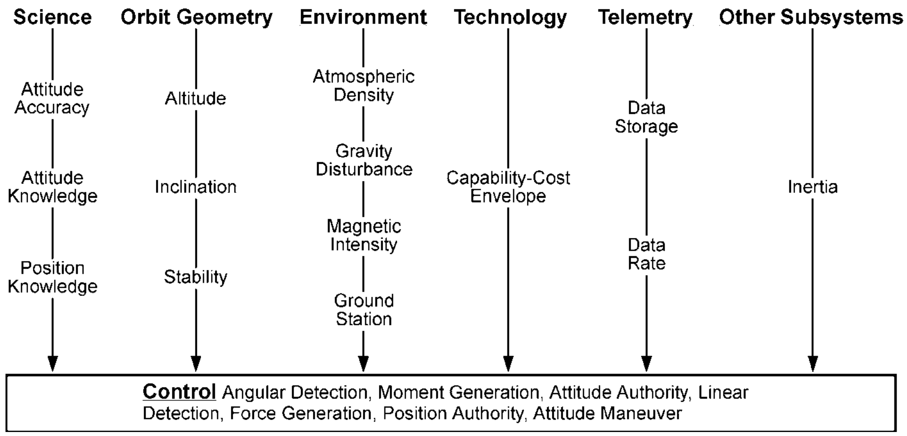


Figure 3. Control Flow-Down Relationships

Science objectives associated with CO measurement data fidelity and associated post-processing mandate knowledge of absolute instrument pointing and position to ±0.5º and ±25 km, respectively, while the accuracy of instrument pointing to a specified direction must be within ±5º. These objectives translate directly to requirements on angular detection, linear detection, and angular authority. These first two requirements (detection) impose conditions solely on the ability of sensor hardware to measure vehicle dynamic state information to sufficient precision (±0.5º and ±25 km). The latter requirement (authority) imposes a condition on the whole attitude control system (sensor, actuator, control logic, software, flight computer, etc.) to achieve and maintain a vehicle attitude state to within a specified tolerance (±5º). This requirement could impose further requirements such as a need for integral control logic to eliminate steady error in the presence of disturbances and sufficiently small nonlinear actuator traits such as dead zones to prevent transients outside the ±5º limit. Note, there is no direct requirement on position authority. However, orbit stability imposes a mild requirement for orbit inclination and altitude maintenance. Environmental constraints associated with atmospheric density and gravitational disturbances influencing the spacecraft trajectory (as well as moment disturbances from atmospheric, gravitational, and magnetic sources) require certain levels of force- and moment-generating capability from the control actuator hardware. Aerodynamic moment dominates below 400 km and requires a moment-generation capability of 5×10^{-3} Nm at 200 km and decreases to 8×10^{-5} Nm at 400 km, while magnetic moment dominates above 400 km, requiring a constant 8×10^{-5} Nm moment level. These requirements are influenced by orbit altitude and inclination, as indicated in Figure 3. Force-generation requirements are considered under Propulsion. Based on the Telemetry data rate and storage requirements, and the frequency of downlink opportunities to ground stations, which is influenced by Environment and Orbit Geometry factors (see Figure 3), a requirement to point periodically to ground stations may be needed. Any related requirements for attitude maneuvers are left as "To Be Determined." Note, inertias from the Other Subsystems (Structure) would strongly influence these requirements. Technology constraints impose additional requirements associated with the currently available capability-cost envelope, and these constraints are left unspecified at this time. All of these flow-down relationships are illustrated in Figure 3. Resulting requirements are summarized below.

Control

- Angular detection: ± 0.5º
- Moment generation: 5×10^{-3} Nm at 200 km (aerodynamic),
  8×10^{-5} Nm at 400 km (aerodynamic),
  8×10^{-5} Nm at 400 km and above (magnetic)
- Attitude authority: ± 5º (Ref. [1] lists ± 2.5º)
- Linear detection: ± 25 km
- Force generation: See Propulsion
- Position authority: 200 km or higher altitude for 3 years,
  75º or higher inclination for 3 years
- Attitude maneuver: to be determined

Next, consider development of requirement flow-down relationships for Propulsion. Within the propulsion subsystem, only requirements for thrust level and total impulse are considered here. Figure 4 shows the most significant mechanisms affecting requirements for these propulsion system characteristics. Objectives and constraints from Instrument, Environment, Control, Technology, and Orbit Geometry are the most significant factors here. The primary function of the propulsion system is to maintain orbital altitude and inclination stability over the mission life. Inherent natural stability will most likely be sufficient for most orbit initializations lying within requirements noted previously. How-

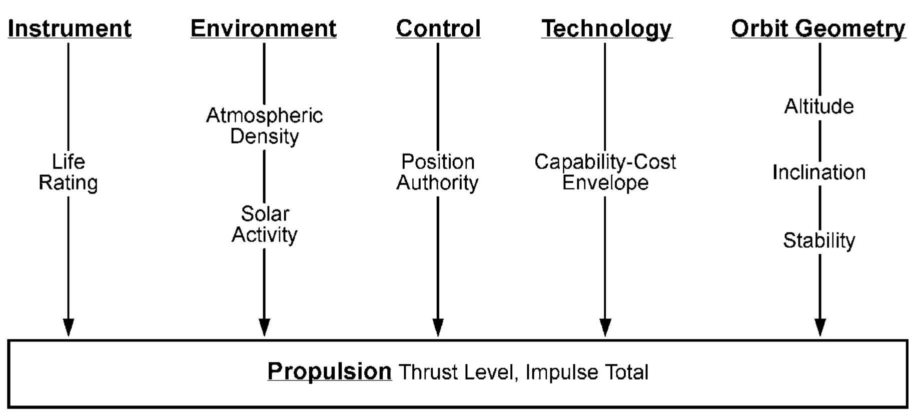


Figure 4. Propulsion Flow-Down Relationships

ever, for initial orbit altitudes below approximately 300 km, depending upon the solar cycle phasing during the mission, the orbital decay rate compromises the mission before the end of the 3-year instrument life. The orbital decay rate is computed by the method suggested in Reference 6 with an ample safety margin for uncertainty. Thus, a propulsion system is required for orbits below 300 km, and not required otherwise. A mission starting 3 to 5 years from the present should experience a period of decreasing solar activity, lessening the need for a propulsion system. At the minimum acceptable orbit altitude of 200 km, the drag force is projected to be 0.021 N assuming the worst-case atmospheric density, reference area of 1 m^2, and drag coefficient of 2. At 300 km, the drag force would be 0.0015 N. Thus, Environment and Orbit Geometry constraints require a thrust level of at least 0.021 N at 200 km and 0.0015 N at 300 km, respectively, to maintain altitude. For a 3-year mission, these conditions translate to total impulse requirements of at least 1,987 kNs (200 km) and 141.9 kNs (300 km). These requirements are influenced by orbit altitude, stability, atmospheric density, solar activity, and position authority, as indicated in Figure 4. Technology constraints impose additional requirements associated with the currently available capability-cost envelope, which are left unspecified at this time. All of these flow-down relationships are illustrated in Figure 4. Resulting requirements are summarized below.

Propulsion

- Thrust level: 0.021 N for 200 km, 0.0015 N for 300 km, 0 N above 300 km (min)
- Impulse total: 1,987 kNs for 200 km, 141.9 kNs for 300 km, 0 kNs above 300 km (min)

Next consider development of requirement flow-down relationships for Electrical. Within the electrical subsystem, only requirements for power generation, energy storage, and surface area are considered here. Figure 5 shows the most significant mechanisms affecting electrical system requirements. Objectives and constraints from Instrument, Environment, Technology, Orbit Geometry, and major power consumption Subsystems (including Control, Propulsion, Telemetry, and Others [Thermal]) are the most significant factors here.

Power generation is one of the most straightforward requirements to be considered. An estimate of the system power budget translates directly to power generation demands. Total power consumption of approximately 300 W (no energy storage) is projected with contributions to the total consisting of 24 W for Instrument, 60 W for Control, 100 W for Propulsion, 10 W for Telemetry, and 100 W for Thermal. Therefore, a minimum requirement for 300 W power generation (assuming no energy storage) due to the Instrument and Subsystems is established, as

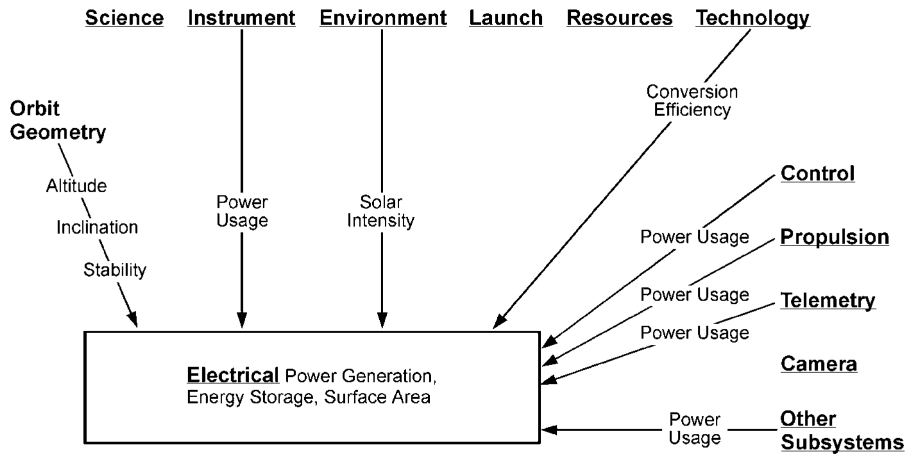


Figure 5. Electrical Flow-Down Relationships

indicated in Figure 5. In Figure 5, also note that Orbit Geometry factors can influence the power generation requirement by determining the level of Control-Propulsion power consumption needed to maintain orbital altitude.

There are two main options for generating this power: fuel cells or solar arrays. Because fuel cell consumables and complexity may drive the spacecraft mass and design beyond practical limits, this option is not considered further. Using spacecraft lighting estimates and solar energy conversion trends, requirement flow-down relationships for energy storage and surface area can be further established. Spacecraft passage within the Earth shadow mandates a need for energy storage. Assuming an asynchronous, high-inclination, low-altitude orbit, the percentage of time corresponding to darkness is a worst-case value of approximately 30%, or 0.45 hr for a 1.5-hr orbit period. Using a 10% nominal battery discharge depth, an energy storage requirement for 1,350 Wh is formulated. Note, an additional 135 W of power generation capability is required, leading to a revised requirement of 435 W (including energy storage). Finally, assuming solar conversion efficiency of approximately 25% (a Technology constraint), a requirement for 1.27 m^2 of surface area is established. Resulting requirements are summarized below, and Figure 5 shows the electrical flow-down relationships.

Electrical

- Power generation: at least 435 W
- Energy storage: at least 1,350 Wh
- Surface area: at least 1.27 m^2

Finally, consider development of requirement flow-down relationships for Telemetry. Only data rate, data storage, downlink power, and antenna gain are considered in this analysis. Figure 6 shows the most

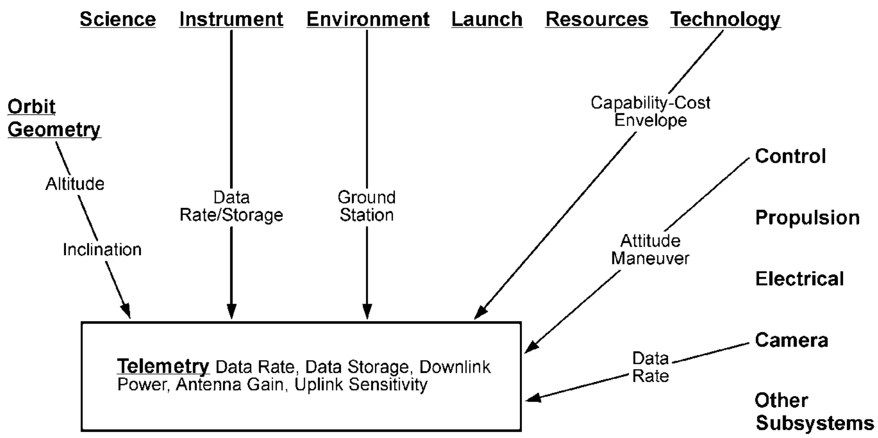


Figure 6. Telemetry Flow-Down Relationships

significant mechanisms affecting requirements for these telemetry system characteristics. Objectives and constraints from Instrument, Environment, Technology, Orbit Geometry, and Camera are the most significant factors here. Instrument data generation rate after compression is 40 bits/s = 0.432 Mbyte/day. Furthermore, the Instrument has a storage buffer capacity of 0.432 Mbyte. Thus, a minimum requirement for telemetry downlink data rate is 0.432 Mbyte/day (no camera). However, a maximum buffer content of only 25% at any given time is highly desirable to prevent scientific data loss if unexpected perturbations to the downlink were experienced. Thus, a more stringent requirement for data handling is 1.73 Mbyte/day (data rate) using the current storage buffer capacity. As discussed in Section III, if Earth image data of sufficient resolution must be downlinked also, the data rate and/or storage requirements could be much higher. Requirements for data handling with a camera are not considered here. Assuming a high-inclination low-altitude orbit with a period of 1.5 hr, and based on the NASA STDN S-band ground station geographic distribution and Earth spin rate, to ensure a downlink opportunity every 6 hr (0.25×24 hr) the downlink antenna beam width should be approximately 30º or larger. Assuming a conical beam shape, the corresponding antenna gain should be at least 60 = 35 dB (see Refs. [6],[8]). Using standard communication models for S-band telemetry[6],[8], the product of antenna gain with transmitter downlink power is estimated to be 230 W. Thus, a minimum requirement for downlink power is 4 W. Higher data rate or lower antenna gain and downlink power requirements could be accommodated with attitude maneuvers for ground-station pointing. Design freedoms of this type are not considered here. Figure 6 illustrates these flow-down relationships. Resulting requirements are summarized below.

Telemetry

- Data rate: 1.73 Mbyte/day (no camera)
- Data storage: 0.432 Mbyte
- Downlink power: at least 4 W
- Antenna gain: at least 60 = 35 dB

Orbit Selection

Orbital parameters will be calculated based on the scientific mission objectives. An algorithm has been developed to rapidly and roughly calculate a suitable orbit for a given set of objectives analytically. Only the J_2 gravitational perturbation is taken into account. A software tool that performs these calculations is available. Curves that illustrate the change in orbit altitude with variation of user objectives is presented. Two types of orbits are investigated. The first is the Earth-Sun synchronous orbit and the second is the Earth synchronous orbit. All orbits are assumed to be circular.

Science objectives require the instrument to collect the CO-distribution picture for the Earth at least once every season, or every 90 days. However, more frequent CO-distribution pictures for the Earth are certainly desirable. "Revisit Time" is defined as the period required to obtain a complete global measurement of the CO distribution. Beyond the revisit time, additional measurements begin to overlap earlier measurements. From the way the MicroMaps data will be processed, one can deduce that no need exists to measure every point on the globe; rather the Earth surface is divided into boxes, and the information for each box is considered uniform over the box. The size of a box is 5º longitude × 5º latitude.

The size of each box is equivalent to a rectangle with dimensions that vary according to the latitude of the box. At the equator, the box dimensions, X_{LA} and X_{LO}, are approximately $X_{LA} = X_{LO} = 556.6$ km. At latitude 80º, the rectangular dimensions are $X_{LA} = 556.6$ km and $X_{LO} = 96.6$ km.

Prediction of single measurement corruption due to cloud obscuration is not possible. However, statistical information can be used to calculate the number of measurements per box required such that at least three of them are cloud free.

An approximate estimate for cloud statistics is that 30% of all measurements will be randomly obscured. Assume for the moment that 10 measurements per box are required so that at least three of them will be cloud free. If it is sufficient to have a single path over each box in the revisit period, then the ground distance between tracks (i.e., the swath width) can be taken as 556 km at the equator. However, for more reliable performance, each box should be visited more than once in the revisit period. Assume that each box should be visited four times so that measurements can be obtained in any of the four visits. Thus, the swath width is around 120 km. Regarding the revisit time, a complete set of data will constitute a global picture for CO distribution, and this set of data is likely to be obtained with at least a seasonal temporal resolution. Reasonable orbits can be found with revisit time periods of around 20 days.

Earth-Sun Synchronous Orbit

Since the orbit is circular and Earth-Sun synchronous, defining the altitude will completely specify the orbit. The main idea is that an initial altitude is calculated based on a given swath width and revisit time taking into account only the condition of Sun synchronization. Then this initial altitude is corrected to the nearest altitude satisfying the Earth synchronization condition. A satellite flying at the new altitude will have a revisit time equal to that for the initial altitude but a slightly different swath width, as will be seen.

First, an initial altitude for the given swath width *(S)* and revisit time (m) are computed as follows. The distance on the ground between successive orbits *(D)* is related to *S* and *m* by

$$D = S \times m \qquad \text{Equation (1)}$$

The required change in longitude $\Delta\Phi$ on the equator between successive orbits is

$$\Delta\Phi = D/(R_e \cos\Lambda) \qquad \text{Equation (2)}$$

where R_e is the Earth radius and Λ is the latitude of the Earth location of interest. For a Sun synchronous orbit, Equation (2) can be expressed as

$$\Delta\Phi = 2\pi\ \tau\left(\frac{1}{\tau_e} - \frac{1}{\tau_{es}}\right) \qquad \text{Equation (3)}$$

where τ is the satellite orbital period, τ_e is the Earth period through one revolution, and t_{es} is the Earth orbital period around the Sun. For details on the preceding relationship derivations, refer to [8,9]. The required satellite orbit period τ can be calculated from Equation (3). τ is a function only of altitude, so the altitude *(H)* of the satellite can be computed from

$$\tau = 2\pi \sqrt{\frac{a^3}{\mu}} \qquad \text{Equation (4)}$$

$$H = \sqrt[3]{\left(\frac{\tau}{2\pi}\right)^2 \mu} - R_e \qquad \text{Equation (5)}$$

In Equations (4) and (5), μ is the Earth gravitational constant, and *a* is the orbit semi-major axis ($a = R_e + H$ for the assumptions made here).

Second, the condition of the Earth synchronous orbit is checked to determine the appropriate altitude. This will be done as follows. It can be proved that for Earth-Sun synchronous orbits,

$$2\pi n\sqrt{\frac{H^3}{\mu}}\,(1-\frac{\tau_e}{\tau_{es}}) = m\,\tau_e \qquad \text{Equation (6)}$$

where n is the total number of orbits before an identical ground track occurs, m is the revisit time, and H is the altitude. Note, variables m and n are integers. For the initial altitude, n is calculated. In general the calculated n will not be an integer, which means that this altitude does not satisfy the condition of Earth-Sun synchronous orbit. To enforce the Earth-Sun synchronous condition, the computed n is rounded to the nearest integer value and Equation (6) is used to compute the corresponding new altitude while holding the value of m constant. In this way, a value for altitude that satisfies the condition of Earth-Sun synchronous orbit is obtained and is the nearest one to the requirements of the user. The new altitude is usually very near to the initially calculated one, and resulting changes do not significantly impact mission objectives.

Given the satellite altitude, the swath width is calculated as follows. The orbit period is calculated from Equation (4). The change in longitude on the equator is calculated from Equation (3). Finally, the swath width is calculated from combining Equations (1) and (2), or

$$S = \Delta\Phi R_e \cos\Lambda / m \qquad \text{Equation (7)}$$

Several numerical calculations are done using a software tool to explore alternative altitudes for different values of swath width and revisit time. Figures 7 and 8 show some possible orbits for different mission objectives.

Ground Track Pattern

Results from the previous subsection showed there are some orbits that are suitable for the MicroMaps Mission for the given requirements.

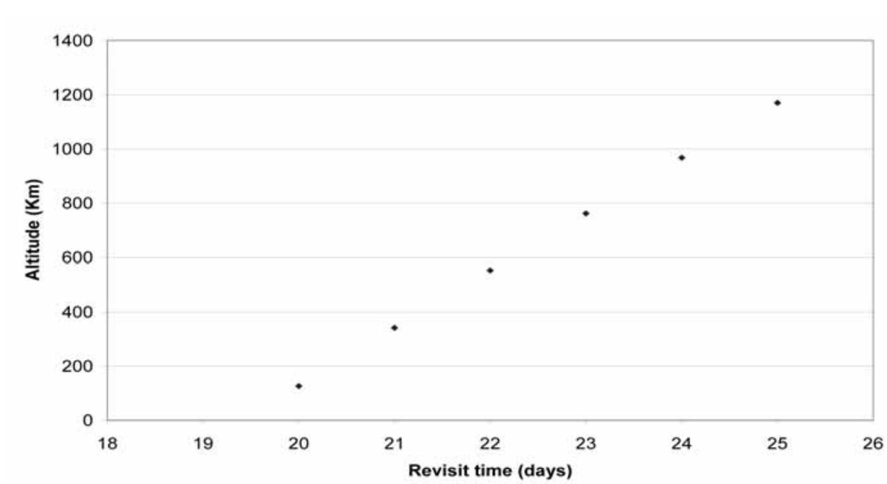


Figure 7. Altitude vs. Revisit Time Chart (Swath Width = 121 km)

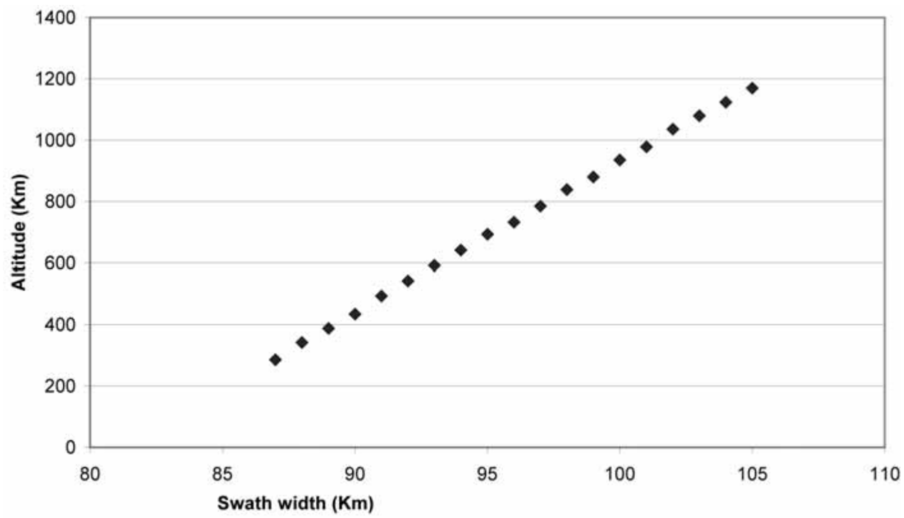


Figure 8. Altitude vs. Swath Width Chart (Revisit Time = 25 days)

In this subsection, the corresponding repass day pattern is determined. Repass day pattern refers to the number of days in which the satellite will pass over a certain area and the number of days in which the satellite will not pass over it. This information may be given, for example, in the format shown in Figures 7 and 8.

For a certain orbit, assume the satellite will pass over the area of interest in the first 2 days; then it will not pass over it in the next 3 days; then it will pass over it in the next 2 days, and so on. This information will be useful in selecting the most suitable orbit among the above possible orbits; for this information will determine the schedule by which the satellite will pass over certain ground stations or specific CO sources/sinks. Repass day pattern is a criterion to select among the possible orbits. The following discussion demonstrates the basic concept of how this criterion will be calculated.

A typical ground track is plotted in Figure 9. Assume that the satellite passes over track 1 and track 18 in the same day. The satellite passes over the tracks 2, 3, 4, 5, …, 17 in the following days. If the satellite passes over track 2 in the second day, and on track 3 in the third day, and so on, then the orbit of the satellite is called a minimum drift orbit. If the satellite passes over track 2 in the second day and on track 5 in the third day, or in any other order of tracks in the days subsequent to the first, then the orbit of the satellite is called a non–minimum drift orbit. For a minimum drift orbit, the repass day pattern is obvious. If, for example, the whole period of revisit time is 53 days, the satellite passes over a certain area every day for certain number of days and then does not pass over it for the rest of the period of revisit time. For a non–minimum drift orbit, some calculations must be done to determine the repass day pattern. These calculations are considered next.

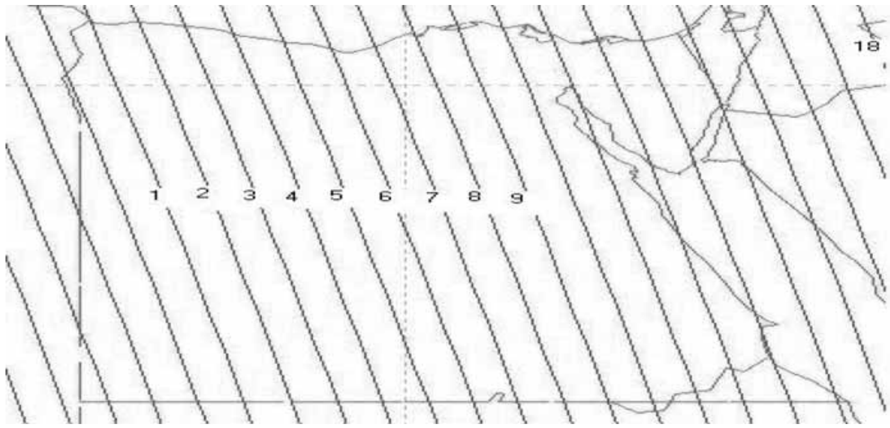


Figure 9. Typical Ground Track for Low Earth Orbit Satellite

Let the number of orbits that a satellite performs in one day be n. In general, n is not an integer. Since the satellite is orbiting in an Earth-Sun synchronous orbit, then the satellite revisits a certain point on the ground every certain number of days. Let it be M days. M is an integer. During these M days, the satellite performs N orbits. The condition of Sun synchronization implies that N is an integer also.

$$n = \frac{N}{M} \qquad \text{Equation (8)}$$

Now, assume that $n = j + x$, where j is an integer that represents the number of complete orbits performed in one day. Parameter x is a fraction less than 1; let it be K/M. This parameter represents the part of the orbit that is performed after j orbits are performed to complete one day of orbiting. As an example, if $N = 800$ and $M = 53$, then $n = 800/53 = 15 + 5/53$. Thus, $j = 15$ and $x = 5/53$. After a complete day of orbiting, the satellite performs a complete 15 orbits plus 5/53 of an additional orbit.

Now, return to Figure 9. The satellite passes on track 1 and on track 18 in the same day; it passes on track 1 in the first orbit and on track 18 in the second orbit of the same day. The distance on the ground between track 1 and track 18, call it S_t, is then the distance scanned in one orbit of the satellite motion. After one day the satellite does not pass on track 1 but on a track that is shifted from track 1. This shift is due to the fraction x of the orbit, that a satellite performs to complete one day of orbiting. If $x = 0$, the satellite repeats track 1 after one day.

Thus, after one day, the satellite passes on a track that is shifted a distance $x \times S_t$ on the ground from track 1. After two days the satellite passes on a track that is shifted a distance $2x \times S_t$ from track 1. After M days the satellite passes on a track that is shifted a distance $Mx \times S_t$ from track 1. Recall that $Mx = K$, which is an integer value.

Now, assume (without loss of generality) that the first track of the first day passes over the area under consideration. One can calculate the

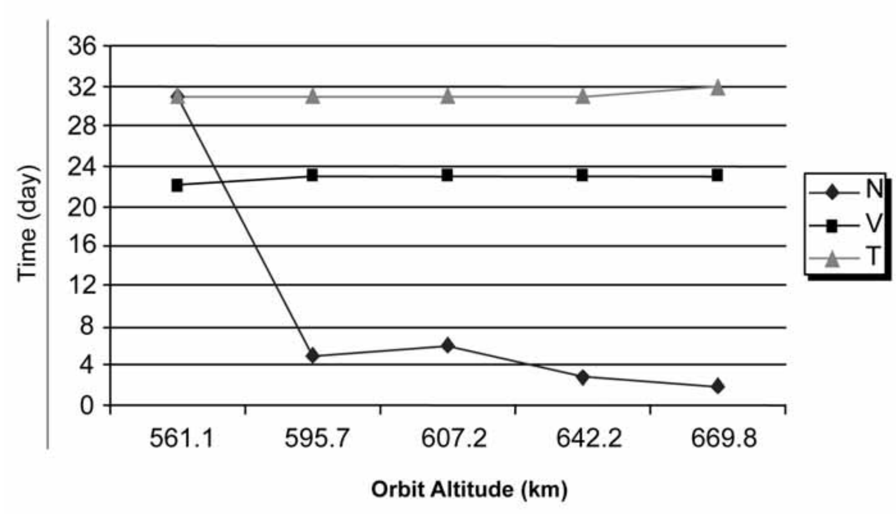


Figure 10. Ground Track Pattern for Various Altitudes

pattern of repass days as follows. Calculate the distance shift of the first track in the second day from the first track in the first day and check if it is within the band of that area or not, and repeat for the first track of the third, fourth, … day until the satellite completes the period of revisit time M. These calculations are programmed on a computer. As an example, a certain area of 1,000 km × 1,000 km is considered, and the results are plotted in Figure 10. In Figure 10, N is the maximum number of days of not visiting the area, V is the total number of days of visiting the area, T is the total number of days of not visiting the area, and $T + V$ is the revisit period.

Earth Synchronous Orbit

For the MicroMaps Mission, it is not scientifically required to fly the instrument in an Earth-Sun synchronized or Earth synchronized orbit. However it could be advantageous to fly the instrument in an Earth synchronous orbit for engineering purposes. In this case, the number of equations is less than the number of unknowns (for circular orbits) yielding many solutions for a single set of mission objectives. This fact is especially important considering that the launch conditions are not well defined at this time. In this subsection, a quick and rough approach is developed to calculate the possible orbits for a single set of objectives. The mathematical algorithm starts by specifying the requirement set for revisit time and swath width. The initial steps are to calculate D from Equation (1) and $\Delta\Phi$ from Equation (2).

Next compute n using Equation (9).

$$n\,\Delta\Phi = 2\pi\, m \qquad \text{Equation (9)}$$

Next correct n to the nearest integer. Recompute $\Delta\Phi$ using Equation (9), recalculate D with Equation (2), and recompute S using Equation

(1). Now select a value for orbit altitude H, and compute the orbit inclination i for the specified H using the following relationships.

$$\tau = 2\pi \sqrt{\frac{a^3}{\mu}} \quad \text{Equation (10)}$$

$$\Delta\Phi_1 = 2\pi\ \tau (\frac{1}{\tau_e}) \quad \text{Equation (11)}$$

$$\Delta\Phi_2 = \Delta\Phi - \Delta\Phi_1 \quad \text{Equation (12)}$$

$$\Omega = \Delta\Phi_2 / \tau \quad \text{Equation (13)}$$

$$\cos(i) = -\Omega \frac{2}{3} \frac{a^2}{2\pi / \tau} \frac{(1-e^2)}{R_e^2 J_2} \quad \text{Equation (14)}$$

For circular orbits, eccentricity e will equal zero. A family of solutions is obtained by using different values for H. This algorithm is implemented in software, and Figure 11 illustrates results for selected cases.

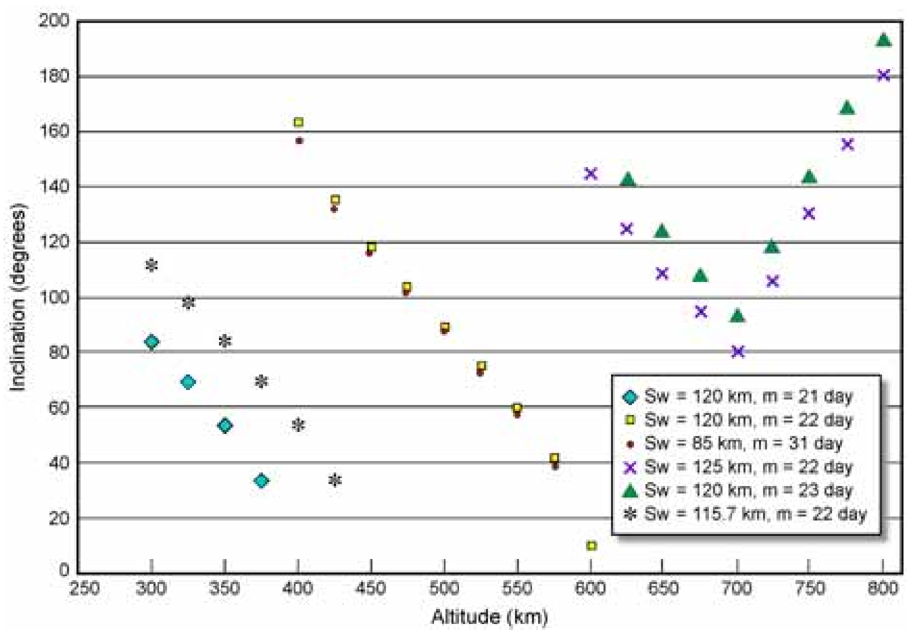


Figure 11. Candidate Earth Synchronous Circular Orbits for MicroMaps

III. DEDICATED SPACECRAFT—SUBSYSTEM STUDIES.

Attitude Sensing and Control

Two options are considered. The first is sending MicroMaps on a dedicated satellite to orbit. The second is sending a satellite with MicroMaps as the primary payload and a camera as a secondary payload.

There are four main sources of external environmental torques: solar pressure T_s, gravity gradient T_g, Earth magnetic T_m, and atmospheric T_a. Disturbance torques versus altitudes are plotted in Figure 12. It is clear that the magnetic torque values are to be used in the design process. All other torques, even if summed together, are negligible compared to the magnetic torque. This observation is true for orbits higher than 400 km. For lower orbits, which are less likely for adoption, the aerodynamic torque is the driving factor. Because the magnetic torque is relatively invariant to altitude (see Figure 12), and does not depend strongly on satellite configuration, altitude is not a significant driving factor in choosing an orbit for this mission, based on control disturbance rejection considerations. Due to the nature of the mission, a limited number of orbits might be available. The disturbance torque is not a limiting factor in the choice of the orbit. This observation must not be confused with the fact that the mission is already limited to available launches within a 3- to 5-year window referenced to the present time with prespecified orbits.

A suitable hardware configuration for a small, three-axis stabilized satellite uses three reaction wheels for attitude changing (a fourth redundant reaction wheel is also added), three torque rods or magnetic coils for momentum dumping from the reaction wheels, a three-axis magnetometer, and a pair of attitude sensors such as a Sun sensor and an Earth

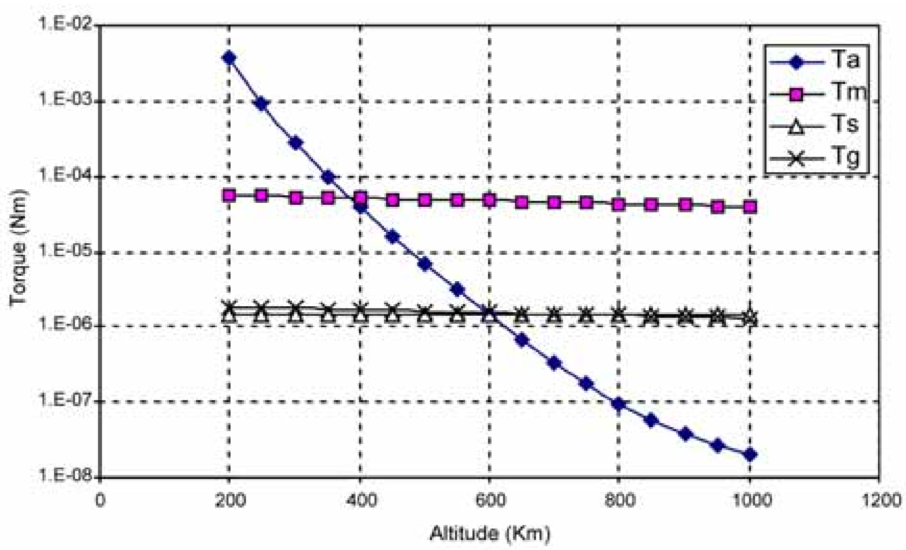


Figure 12. Disturbance Torques Affecting Satellite Pointing

sensor (another pair is also used for redundancy). If the rates are to be measured (rather than calculated using differentiation algorithms from position sensors), then rate gyros must also be added. An alternate configuration uses the torque rods for attitude control and discards the reaction wheels. This arrangement is a suitable configuration only for very small satellites in low Earth orbits. An algorithm must be developed for the attitude control system. This system might be adequate for the first mission option, i.e., the option to build a satellite dedicated entirely for MicroMaps without a camera and associated slew maneuvers.

Orbital Adjustment and Maintenance Propulsion

In order to extend the useful life of a satellite and avoid the problem of shifting orbits, mission designers can add a propulsion system to the satellite as a way of countering drag and gravitational perturbations. The total impulse, the force of drag over a known period of time, will determine how much fuel is needed for the mission, which is projected to last from 3 to 5 years. Table 1 shows the values of velocity and total impulse to maintain orbit at different altitudes, assuming an average cross-sectional area of 1 m^2, an average drag coefficient of 4, and an average density over solar and day cycles [10]. Table 1 also shows the orbit lifetime when no impulse is applied. Reference [6] is used to generate these lifetime estimates.

Table 1: Orbital Characteristics for MicroMaps

Altitude (km)	Velocity (km/s)	3-year Impulse to Maintain Orbit (kNs)	Lifetime with No Impulse (years)
200	7.784	2,912	0.25
300	7.726	217	0.73
400	7.669	57	2.98
500	7.613	15	10.45

The 200- and 300-km orbits will decay to reentry before the instrument life is up, while above 300 km the instrument life is reached before reentry. Higher altitude orbits would lessen the need for a propulsion system. The problem with increasing the altitude for the MicroMaps orbit is that at 800 km the craft begins to enter the Van Allen Belt, a region inside the Earth's magnetosphere where radiation is trapped. If the nominal orbit were to lie in this region, design complexity and cost to ensure avionics reliability and integrity would increase dramatically.

Table 2 compares electric rocket propulsion system options [11] on several criteria. According to the study, the best engine to use overall

Table 2: Electric Propulsion Performance Chart

Engine Type	Long Duration Low Thrust	Power Consump-tion	Fuel Require ments	Simplicity Reliability	Space Heritage
Resistojet	×	✓	×	✓	✓
Arcjet	×	✓	×	*	*
Ion Engine	✓	*	✓	×	✓
PPT	✓	✓	✓	*	*
MPD	✓	×	✓	✓	×
Hall Thruster	✓	*	✓	*	✓

✓ high rating, * medium rating, × low rating, MPD MagnetoPlasmaDynamics

is the pulse plasma thruster (PPT) followed by the Hall thruster. The critical design criterion is whether or not the engine produces long duration, low thrust. Long-duration, low-thrust settings were chosen as opposed to high-thrust burns for reboost because it is easier to perform satellite tracking if the thrust is just enough to cancel out atmospheric drag. Also, high-thrust reboosts could hinder observations made by the satellite.

A product search for companies providing electric rockets narrowed the engine choices down to two: the CU Aerospace PPT-8/9 and the Busek Tandem-200 [12,13]. Table 3 summarizes each engine's performance, fuel requirements, and cost. The first rocket engine, the Tandem-200, by Busek, is a Hall thruster. This system has a higher specific impulse, four times greater total thrust, and a larger range of operating power than the PPT-8/9 by CU Aerospace. However, the Tandem-200 has a higher power requirement, higher thruster mass, and need of a fuel tank and feed system. Also, the cost of xenon is high, while the fuel for the PPT-8/9 is included inside the engine at a lower expense. The cost of the Tandem-200 varies between $100k and $1M depending upon the amount of xenon fuel and system options, including control electronics and power supply. The cost of the PPT-8/9 is $30,000 per thruster, excluding electronics and power units. Individual thrusters are integrated into sets and used together. The electronics and power source can be operated in series, allowing the thrusters to use the same power and electronics for operation. Thruster sets are placed together and fired one at a time to avoid unbalanced thrust in the electrical rocket. The price of the Tandem-200 was quoted to be several hundred thousand dollars for the engines, electronics, and power source.

Table 3: Candidate Electric Rocket Engines

System	Tandem-200	PPT-8/9
Company	Busek	CU Aerospace
Type	Hall Thruster	Pulse Plasma Thruster
Thrust (mN)	12.4	2.9
Total Impulse (Ns)	15,680/kg of fuel	1,225
Specific Impulse (s)	1600	550
Power Consumption Nominal (W)	200	120
Power Consumption Range (W)	50–300	100–150
Mass (kg)	0.9	0.4 (with integrated fuel)
Mass Flow Rate Nominal (mg/s)	0.94	0.538
Fuel	Xenon	Teflon
Cost	$100,000 to $1M for integrated system	$30,000 per thruster maximum 8 = $240,000 $500,000 electronics

Due to the simplicity of integration and use, the PPT-8/9 is preferred over the Tandem-200. However, the deciding factor will be the orbit of the spacecraft because required thrust decreases rapidly with increasing altitude. Table 4 compares the amount of fuel for the Tandem-200 and the number of PPT-8/9 units needed for a 3-year mission to maintain constant altitude, even though above 300 km orbital decay can be tolerated and no propulsion system is required. As seen here, the PPT-8/9 becomes impractical under an altitude of 400 km, based on the required number of units. At those altitudes, only the Tandem-200 could produce sufficient total impulse over the duration of the mission. Furthermore, at extremely low altitudes (200 km), the required fuel mass makes even the Tandem-200 impractical.

Table 4: Constant Altitude Demands for the Tandem-200 and PPT-8/9 Engines

Altitude (km)	3-Year Impulse (kNs)	Tandem-200 Fuel Mass (kg)	Number of PPT-8/9 Units
200	2,912	185	2,377
300	217	13.8	177
400	57	3.6	46
500	15	1.0	12

Electrical Power Generation and Storage

In all likelihood, power generation for the MicroMaps small dedicated space platform will be implemented with solar energy conversion. The solar photovoltaic cells for the MicroMaps satellite will be placed along the surface of the satellite.

Consequently, the satellite will most likely be hexagonal-shaped. Because the satellite must keep the CO measurement instrument facing the Earth in the nadir direction at all times, the satellite will shift its position relative to the Sun. Photovoltaic cells should be perpendicular to the source of light for the maximum amount of light to be absorbed and converted to electricity. This is why a hexagonal shape is expected for the satellite's structure. As the satellite shifts position in relation to the Sun, there will always be enough photovoltaic cells in the proper position to produce the necessary amount of electricity.

The amount of power needed for MicroMaps is estimated at 300 W, including a minimum of 100 W for the propulsion system. Also, the solar cells must generate an additional 135 W for charging the nickel-cadmium (NiCd) batteries used during periods without sunlight. A high conversion rate for a satellite solar cell is approximately 26.5% [14], and the intensity of sunlight at Earth's orbit is 1,371 W/m^2. The area of solar panels needed for powering the spacecraft is thus about 1.2 m^2. This value would have to be increased further to account for degradation of solar cell performance over the mission lifetime. Of course, this is the area of solar cells exposed to direct sunlight at a perpendicular angle to the Sun. Thus, the actual surface area must be larger than 1.2 m^2. Surface area of the satellite is difficult to estimate at this stage without further studies of satellite configuration design and orientation in relation to the Sun. There are also cost and mass restraints to deal with. NiCd batteries have a specific energy, the amount of energy that is stored per unit mass, of 219 Wh/kg. In order to provide 300 W for over seven hours of darkness, there must be 10 kg of NiCd batteries onboard the satellite. The estimated mass of the satellite is 50 kg. Battery mass alone would be 20% of the satellite's mass. Also, fully integrated photovoltaic cells are estimated to cost $700/W [14]. For MicroMaps, the cost would be over $300,000. The dollar amount will be several times larger since there will be more than 1.2 m^2 of satellite surface area to be covered in photovoltaic cells. Another option is to place a smaller area of solar cells on movable panels. However, this would cause complexities in attitude vibrations through the satellite, creating jitter motions in the instrument observations and corrupting the measurement data. In order to avoid this, the control system would have to be more complex, adding cost to the mission. Finally, articulating panels would increase satellite drag, which would increase the fuel, mass, and power needed for the propulsion system in low-altitude orbits.

In order to reduce the power requirements for the satellite, system designers could leave the propulsion system off during days with the shortest periods of sunlight, or during times without any sunlight. Just keeping the propulsion system off during periods without sunlight could reduce the NiCd battery mass to 6.6 kg. If the satellite is in sunlight at least 80% of the day before the propulsion system is operated, and if the propulsion system is operated only during the day, the effective area needed for the photovoltaic cells would be only 0.95 m^2. Power management will be very important for the mission to reduce cost, mass, and system complexity.

Vehicle-Ground Communication and Telemetry

Communication between the proposed spacecraft and the ground stations will at a minimum consist of MicroMaps Data Downlink, Camera Data Downlink, Spacecraft Telemetry Downlink, and Ground Command Uplink. Sizing the communication system will affect the overall sizing of the satellite and will also be affected by some key system drivers. This subsection briefly presents some of the system drivers that affect communication system size.

The MicroMaps data is generated and compressed inside the instrument itself. The data rate coming from the instrument is 0.432 Mbyte/day, which is 40 bits/s. Assume the telemetry data rate is 30 kbits/s, the command data rate is 3 kbits/s, and the camera imaging data rate is 25 Mbits/s. Another assumption made here is that one ground station is available to receive MicroMaps data. As the satellite flies around the Earth, MicroMaps always collects data and stores it in mass memory on board. The satellite downlinks data each time the ground station is available. The satellite collects a complete set of data for the whole Earth in m days, the revisit time period. The satellite is required to downlink the complete set of data for the whole Earth in m days also. Total MicroMaps data stored in m days equals $3.3m$ Mbits.

A gain-shaped antenna (a directional antenna with a tailored beam boresite power level) can be used to cover the whole horizon under the satellite with an elevation angle of 5º. Four different orbits are investigated. For each orbit, the beam width of the antenna is calculated, and the time available for downlink of MicroMaps data, T, is also calculated. From that, the bit rate for downlink is calculated for MicroMaps data. Results are listed in Table 5. From the computed values of the downlink bit rate, one can use a single set of transmitter and antenna components for both telemetry and MicroMaps data. Either UHF or S-band frequencies can be used. Both frequency bands may be used for redundancy. To estimate the required memory for MicroMaps data, notice that the maximum period for the satellite in which it cannot see the ground station is 2 days; the required mass memory is thus $2 \times 24 \times 3{,}600 \times 40 =$

Table 5: MicroMaps Data Downlink Bit Rate

Altitude (km)	*m* (day)	Beam Width (deg)	*T* (min)	Data Rate (kbits/s)
461	20	136	93.8	12.0
542	25	133	174.2	8.1
676	23	129	195.4	6.6
776	26	125	257.4	5.7

6.6 Mbits. The downlink bit rate also affects the power consumption of the communication system.

Earth Observation Camera

A parametric study can be conducted to uncover effects from some of the design parameters. Figure 13 illustrates the variation of the imaging data rate vs. the resolution for different values of the camera swath width. This chart is useful for deciding whether to look for global image coverage of the Earth or accept a certain coverage percentage. Varying the swath of the camera for a certain resolution affects not only the data rate, but also the size of the camera itself, and of course camera cost will be affected. One can get a sense of how the camera size will increase by calculating the required number of detectors. Figure 14 illustrates the variation of the number of detectors for one color by varying the camera swath and the resolution.

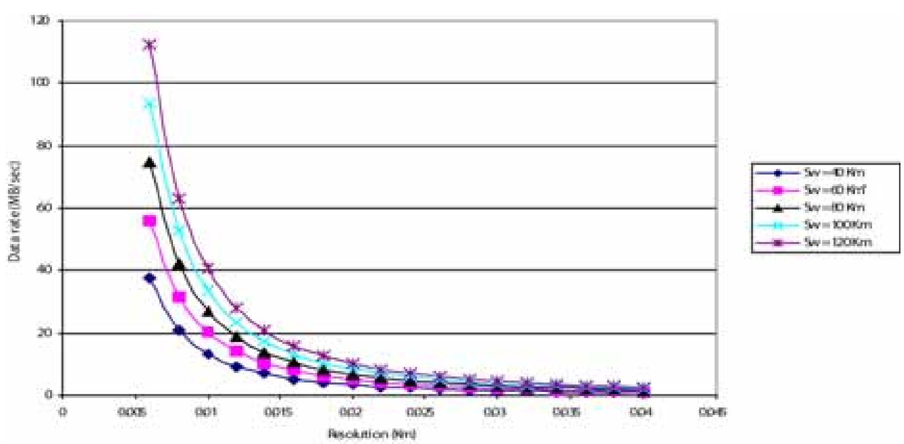


Figure 13. Data Rate vs. Resolution (Constant Swath Width)

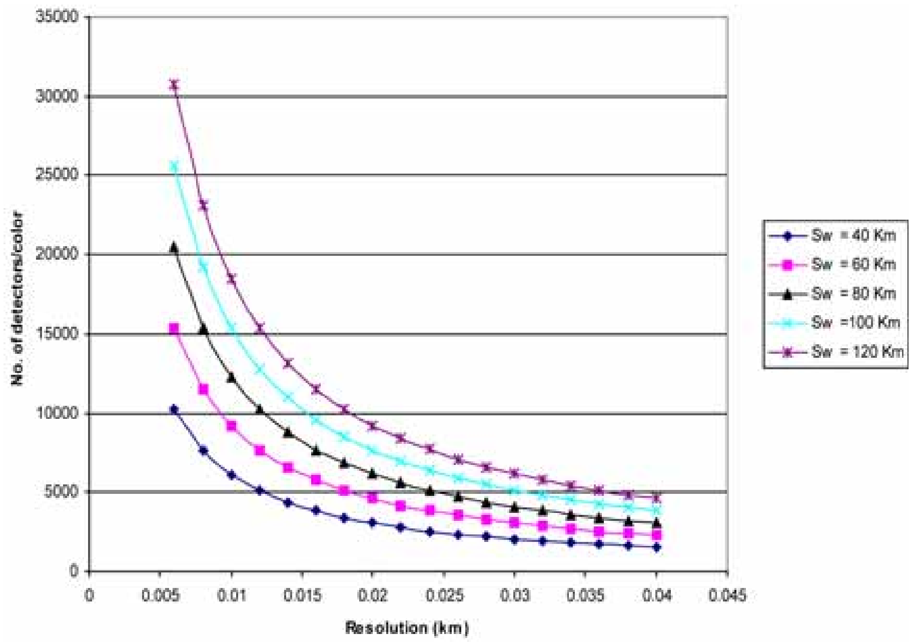


Figure 14. Number of Detectors vs. Resolution (Constant Swath Width)

IV. INTERNATIONAL SPACE STATION—KEY ISSUES.

Earth Surface Coverage

The International Space Station orbit covers nearly all continental land mass in the southern hemisphere, and all land south of latitude 51.6º in the northern hemisphere. Unfortunately, all of Russia, most of Canada, northern parts of Europe, northern parts of Alaska, and all of the polar regions are not covered. This coverage gap means that if a significant CO source or sink is positioned in any of these regions, its characteristics will not be directly detected by MicroMaps. Yet, because of global-scale atmospheric air motion, weather, wind, etc., indirect effects from this CO source or sink would be detected elsewhere in the covered regions. Incomplete Earth surface coverage can lead to uncertainty in global atmospheric models and climate projections. Furthermore, geographic regions that would be covered by the Space Station orbit were previously studied during Space Shuttle Measurement of Air Pollution from Satellites (MAPS) missions [2,3]. A complete global map of CO distribution will not be possible, only a partial map between latitudes 51.6º north and south.

As documented in Section II, scientific objectives highly emphasized a global CO distribution measurement as opposed to just covering lower latitudes. Mission scientists have underscored this objective on several occasions. Therefore, based on orbit suitability and associated Earth surface coverage, flying MicroMaps on the International Space Station platform is not recommended.

Attitude and Vibration Transients

One of the major roles the International Space Station will fulfill is to serve as a multi-user platform for long-term atmospheric, ocean, land, and astronomical scientific investigation. Additionally, exploitation of the microgravity and/or vacuum space environment for scientific and commercial purposes is expected. Unfortunately, the Space Station will be a dynamic platform that experiences attitude and vibrational motion transients originating from a multitude of operational constraints that may corrupt or compromise the user requirements depending upon the application. Evaluating the Space Station attitude and vibrational dynamic characteristics against the MicroMaps requirements will therefore be addressed in this subsection.

Figure 15 shows the fully operational Space Station configuration [15]. The vehicle is characterized by a long, slender truss structure serving as a backbone with numerous facilities, modules, and solar arrays attached along its length. The span of this truss structure is approximately 108 m while the transverse attachments are about 80 m long. Users will attach hardware to available pallets that are located along the truss structure. These pallets are oriented in both the +Z and –Z directions, and can be located a significant distance from the vehicle mass center.

The Space Station will be flown in several operational modes with varying orientations; the solar panels will be actively articulated for optimum solar tracking; robotic arms and track vehicles will be per-

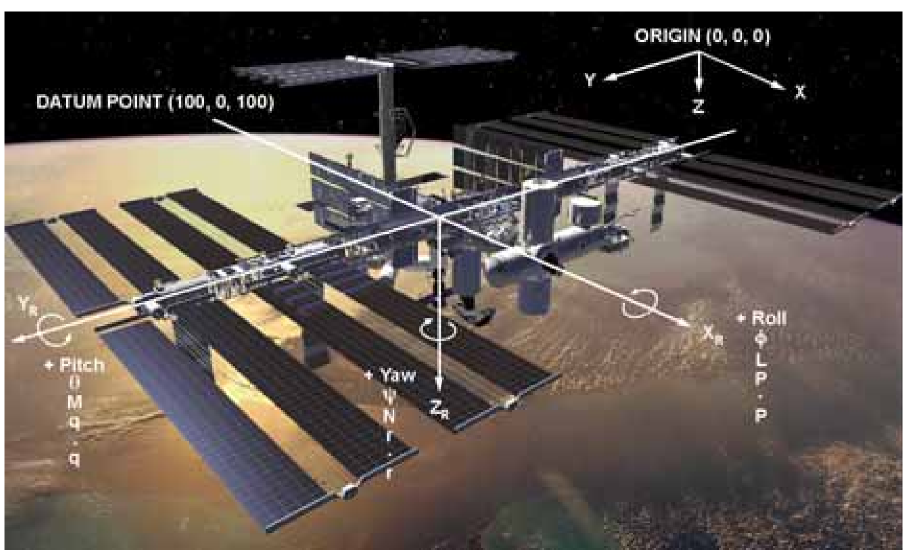


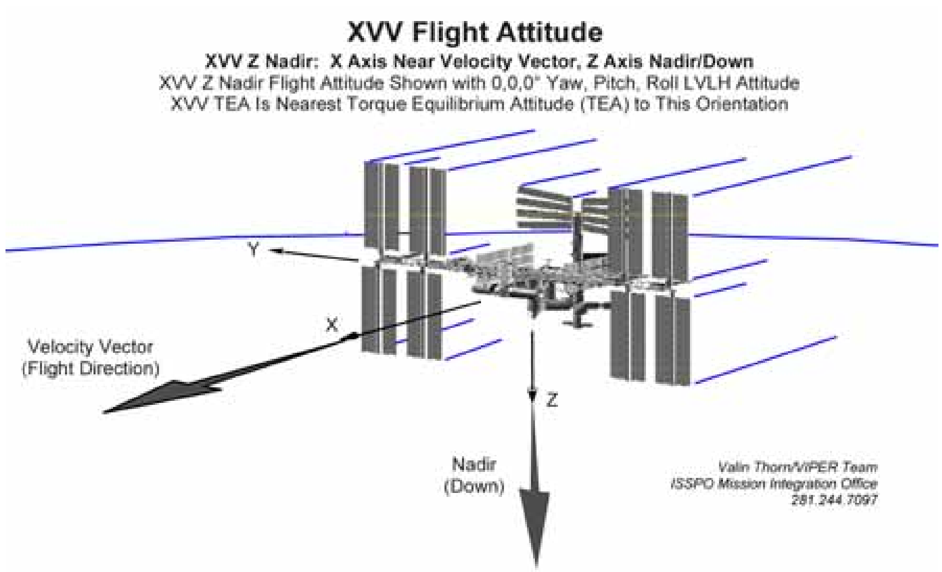
Figure 15. International Space Station Configuration

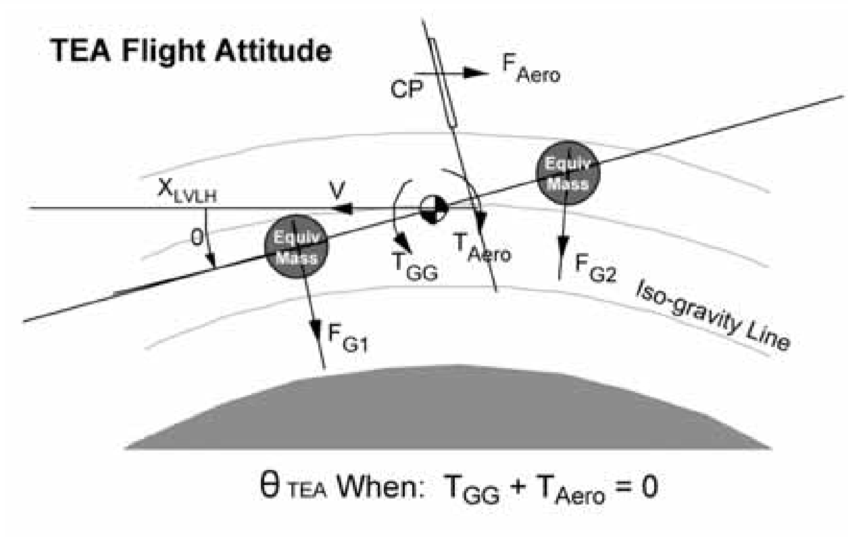
forming construction and maintenance duties; service and supply vehicles will be docking frequently; periodic orbit boost maneuvers will be executed; angular momentum control devices will be in operation; and the vehicle is a large, lightweight, flexible structure susceptible to disturbance propagation. In addition, the current configuration will undergo many on-orbit modifications over the next several years before achieving the fully operational configuration of Figure 15. These configuration modifications encompass large changes in inertia and attitude control capability. In summary, the International Space Station has a potential for exhibiting significant attitude and vibration transients that could compromise the scientific integrity of data collected by MicroMaps.

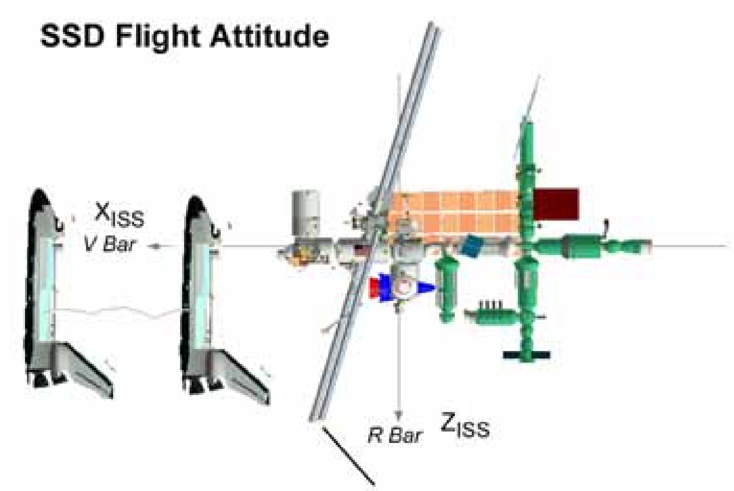
During the next several years, the Space Station will be operated in various flight attitude modes [16]. These modes are summarized in Table 6 and Figure 16. Mode XVV is a flight attitude where the X axis is near the Velocity Vector. This mode minimizes aerodynamic drag and is used to achieve microgravity conditions and for orbit boost maneuvers. Mode XPOP is a flight attitude where the principal X axis is Perpendicular to the Orbit Plane. This mode simultaneously provides for optimum solar collection and power generation and minimizes the gravity gradient torque. Mode TEA is a flight condition where environmental torques are in approximate balance, i.e., Torque Equilibrium Attitude. This mode balances aerodynamic torque and gravity gradient torque and is

Table 6: Space Station Flight Attitude Modes

Mode	Description	Yaw—Z (deg)	Pitch—Y (deg)	Roll—X (deg)
XVV	X Axis near Velocity Vector Min Aero Drag, Microgravity, Orbit Boost	+ 15.0 – 15.0	+ 15.0 – 20.0	+ 15.0 – 20.0
XPOP	Xp Axis Perpendicular to Orbit Plane Min Gravity Torque, Max Solar Collection	+ 10.0 – 10.0	+180.0 –180.0	+ 10.0 – 10.0
TEA	Torque Equilibrium Attitude Aero-Gravity Torque Balance, Microgravity	+ 13.1 – 12.0	+ 2.8 – 19.1	+ 1.2 – 2.6
SSD	Space Shuttle Docking Shuttle Docking Procedures, Similar to XVV	+ 0.0 – 0.0	+ 0.0 – 0.0	+ 0.0 – 0.0
SVD	Service Vehicle Docking Service Docking Procedures, Similar to TEA	+ 0.0 – 0.0	+ 15.0 – 20.0	+ 15.0 – 20.0


XVV Flight Attitude
XVV Z Nadir: X Axis Near Velocity Vector, Z Axis Nadir/Down
XVV Z Nadir Flight Attitude Shown with 0,0,0° Yaw, Pitch, Roll LVLH Attitude
XVV TEA Is Nearest Torque Equilibrium Attitude (TEA) to This Orientation
Y
X
Velocity Vector
(Flight Direction)
Z
Nadir
(Down)
Valin Thorn/VIPER Team
ISSPO Mission Integration Office
281.244.7097


TEA Flight Attitude
CP
F Aero
X LVLH
V
θ
Equiv Mass
Equiv Mass
T GG
T Aero
F G2
F G1
Iso-gravity Line
θ TEA When: T GG + T Aero = 0


SSD Flight Attitude
X ISS
V Bar
R Bar
Z ISS

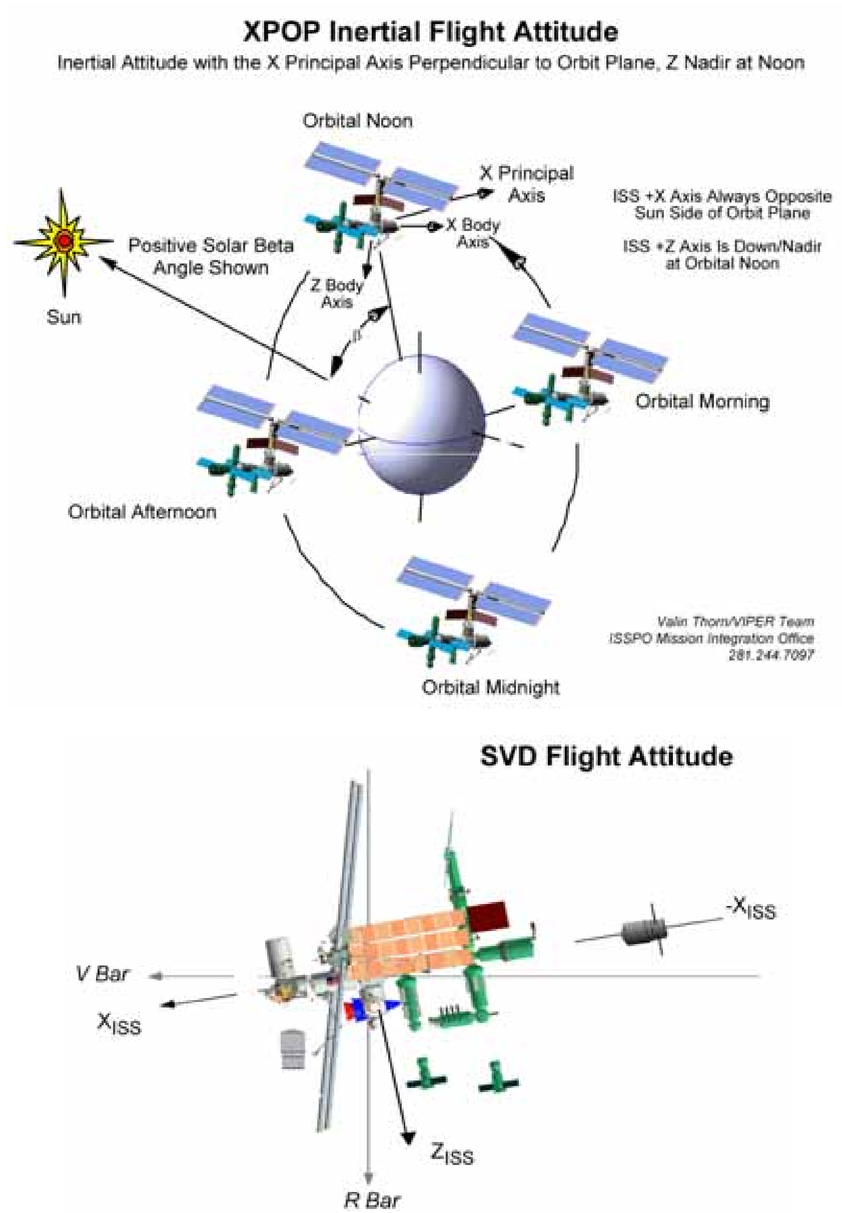


Figure 16. Space Station Flight Attitude Modes.

used to achieve microgravity conditions. Mode SSD is a flight attitude used for Space Shuttle Docking, while Mode SVD is a flight attitude used for Service Vehicle Docking.

The yaw-pitch-roll attitudes (referenced to a local level frame) listed in Table 6 indicate nominal operating ranges expected over the next several years as the platform configuration undergoes modification and expansion. For example, in the XVV mode, the pitch angle will lie somewhere between +15 and –20°, such as +10°, for an extended period of time on the scale of many months. After this extended period, the platform is modified and a new pitch angle results, such as –5°. This new angle remains until the next modification. The ranges listed in Table 6 should not be interpreted as bounding continuous transients occurring on an hourly-daily scale. The one exception to this interpretation is pitch angle range for the XPOP mode. The XPOP mode holds an inertial orientation, thus yielding a ±180° pitch angle variation during each orbit. Returning to the XVV mode +10° pitch angle example, the Space Station will not hold this precise attitude in the short term either. Due to previously listed disturbances, actuator and/or sensor hardware limitations, control performance, and structural dynamics, small pitch transients will continuously occur around the nominal value. References [17] and [18] specify maximum transients to be ±3.5 to ±5°, depending upon platform configuration. Although the Space Station is a large, flexible structure, the attitude control and disturbance bandwidths are well below the structural dynamic resonant frequencies. As shown in [19], pointing disturbances due to structural vibrations are projected to be minor.

Now interpret the information in Table 6 with respect to the MicroMaps Space Mission requirements. MicroMaps would be mounted to the Space Station such that the instrument would look along the +Z axis (see Figure 15) since +Z is approximately oriented along nadir in most cases. With this arrangement, only the pitch-roll variations in Table 6 are of concern. Yaw would not affect nadir viewing or measurements. Optimum pitch-roll values are 0 and 0°. Recall the nadir pointing accuracy requirements for MicroMaps from Section II: ±5° (Ref. [1] lists ±2.5°). The pitch-roll variations listed in Table 6 suggest that atmospheric CO measurement data would be severely compromised due to Space Station attitude variations if MicroMaps was rigidly fixed to the platform structure. Theoretically, the Space Station configuration will stabilize after operational capabilities are achieved, or funding evaporates, and the attitude variations listed in Table 6 will become fixed biases that could be corrected with counter-bias mounting. However, the platform attitude control performance of ±3.5 to ±5° would still significantly contaminate the scientific data. Even in this scenario, angular variations will occur when the flight attitude mode is switched from the various options. To fully resolve these issues, MicroMaps would have to be mounted on an active pointing and/or tracking system. A pointing system with ±30° azimuth-elevation range could correct for the Table 6 and platform control performance variations in all flight attitude modes except XPOP. The XPOP mode demands a full 360° pointing capability (only 180° of usable pointing exists due to pallet viewing blockage).

Scientific objectives emphasized precision nadir measurements of atmospheric CO vertical profiles. Mission scientists have underscored this objective on several occasions. Therefore, based on platform suitability and associated attitude transient motions, flying MicroMaps on the International Space Station platform is not recommended. It would require an expensive and complex active pointing system.

V. LAUNCH OPPORTUNITIES.

Future Space Missions

Due to budget constraints, gaining access to the space environment for MicroMaps, regardless of the platform option selected, must be achieved by flying the instrument as a secondary payload on board an already scheduled flight. Domestic government and/or commercial launches to low Earth orbit that have appropriate schedules and satisfy the MicroMaps Mission requirements are highly desirable and sought after. The initial task is simply to collect a database of future space missions form which appropriate launch opportunities can be identified. A modern, computerized search strategy easily identified numerous lists of scheduled launches extending approximately 1 year from the present. However, all of these launches are inconsistent with a MicroMaps Mission start date of 3–5 years from the present. Consequently, a more refined search strategy was required. This strategy concentrated on identifying

1. multi-decade space missions requiring multiple launches,
2. studies addressing demands for future space launch infrastructures, and
3. individual, single-launch space missions one by one, followed by focused searches on the identified items.

This approach was successful in finding a large database of candidate missions.

Results of this effort show that if the International Space Station platform option was selected for MicroMaps, numerous Space Shuttle launch opportunities for access to low Earth orbit are available [20]. Reference [21] lists many future, mostly domestic governmental and commercial science missions that may offer suitable launch opportunities for a small, dedicated spacecraft serving as the platform for MicroMaps. Finally, several future space missions currently under development with launch dates well beyond the current time frame were identified on a case-by-case basis and may also be suitable for the MicroMaps Mission.

Candidate Launch Assessments

Now that a healthy database of future space missions is available, the next task is to extract a subset of associated launch opportunities to low Earth orbit that are approximately in alignment with the MicroMaps

Space Mission requirements. Top-level (general) criteria, such as launch year and orbit type, can be used to narrow the database down to several competing launches. Then, a more refined assessment using additional lower-level (specific) criteria (such as launch date, orbital parameters, launch vehicle constraints, cost, and cooperation) can be used to identify the optimum launch opportunity. This two-stage approach is used here. Before conducting the analysis, additional comments on this process are offered.

There are several considerations that must be made in order to launch a satellite as a secondary payload. The first considerations are the maximum payload capacity of the launcher, and the mass of the primary payload and any additional secondary payloads previously scheduled. Next, the altitude must be approximately the same as desired. Now, this parameter is flexible since most geostationary-orbiting satellites start off in low Earth orbit and are taken to geostationary Earth orbit by a separate booster rocket. However, inclination is different in that this is an inflexible parameter. Inclination changes are more difficult and expensive to make than are altitude changes. Because satellites in geostationary Earth orbit have a low inclination, taking MicroMaps up as a secondary payload on a geostationary satellite launch is not appropriate to mission objectives. Of course, this is a conservative assessment of possible launch windows for MicroMaps. With an electric propulsion system, it would be possible to make changes in altitude and inclination. While it is desirable to find a launch with the correct orbit and inclination, there is some degree of freedom offered by an electric propulsion system.

Now return to the assessment task. In this analysis, the small, dedicated spacecraft platform option for MicroMaps will be assumed. Further, recall the requirement for high orbital inclination. With this information, all Space Shuttle flights to the International Space Station are eliminated from consideration. Further, scientific Space Shuttle flights are also eliminated because they do not offer high orbital inclination. Now, recall that the intended launch date for the MicroMaps Mission is 3–5 years from the present. Thus, only flights with a launch date lying approximately within the 2006–2008 window are retained. Additionally, only high-inclination, low-altitude flights are retained. The remaining launch opportunities that are potential contenders for the MicroMaps Mission are listed in Table 7. The Table 7 launch opportunities were generated in this initial assessment stage.

Table 7 lists the high-potential launch opportunities for MicroMaps with additional detail information on each mission, including orbital inclination, orbital altitude, mass constraints from the launch vehicle lift performance minus primary payload mass, size constraints from the launch vehicle fairing dimensions minus primary payload size, ascent constraints from the launch vehicle vibrational environment, cost sharing, and willingness for cooperation, where available. Data that is either not available, could not be found, or that must be collected in further studies, is designated “to be determined” (TBD). Because of the incomplete data, a final selection for the MicroMaps launch opportunity cannot be made at this time. However, several important observations can be made, and the steps necessary to complete this process at a later date are clear.

A number of listed missions have orbital geometries that can satisfy the MicroMaps Mission requirements: Meteorological Operational (MetOp) satellite, Solar-B, Gravity Field and Steady-State Ocean Circulation Explorer (GOCE), Astrometric Interferometry Mission (AIM), NPOESS [National POES System] Preparatory Project (NPP), Hydrosphere State (HYDROS) mission, Global Electrodynamics (GED) mission, Polar-Orbiting Operational Environmental Satellites (POES), Orbiting Carbon Observatory (OCO), and Aquarius. Orbital geometry parameters for the Spectrocopy and Photometry of Intergalactic Mediums’ Diffuse Radiations (SPIDR) mission are unknown at this time. However, with the launch vehicle designated as the Pegasus XL, orbital altitude will be low, and the inclination could be high. Thus, the SPIDR mission was retained in the final list. To discern among these missions, additional criteria must be considered. For example, higher-altitude orbits could eliminate the need for a propulsion system and simplify the platform design. At this time, no attempt was made to quantify the constraints imposed by the launch vehicle, which could also expose the better opportunities. The AIM and SPIDR missions have been rated with minimal cost sharing and high cooperation because of their designation as low-cost NASA Explorer Program missions (SMEX), which foster a spirit of cooperation in pursuing important but small-scale scientific pursuits from space. In other words, an environment exists that facilitates secondary payloads to piggyback into space for minimal cost. AIM may hold unique advantages in these latter criteria. This mission is being led by the Center for Atmospheric Sciences at Hampton University. The MicroMaps university team members and NASA Langley have a strong record of cooperation and close proximity with Hampton University and their atmospheric sciences program.

The various MIDEX and SMEX missions in Table 7 are NASA Explorer Program flights that are slated for future launch, but have not yet been awarded to a specific proposal. The mission of the Explorer Program is to provide frequent flight opportunities for scientific investigations from space. The Explorer Program enables the definition, development, and implementation of mission concepts through a variety of modes to meet the needs of the scientific community and the NASA space science enterprise. The missions are characterized by relatively moderate cost, and by small to medium-sized spacecraft that

Table 7: Potential Launch Opportunities for MicroMaps

Mission	Schedule (yr/mth)	Inclination (deg)	Altitude (km)	Mass Constraint (kg)	Size Constraint (m)	Ascent Constraint (g)	Cost ($)	Cooperation (-)
SPIDR	2005	TBD	TBD	Pegasus XL, Prime	Pegasus X, Prime	Pegasus XL	0	High
MetOp	2005 July	98.7, Sun Sync	796 × 844	Atlas II, Prime	Atlas II, Prime	Atlas II	TBD	TBD
Solar-B	2005 Sep	97.9, Sun Sync	600 km	M-V, Prime	M-V, Prime	M-V	TBD	TBD
GOCE	2006	96.5, Sun Sync	250	Rockot, Prime	Rockot, Prime	Rockot	TBD	TBD
AIM	2006	Polar Inclination	Low Altitude	TBD, Prime	TBD, Prime	TBD	0	High
MIDEX	2006	TBD	TBD	Medium-Light, Prime	Medium-Light, Prime	Medium-Light	0	High
SMEX	2006	TBD	TBD	TBD, Prime	TBD, Prime	TBD	0	High
NPP	2006 May	Polar Inclination	824	TBD, Prime	TBD, Prime	TBD	TBD	TBD
HYDROS	2006 Jun	Polar, Sun Sync	670	Taurus, Prime	Taurus, Prime	Taurus	TBD	TBD
GED	2007	Polar Inclination	350 × 2,000	Medium-Light, Prime	Medium-Light, Prime	Medium-Light	TBD	TBD
SMEX	2007	TBD	TBD	TBD, Prime	TBD, Prime	TBD	0	High
MIDEX	2008	TBD	TBD	Medium-Light, Prime	Medium-Light, Prime	Medium-Light	0	High
SMEX	2008	TBD	TBD	TBD, Prime	TBD, Prime	TBD	0	High
POES	2008 Mar	Polar Inclination	Low Altitude	Delta II, Prime	Delta II, Prime	Delta II	TBD	TBD
OCO	TBD	Polar Inclination	705	Taurus, Prime	Taurus, Prime	Taurus	TBD	TBD
Aquarius	TBD	Polar, Sun Sync	600	TBD, Prime	TBD, Prime	TBD	TBD	TBD

are capable of being built, tested, and launched in a short time interval compared to the large observatories. The three mission categories include Medium-class Explorers (MIDEX), where NASA expenses are not to exceed $150M, Small Explorers (SMEX), where NASA expenses are not to exceed $75M, and University-class Explorers (UNEX) where NASA expenses are not to exceed $15M. Therefore, the generic MIDEX and SMEX launch opportunities listed in Table 7 are projected to offer unique advantages as well. When the MIDEX/ SMEX awards are announced, their associated orbit requirements should be reviewed, and any that have been found consistent with the MicroMaps requirements should be approached early on for future collaboration.

The launch opportunity for the MicroMaps Space Mission could very well come from this final list (Table 7). Given that 16 strong possibilities were identified in a preliminary study, securing a suitable launch for MicroMaps should be feasible. With additional information, possibly obtained from communicating with the mission lead personnel, the optimum launch opportunity can be identified. Another main point to make is that MicroMaps Mission planning and design should continue, so that when a launch opportunity presents itself, the MicroMaps team can quickly respond and take advantage of this opportunity. The MicroMaps team should be ready when these opportunities arise.

VI. Conclusions and Recommendations. A mission planning process was outlined and applied to specific aspects of the MicroMaps Space Mission. All constraint and objective information from various sources was quantified, documented, and mapped into requirements for orbital geometries and spacecraft subsystem characteristics. Further sizing and definition studies in these areas for a small, dedicated spacecraft serving as the MicroMaps platform revealed no obvious critical requirements that would prevent a successful mission design and implementation. The most revealing result is an understanding of critical factors that impact the overall system design, and the key relationships among requirements, objectives, and constraints. Such understanding will be important when final engineering trades and program decision options are made.

This study provides a framework that can be revisited when more detailed information is available in more advanced planning stages. The feasibility of using the International Space Station as a space platform for MicroMaps was evaluated in specific areas, and those evaluations revealed deficiencies for this option in Earth surface coverage, attitude and vibrational transients, and the need for an active pointing system. Some of these deficiencies could be overcome, but with associated cost and complexity. Other deficiencies are simply not correct-able. Based on these results, flying MicroMaps on the International Space Station is not recommended. A small, dedicated spacecraft with a single function of supporting MicroMaps objectives is recommended. A large final list of launch opportunities with orbital characteristics and launch windows consistent with the MicroMaps Mission requirements was identified and described. Additional data and study will be needed to identify the optimum launch opportunity. The AIM mission, and future MIDEX/SMEX missions, offer unique advantages for MicroMaps. Although a specific launch opportunity has not been recommended, results indicate that finding such an opportunity should be feasible.

References

(Acknowledgement and appreciation is given to Dr. Roscoe Lee, Northrop Grumman, and Dr. Khaled Abdel-Motagaly, Titan Systems, for providing References 16–18.)

[1] Walberg, G., Reichle, H. G., and Morrow, W. H. (1999), *Design and Testing of the MicroMaps Gas Filter Correlation Radiometer,* NASA Final Report, Contract SSTI:94C:004, Langley Research Center, Hampton, Virginia.

[2] Anselmo, J. C. (1997), "TRW Scrambles to Save Satellite," *Aviation Week & Space Technology,* Vol. 147, No. 9, September 1, 1997, p. 31.

[3] Editorial Staff (1998), "NASA Cancels Clark Satellite," *Aviation Week & Space Technology,* Vol. 148, No. 9, March 2, 1998, p. 36.

[4] Reichle, H. G., et al. (1999), "Space Shuttle Based Global CO Measurements During April and October 1994, MAPS Instrument, Data Reduction, and Data Validation," *Journal of Geophysical Research*, Vol. 104, No. D17, pp. 21,443–21,454.

[5] Connors, V. S., Novelli, P. C., and Reichle, H. G. (1996), "Space Shuttle Views Changing Carbon Monoxide in Lower Atmosphere," *Transactions of the American Geophysical Union*, Vol. 77, No. 47, pp. 466–467.

[6] Griffin, M. D. and French, J. R. (1991), *Space Vehicle Design*, American Institute of Aeronautics and Astronautics, Washington, District of Columbia.

[7] Brown, C. D. (1992), *Spacecraft Mission Design,* American Institute of Aeronautics and Astronautics, Washington, District of Columbia.

[8] Fortescue, P. and Stark, J. (1992), *Spacecraft Systems Engineering,* John Wiley and Sons, New York, New York.

[9] Larson, W. J. and Wertz, J. R. (1992), *Space Mission Analysis And Design,* Microcosm and Kluwer Academic Publishers, Torrance, California.

[10] Craig, R. A. (1965), *The Upper Atmosphere: Meteorology and Physics,* Academic Press, New York, New York.

[11] Sutton, G. P. and Biblarz, O. (2000), *Rocket Propulsion Elements*, John Wiley and Sons, New York, New York.

[12] Technical Staff (2002), *PPT-8/9 Electric Propulsion System,* Product Specification Sheet, CU Aerospace Inc., Urbana, Illinois.

[13] Technical Staff (2002), *Tandem-200 Electric Propulsion System,* Product Specification Sheet, Busek Inc., Natick, Massachusetts.

[14] Technical Staff (2002), *Rectangular Dual Junction GaAs/Ge Solar Cell,* Product Specification Sheet, Spectrolab Inc., Sylmar, California.

[15] Technical Staff (1997), *International Space Station: Assembly Complete,* NASA-LG-1997-07-456-HQ, NASA Headquarters, Washington, District of Columbia.

[16] Thorn, V. (1999), "ISS Flight Attitudes," *Proceedings of the AIAA International Space Station Service Vehicles Conference,* Houston, Texas.

[17] Technical Staff (1994), *Space Station Program (SSP) 41162, United States On-Orbit Segment (USOS) Specification,* NASA JSC Report SSP-41162, Revision AE, Johnson Space Center, Houston, Texas.

[18] Technical Staff (1994), *Space Station Program (SSP) 41163, Russian Segment (RS) Specification,* NASA JSC Report SSP-41163, Revision H, Johnson Space Center, Houston, Texas.

[19] Woods-Vedeler, J. A. and Rombado, G. (1997), *DAC-3 Pointing Stability Analysis Results for SAGE III and Other Users of the International Space Station (ISS) Payload Attachment Sites (PAS),* NASA-TM-110318, Langley Research Center, Hampton, Virginia.

[20] Technical Staff (2002), "ISS Missions 1998–2006, Space Station Assembly Sequence," *Space Station User's Guide Data Base,* SpaceRef.com, Reston, Virginia, http://www.spaceref.com/iss/schedules/10.30.00.as.html.

[21] Technical Staff (1999), "NASA/Commercial/DOD Baseline Mission Models Through 2020," *Space Transportation Architecture Study—Document B,* NASA Headquarters, Washington, District of Columbia, http://www.hq.nasa.gov/office/codeae/documentb.html.

Authors

Ossama Omar Abdelkhalik is a doctoral candidate in the Department of Aerospace Engineering, Old Dominion University. His interests lie in remote-sensing space mission design and control of satellite formation flight. Before this position, Mr. Abdelkhalik held an engineering staff position with Carlo Gavazzi Space Corporation, Italy, where he designed control systems for low Earth orbit satellites. He was awarded BS and MS degrees from the Department of Aerospace Engineering, Cairo University, Egypt.

Bassem Nairouz is a doctoral candidate in the Department of Aerospace Engineering, Old Dominion University. His interests lie in attitude dynamics and control of mini-micro class spacecraft. He was awarded BS and MS degrees from the Department of Aerospace Engineering, Cairo University, Egypt.

Timothy Weaver is a masters candidate in the Department of Aerospace Engineering, Old Dominion University. His interests lie in asteroid space mission design and advanced electric propulsion concepts. He was awarded a BS degree from the Department of Mechanical and Aerospace Engineering, University of Alabama at Huntsville.

Brett Newman is an associate professor in the Department of Aerospace Engineering, Old Dominion University. He has been principal investigator for numerous government-industry sponsored projects concerning the dynamics and control of spaceflight and atmospheric flight vehicles. He has authored over 50 technical publications and secured approximately $0.9M in externally supported research in this field. Before this position, Dr. Newman held the position of guidance and control engineer with Orbital Sciences Corporation, where he was involved with the design and implementation of various guidance and attitude control systems for several suborbital missions. He was awarded BS and MS degrees from the School of Mechanical and Aerospace Engineering, Oklahoma State University, and a PhD degree from the School of Aeronautics and Astronautics, Purdue University.

Building a Pathway to Mars: Technology Investment for Science Return

Jeffrey H. Smith, Julie Wertz, and Charles Weisbin

Abstract

The exploration of Mars has been the focus of increasing scientific interest about the planet and its relationship to Earth. A multi-criteria decision-making approach was developed to address the question, "Given a Mars program composed of mission concepts dependent on a variety of alternative technology development programs, which combination of technologies would enable missions to maximize science return under a constrained budget?" The scientific value of each portfolio was used to compute each portfolio contribution to a strategic exploration goal. Solutions were found by searching all possible portfolios for the maximum science value within budget constraints.

I. INTRODUCTION. There has been considerable interest in the scientific community and at NASA in addressing fundamental questions about the planet Mars [1–3]. NASA's program for the exploration of Mars is linked to a need for numerous enabling technologies that must be developed in order to proceed with the variety of missions planned.

A diverse mixture of programmatic issues faces the Mars Exploration Program. The complex interactions between scientific interests, missions, technologies, and budgets has amplified the need for an organizing structure to provide insights about high-value technologies and mission sensitivities to technology development uncertainties and budget constraints. The purpose of this paper is to describe such an organizing structure used to address this problem. A combined approach was developed for analyzing portfolios of technology investments using multi-criteria decision analysis, Monte Carlo simulation, and mathematical programming techniques [4–6]. The approach enumerated every possible technology portfolio combination to identify sets of highest science-value missions and technologies that could be funded within a specified budget. This was done in a stepwise fashion by simulating the uncertainties in every technology required by every mission. If, during the simulation, a technology development failed, its parent mission was removed from the portfolio. The science value of the remaining missions was then computed and the total technology cost by year was compared to the budget for feasibility. The process was repeated to obtain the probabilistic uncertainties and their impacts on technology outcomes. The resulting outcomes were sorted by science value, technology value, cost feasibility, and (in some cases) minimum cost and maximum number of enabled missions.

The approach and results obtained were viewed to have value in unraveling the complex interdependencies, risks, and uncertainties pertinent to the Mars Exploration Program. Many of the varied planning concerns (mission candidates, science values, technology risks, uncertainty, investment costs, budgets, and timing) were aggregated in a fashion that allowed planners to quantify the overall effect of alternative assumptions and possible actions on the Program.

This paper represents a first attempt to apply multicriteria decision analysis techniques to the Mars technology R&D program. A brief description of the Mars missions, technologies, and cost assumptions is presented first. The next section describes the approach, followed by the results obtained. The last section provides a discussion of these results and the conclusions.

II. Finding the Pathway. Finding a pathway to Mars in the context of conflicting science objectives, mission requirements, uncertain technologies, and limited resources is fraught with innumerable possibilities. As a first step, the scope of the problem was defined in terms of science objectives, the missions candidates, the technologies required, and the assumptions made.

The science objectives for the Mars Exploration Program were, at the time of the study, divided into three categories aimed at addressing three overarching questions:

1. Is there life on Mars?
2. If not, has there ever been life on Mars?
3. What happened to the global climate on Mars?

These questions had been translated into a number of strategic "pathways" designed to address each question through scientific measurements [3]. The emphasis of the pathways was a weighted sum of eight levels of priorities assigned to 192 scientific measurements. The three pathways included a Mars in situ strategy, a Mars sample return strategy, and a global cycles climate strategy. This paper reports on the results of a combined strategy that was a weighted combination of the three science pathways. Because each science pathway emphasized a different set of scientific measurements (and missions), different pathway strategies (in situ, sample return, global climate) could be amplified by increasing the weighted value of those measurements. For example, emphasis on the sample return science pathway utilized a 20%, 60%, 20% allocation of science measurements to the in situ, Mars sample return, and global climate pathways so that twenty percent of the total number of 192 measurements (115) was allotted to in situ missions, sixty percent to the sample return mission, and twenty percent to the global climate missions. In this manner, programmatic inputs were used to examine the impacts of focusing on different exploration paths. Additional cases included a variety of in situ pathway assumptions.

The missions considered for implementing each scientific pathway are summarized in Table 1. The alternatives included three lander/rover missions, four orbiter missions, one Mars sample return mission, and one lander/drilling mission,

It should be noted, in some cases the missions in Table 1 were candidate missions that served as placeholders for evolving mission concepts and science pathways. In some cases, only one of two orbiter concepts might be chosen or two of three landers were planned. The determining factor in such cases was often the technology development cost or cost

Table 1: Mars Mission Candidates

Mission Name	Description
Mars Science Laboratory	Mission to measure science measurement in situ with a rover
Volcanology Rover	Rover mission to characterize volcanic region with in situ sampling
Polar Layer Deposit Rover	Rover mission to characterize polar regions with in situ sampling
Synthetic Aperture Radar Orbiter	Orbiter sounding for surface science experiments and mapping
Imaging/Atmospheric Sounding Orbiter	Next-generation remote-sensing orbiter (imaging and atmospheric sounding)
G. Marconi Orbiter	Telecommunications orbiter relay for high-data-rate communications
Telesat Orbiter	Small Mars telecommunications orbiter for high-data-rate communications
MSR Sample Lander	Sample return with a Mars ascent vehicle
Wildcat Lander	Lander with 30-m depth drilling system

coupled with the technology development requirements and development challenges (chance of success).

Each of these missions had a variety of requirements for enabling technologies. A list of 110 technologies was divided into 18 representative categories. A performance attribute was defined to characterize each technology category requirement and corresponding technology development task. Table 2 lists the high-level attributes and their definitions.

The technology capabilities in Table 2 were then mapped to the missions in Table 1 to define a roadmap of enabling technologies by mission. The nine missions mapped 18 technology capabilities to a total of 31 unique technology requirements. This was due to sharing of common requirements by some missions and a natural partitioning between rover, lander, and orbiter missions. In each of the 31 cases, data were obtained from technologists, mission designers, and program documentation.

Additionally, a number of technologies (17) were dependent upon the success of predecessor technologies. These dependencies were differentiated as technology, mission, and cost dependencies. Technology dependencies were used to disable dependent technologies if their predecessor failed. Mission dependencies were used to disable missions dependent on the successful development of a precursor mission. Cost dependencies were used to carry forward technology funding from an unsuccessful technology to an enhancing dependent technology. Table 3 lists the data items gathered for each technology attribute.

Finding a feasible path through the large number of possible technology investments required combining Tables 1, 2, and 3 in a manner that would sort out the high-science-value, high-technology-capability, low-risk, and low-cost technologies while discounting the less promising (i.e., lower performing and risky) and more expensive technologies.

A systematic approach was developed to address the question of identifying high-value technology investment portfolios by enumerating every possible technology portfolio combination and searching for the lowest technology cost portfolio that enabled the most science within the mission budget constraint. The resulting technology portfolios provided guidance on where technology investments should be made for the chosen science pathway strategy. The next section describes the approach used to find this pathway.

III. Approach. Figure 1 illustrates the process used. The first two steps (1, 2) culminated in Table 1, the next two steps (3, 4) produced Table 2, and step 5 was captured by Table 3. The focus of this section is on step (6) and the procedure for evaluating the alternative portfolios.

The process outlined above can be restated in the following mathematical terms. Let the technology attributes be defined as random variables $x_1, x_2, \ldots, x_n$ each with probability density functions $f_1(x_1)$, $f_2(x_2)$, …, $f_n(x_n)$. Let the technology capability value for each attribute be rep-

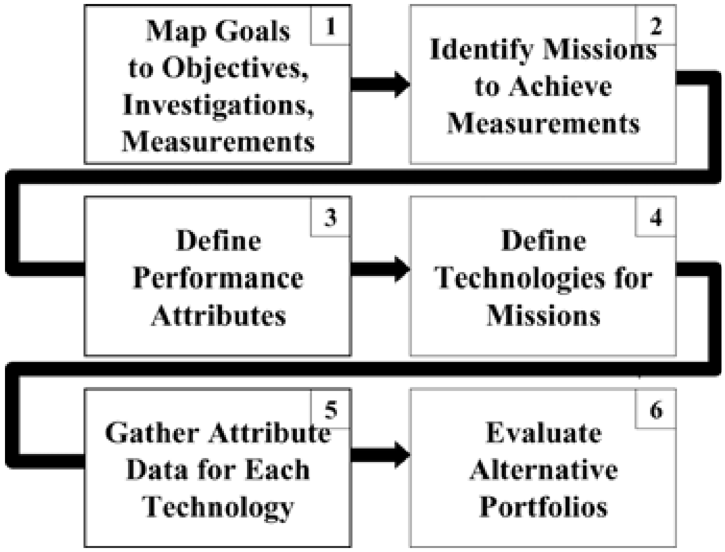


Figure 1. Mars Portfolio Analysis Approach

resented by an attribute value function that maps the range of each attribute, x, to a value, $v(x)$, between zero and one, $0 \le v(x) \le 1$. Using a multiattribute decision analysis approach,[5] the best state of each attribute was scaled to a value of one and the worst state of the attribute was defined as having zero value. It can be shown that attribute value functions $v_1(x_1)$, $v_2(x_2)$, …, $v_n(x_n)$ can be used (under an assumption of preferential independence) to compute a multiattribute value function for the portfolio of each technology set within a mission:

$$V(Mission\ \ i) = V(\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n) = \sum_{j=1}^{n} k_j v_j(\underline{x}_j) \qquad \text{Equation (1)}$$

where the $\underline{x}_j$ represents the mission-specific realizations of each technology j. To compute a measure of technology value for a mission, i, the values of each randomly sampled attribute would normally be substituted in the corresponding value functions, v_j, and $V(i)$ computed for each mission using a weighted sum. The process would be repeated in Monte Carlo simulation fashion until a statistically representative result was obtained.

The weights, k_j, were set equal in the first application of the methodology to simplify and expedite interpretation of the results. Because the attributes were random variables, the probability distribution uncertainties had to be transformed through the attribute value functions into a probability distribution for each result, $V(i)$. This was done using Monte Carlo simulation to generate expected values that reflected the uncertainties of each technology task. During this process, technology tasks failed in accordance with their estimated task probabilities of success,

Table 2: Technology Performance Attributes

Technology	Attribute Definition
Precision Landing	Semi-major axis ellipse distance, kilometers. Width of landing ellipse with 99% landing probability.
Impact Attenuation	Landing survivability, meters. Free-fall distance at terminal landing phase for pallet-based landers.
Hazard Avoidance	Average size of identifiable rock on 30-degree slope to be avoided during landing.
On-Orbit Science Resolution	Resolution of primary instrument, meters/pixel.
On-Orbit Science Wavelength	Specific wavelength of primary instrument, meters.
Forward Planetary Protection	Number of organisms present on the spacecraft (thousands).
Forward Planetary Protection	Measurement time after cleaning to process spacecraft (hours).
Surface Sample Characterization	Technology Readiness Level of instrument package designed for Mars surface sampling. Measured on 1–9 scale using a narrative definition [7].
Subsurface Access (drilling) Technologies	Achievable depth of drilling subsystem, meters. Two cases: shallow (30 m) and deep (1000 m) technologies.
Surface Mobility	Distance capable of roving, meters per sol (Martian day).
Surface Sample Handling	Sample cross-contamination limit, parts per million.
Back Planetary Protection	Minimum containment size of particle within sample return system, microns.
Mars Proximity Data Rate	Data rate among communications systems (and missions) at Mars, megabits/second.
Mars-to-Earth Data Rate	Data rate for transmission to Earth, megabits/second.
Mars Orbit Rendezvous	Sample capture system time to acquire sample, sols.
Multimission Survivability	Infrastructure technologies to extend component lifetimes, sols. Two cases: on-orbit and surface technologies.
Surface Instrument Approach and Placement	Time for rover to plan, traverse to target, and place instrument on sample, sols.
Mars Ascent Vehicle	Qualification temperature of ascent engines, °C.

Table 3: Data Inputs for Mars Technologies

Data Item	Description
Technology Capability Estimate	Estimate of technology attribute requirement outcome given technology development budget and development task is 100% successful. Value can be a point estimate, range, or probability distribution.
Probability of Success	Estimate of probability of technology development task success (based on likelihood of budget changes, dependencies on external developments, or task complexity).
Default Outcome	Likely value of technology attribute outcome if technology development fails completely or partially. Use state-of-the-art or descope option.
Technology Budget Constraint Profile	Resources planned for development task in 3-year increments over a 12-year planning horizon, real-year dollars.
Dependency	Identifier of parent technology and type of dependency (technical, mission, cost).

and in those cases, the default value was used in place of the sampled value.

Because the technologies were regarded as enabling for the missions that depended on them, a technology failure within a mission was equivalent to removing the mission from the portfolio for a single Monte Carlo trial. The success of each technology was critical to the development of its mission. If any single technology failed to develop within a mission that needed it, then the mission was disabled. This assumption transformed the procedure of Equation 1 into:

$$V(Mission\ \ i) = \begin{cases} 0 & if\ \ any\ \ technology\ \ \underline{x}_1, \underline{x}_2, ..., \underline{x}_n\ \ fails) \\ 1 & if\ \ no\ \ technology\ \ \underline{x}_1, \underline{x}_2, ..., \underline{x}_n\ \ fails) \end{cases} \quad \text{Equation (2)}$$

Sampling of the x_j values was used to identify failures based on the probability of success inputs (Table 3). The missions remaining in the portfolio after each Monte Carlo trial were then recorded and the statistical expected value and standard deviation of the portfolio science return was accumulated prior to sampling for the next iteration. The process was repeated 5,000 times for each portfolio combination. The analysis of 9 missions entailed $2^9 - 1 = 510$ portfolio combinations.

During the Monte Carlo simulation, the dependencies were applied in accordance with the type of dependence. Technology dependencies were applied within each trial of the simulation; mission dependencies were applied at the portfolio level; and cost dependencies were applied at both the trial and portfolio level based on whether the parent technology or mission had failed or was not included in that specific portfolio combination. After each portfolio was simulated, the missions in the portfolio, the expected science value and standard deviation, the portfolio technology versus budget cost profile, and portfolio mission cost versus budget profiles were recorded. After the entire set of 510 portfolios was simulated, the list was sorted first by mission cost feasibility to eliminate those portfolios exceeding the mission budget cap. Next, the list was sorted by technology cost versus budget to eliminate the cases that fit within the mission cost budget while exceeding the technology budget. Finally, the list was sorted by expected science value to reveal the highest expected science value portfolio feasible within the technology and mission cost budgets.

Three budget profiles were examined: 25, 50, and 75 million dollars per year (real-year dollars). A first-order feasibility criterion was used to determine cost feasibility—if the total technology costs exceeded the budget for any year, the portfolio was declared infeasible and discarded. It should be noted that no attempt was made to shift funds and technology costs to resolve feasibility problems.

IV. Results. Although a number of cases and sensitivity studies were examined, this paper reports on the primary results obtained for technology budget profiles of $25M, $50M, and $75M per year. The results provided insights into which technologies were important for strategic funding and also identified missions enabled by those technologies. Table 4 summarizes the baseline results for each of the three budget assumptions.

At the $25M/yr technology budget, only 24 out of the 510 portfolios met the budget constraints. One lander/rover and one orbiter had the lowest technology costs that fit within the budget profile. The striking result was that although this was an in situ science pathway, only one of the in situ options was feasible from a technology- and mission-enabling perspective.

At the $50M/yr technology budget, the number of affordable technology portfolios increased to 288 out of 510 possibilities, and it allowed 15 additional technologies to enter the solution, which enabled three additional missions. From these results it was clear that the $50M/yr budget had opened the trade-off space between technologies and enabled a variety of missions (in situ, sample return, and global orbiters).

All 510 technology portfolios fit within the technology budget constraint at the $75M/yr level, enabling one additional mission. The mission budget constraint—coupled with the higher cost of the Polar Lander and Wildcat missions, their higher risks, dependencies on the Mars Science Laboratory, and lower expected science value—prevented their entry in the optimal (highest science value) solution for the sample return pathway. As would be expected, if the pathway strategy were revised to emphasize in situ exploration, the in situ missions push the sample return mission out of the solution. In that case the mission portfolio solution becomes as follows: Mars Science Laboratory, Volcanology Rover, Polar Layer Deposit Rover, Synthetic Aperture Radar orbiter, Imaging/Atmospheric Sounding orbiter, and Wildcat.

The fact that an additional $25M/yr allowed only one added technology beyond the $50M/yr case was an indication that many of the technology trade-offs were likely to be in the neighborhood of $50M/yr (for example, $40–$60M/yr).

Figure 2 displays all of the 510 portfolios in three dimensions for the $50M/yr case and the sample return pathway. Shown are rectangular polygons representing the expected science value plus or minus one standard deviation for each portfolio. (The expected science value is in the vertical center of each box.) The portfolio science values are positioned on the location corresponding to the total portfolio mission cost and total portfolio technology costs. Also shown are the total budget planes for mission and technology costs. Embedded within the display is the optimum solution presented in Table 4 at $50M/yr. A number of observations are noteworthy.

First, a number of portfolios are too expensive—these can be eliminated from further consideration. Second, the top view shown in Figure 3 reveals that all of the portfolios are potentially affordable at $50M/yr since most of the technology budget violations occur in the first 5 years with excess funds in years 6–12. This highlights the need to reallocate the resources in the long term forward to the present. In fact, the total

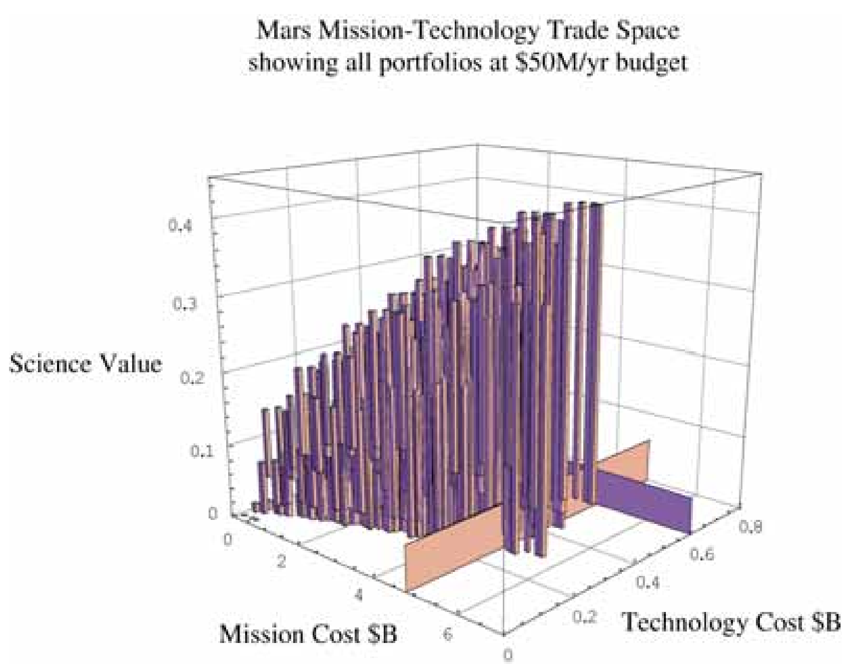


Figure 2. Display of Portfolio Results Showing Expected Science Value versus Mission and Technology Total Costs and Budget Constraints.

Table 4: Mars Technology Portfolio Results for Three Investment Levels Showing Feasible Technologies and Missions Enabled (Sample Return Pathway)[1]

Technology Investment	Technology Portfolio	Missions Enabled
$25M per Year	❑ Sample characterization ❑ Rover mobility at 160–200 m ❑ On-orbit science resolution ❑ Telecom network, Mars-to-Earth	❑ Volcanology Rover ❑ Imaging/Atmospheric Sounding orbiter
$50M per Year	❑ Precision landing ❑ Impact attenuation ❑ Hazard avoidance ❑ Forward planetary protection time ❑ Sample characterization ❑ Mobility at 230–450 m ❑ Sample handling, contamination ❑ Multimission survivability ❑ Approach/instrument placement ❑ Sample characterization ❑ Mobility at 160–200 m ❑ On-orbit science resolution ❑ Telecom network, Mars-to-Earth ❑ Precision landing ❑ Impact attenuation ❑ Forward planetary protection time ❑ Forward planetary protection, number of organisms ❑ Back planetary protection ❑ Mars orbit rendezvous ❑ Mars ascent vehicle	❑ Mars Science Laboratory ❑ Volcanology Rover ❑ Imaging/Atmospheric Sounding orbiter ❑ Mars Sample Return
$75M per Year	All of $50M case plus Synthetic Aperture Radar technology: ❑ On-orbit science, wavelength	❑ Mars Science Laboratory ❑ Volcanology Rover ❑ Synthetic Aperture Radar orbiter ❑ Imaging/Atmospheric Sounding orbiter ❑ Mars Sample Return
$75 M per Year	In situ pathway	❑ Mars Science Laboratory ❑ Volcanology Rover ❑ Polar Layer Deposit Rover ❑ Synthetic Aperture Radar orbiter ❑ Imaging/Atmospheric Sounding orbiter ❑ Wildcat

1. The G. Marconi and Small Telesat orbiters did not appear in any solutions—they had no explicit science value. Subsequent to the study it was determined such “enhancing” missions should have been modeled as technologies for use by one or more science mission candidates.

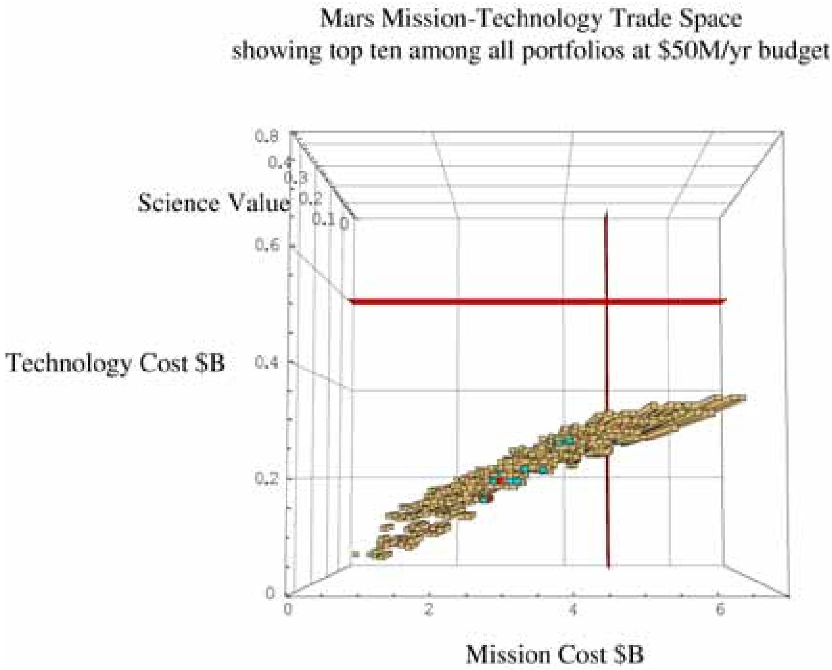


Figure 3. Display of Portfolio Results (Top View) Showing Expected Science Value versus Mission and Technology Total Costs and Budget Constraints.

budget could be much less than 12 yrs × $50M/yr if a more accurate budget profile were used—the budget was fixed in this study at $25M, $50M, and $75M for each as an arbitrary starting point. Third, it can be observed that expected science value, portfolio uncertainty, mission portfolio cost, and technology portfolio costs increase together (although technology costs increase at a diminishing rate). This is due to the increasing number of missions in the portfolio, which adds science value, and carries additional technologies that have added risk.

V. DISCUSSION AND CONCLUSIONS. The results were presented to the Mars Program Systems Engineering Team and endorsed by that group as providing valuable insights and benefits for Mars Program planning. During the course of their review, the team also identified a number of key areas for further improvements.

Benefits

The first benefit of the methodology was in providing a systematic approach that addressed four issues critical to the Mars Exploration Program:

1. identifying key technologies and their risks to candidate mission concepts;
2. linking science objectives to technology selection;

3. including technological uncertainties; and
4. applying costs and budget constraints to the selection of feasible technologies.

In particular, the ability to provide an audit trail through the process from science objectives, to technology capabilities, to enabled missions, and ultimately to the feasible technology portfolios was viewed as a major contribution.

A second benefit was in capturing key aspects of the problem facing Mars Program planners. The relationships between technologies, risks, costs, missions, and budget constraints embodied a complex nest of interactions making it difficult to unravel the effects of adding or deleting technologies, modifying science objectives, or changing budgets and costs. The approach aided in managing these effects by modeling important relationships in a consistent manner that allowed a variety of planning assumptions to be tested.

A third benefit was the ability of the methodology, and particularly the software tool, to enumerate and evaluate rapidly every mission technology portfolio combination. This provided an additional level of confidence in the approach that every case possible had been considered rather than some limited set produced by a time-constrained committee or because of modeling limitations.

A fourth benefit was the enhancement of communication between Mars Program mission planners and technologists. It was observed that mission planners sometimes levied requirements they viewed as goals whereas the technologists viewed the requirements as fixed and had assumptions and constraints about the requirements not communicated clearly to the mission planners. In some cases, missions were surprised to discover they were assumed to be developing predecessor technologies for subsequent missions. Technologists were similarly amazed to find that expectations about their development tasks exceeded their own objectives. The interactive process of gathering the data for Table 3 raised awareness and clarified understanding about assumptions, budgets, and work efforts not clearly understood or defined prior to the exercise.

Limitations and Improvements

Notwithstanding these benefits, the approach did have a number of limitations. The first issue surfaced by the Mars Program Systems Engineering Team involved questions about the uncertainties in technology definitions and data quality. While it was acknowledged that estimation of costs and technology development over a 12-year horizon was difficult, it was argued that having the ability to examine the effects of data variability was at least a first step toward understanding how such estimates might be improved. A second-round analysis was recommended by the Mars Program Systems Engineering Team to refine and improve the definitions of missions, technology attributes, and data values.

A second issue was the effect of temporal dependencies among missions in a portfolio. The sequencing of missions is a process designed to provide "feed-forward" information from one mission to the next. For example, mapping by an orbiter could be used to improve knowledge about future landing sites for landed missions. The current methodology did not attempt to model explicitly this "learning" aspect of mission success.

A related capability to degrade technologies gracefully in the event of failures was also seen as important by the Mars Program Systems Engineering Team for identifying task development shortfalls that provide acceptable technology deliveries. This capability has been added (but not exploited) to address descope options.

A third limitation was the focus on technology investment cost feasibility as simply the difference between total technology cost and budget within each time period. Well-founded techniques to optimize the budget resource profile should be incorporated to allow the movement of excess budget funds (subject to constraints) from adjacent years into years where insufficient funds have identified a potentially viable portfolio as infeasible.

During the course of developing and applying the R&D portfolio model to the Mars Exploration Technology Program, a number of conclusions were drawn.

- At the lowest technology funding levels, the in situ science strategy was not feasible. Low levels of technology funding implied a limited (2 mission) program.
- The highest level of technology funding proved to enable all technologies in the portfolio under the current assumptions. However, pathway strategies, mission costs, and risks reduced expected science values to prevent some enabled missions from entering the solution. Alternative pathways (e.g., emphasis on in situ or global climate cycles) determined by Mars Program scientists and planners will ultimately determine the mission and technology portfolio. As science goals evolve and mission concepts are added, modified, and deleted, different technology portfolios would be derived.
- The inclusion of technology cost profiles and budget constraints immediately focused attention on feasible options by eliminating the portfolios. At the $50M/yr level, 44% of the portfolios were eliminated; at the $25M/yr level, 95% of the portfolios were eliminated.
- The methodology provided a systematic rationale that linked science objectives to enabling technologies to missions and identified high-science-value technology portfolios that minimized technology costs and risks.

- The R&D portfolio approach helped clarify understanding between mission planners and technology developers

The application of the systematic tools and techniques described in this paper to Mars technology and mission planning provided a quantifiable and traceable approach to Mars Program personnel about science, technology, and mission interdependencies. The identification of high-value portfolios was seen as a first step toward making appropriate technology investments for defining the pathway to Mars.

VI. ACKNOWLEDGMENTS. This work was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. The authors gratefully acknowledge the support, both financial and technical, that enabled this work to be performed. In particular, we very much appreciate the sponsorship of Charles Whetsel (Mars Program Chief Engineer), Yuri Gawdiak (Program Manager for the Engineering of Complex Systems Program), Bob Pearce (Manager of the Division of Strategic Planning for NASA Code R, and Steve Prusha, Level 2 Manager for Systems Reasoning and Risk management. The authors would also like to thank Ben Dolgin and Charles Kohlhase for their help and advice.

References

[1] Farmer, J., and D. Des Marais, "Exploring for a record of ancient Martian life." *J. Geophys. Res.* Vol. 104, 1999, pp. 26,977–26,995.

[2] Greeley, R. (Chair), Mars Exploration Program/Payload Analysis Group, Part 2, "Scientific goals, objectives, investigations, and priorities," from *Science Planning for Exploring Mars*, JPL Publication 01-7, Jet Propulsion Laboratory, Pasadena, California, July 2001, pp 10–38.

[3] Mars Exploration and Payload Analysis Group Pathways Science Steering Group Report, "Investigation-Driven Science Pathways for Mars Exploration," Executive Summary, Jet Propulsion Laboratory, Pasadena, CA, June 5, 2002.

[4] McGeary, Michael, and Philip M. Smith, "The R&D Portfolio: A Concept for Allocating Science and Technology Funds," *Science*, Vol. 274, 29 Nov. 1996, pp 1484–1485.

[5] Feinberg, A., and J.H. Smith, "Selection and Scaling of Attributes in Multi-Criteria Decision Making for Ranking Advanced Technology Options," presented at the 1994 Joint National Meeting of TIMS/ORSA: OR/MS Without Boundaries, Boston, Massachusetts, April 24–27, 1994.

[6] Smith, J.H., and A. Feinberg, "Applying Super-computing Methods to Multicriteria R&D Planning Problems," invited paper for the Session on Multicriteria Decision Making, 1993 Operations Research Society of America (ORSA), Phoenix, Arizona, Oct. 31–Nov. 3, 1993.

[7] Mankins, J.C., "Technology Readiness Levels, A White Paper," Advanced Concepts Office, NASA Office of Space Access and Technology, Washington, D.C., April 6, 1995.

Authors

Jeffrey H. Smith is a member of the Mission and Systems Architecture Section at the Jet Propulsion Laboratory. He has developed decision analysis methodologies and tools for the analysis of advanced space missions and automated systems for a variety of NASA's missions. He has authored over 100 technical papers spanning a wide variety of disciplines and applications. Dr. Smith's recent work includes technology investment portfolio analysis and automated resource allocation systems. He has extensive experience in decision support systems and software development tools for technology management. He has a BS degree in Applied Mathematics, an MS degree in Systems Engineering from the University of Arizona, and an MA in Mathematical Economics and PhD in Business Administration from the University of Southern California.

Julie A. Wertz is a doctoral candidate in the Space Systems Laboratory at the Massachusetts Institute of Technology. Her research is focused on risk and reliability analysis for complex space systems—specifically, introducing risk analysis tools in the early stages of the design process. As an application of her research, she is conducting a risk-analysis study for the Terrestrial Planet Finder mission. In addition to her doctoral work, Ms. Wertz is a part-time engineer in the Mission and Systems Architecture Section at the Jet Propulsion Laboratory (JPL). She has worked on a variety of projects at JPL, including technology investment portfolio analysis, flight options analysis, and a technology investment trade study. She has SB and SM degrees in Aeronautical and Astronautical Engineering from the Massachusetts Institute of Technology.

Charles R. Weisbin received his PhD in Nuclear Engineering from Columbia University in 1969. He currently serves as Deputy Program Manager for the Strategic Systems Technology Office of the Chief Technologist at the Jet Propulsion Laboratory. Prior to this, he led the Surface Systems Thrust Area at JPL, within the NASA Cross-Enterprise Technology Development Program. He was Program Manager for Robotics and Mars Exploration Technology programs, and before this, Section Manager for Robotic Systems and Advanced Computer Technology. At the Oak Ridge National Laboratory, he was Director of the Robotics and Intelligent Systems Program and Director of the Center for Engineering Systems Advanced Research. Dr. Weisbin was Associate Professor of Computer Science at the University of Tennessee.

National Aeronautics
and Space Administration

Center for Space Mission Architecture and Design
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

JPL 03-15