JOURNAL OF SPACECRAFT AND ROCKETS
Vol. 39, No. 4, July–August 2002

Space Systems Flexibility Provided by On-Orbit Servicing: Part 1

Joseph H. Saleh,* Elisabeth Lamassoure,† and Daniel E. Hastings‡
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Spacecraft are still the only complex engineering systems without routine maintenance, repair, and upgrade infrastructure. Whereas the technologies for autonomous on-orbit servicing of satellites are emerging, no general conclusions have yet been drawn regarding the cost effectiveness of on-orbit servicing for space systems. The traditional approach to this problem has been to compare the potential cost savings from servicing to the estimated cost of a chosen servicing architecture. A general model is proposed that applies this approach to a large trade space of space missions and servicing infrastructures. When typical results are analyzed, it demonstrates two major limitations of the conventional approach. First, the servicing cost uncertainty is too high to yield meaningful conclusions. Second, this approach does not take into account the flexibility provided by on-orbit servicing to space systems. A new perspective on on-orbit servicing is then proposed. The value of servicing, defined as the maximum price under which on-orbit servicing would be interesting, is studied independent from its cost. In addition, a framework to quantify the value of the flexibility provided by on-orbit servicing to space systems is developed. This framework can be used to identify the space missions for which on-orbit servicing would offer the most potential and should serve as a guide to future technology development.

Nomenclature

A_M = Markov matrix describing the state transitions, year^{-1}
a_0 = baseline constellation circular altitude, m
a_{00} = satellite launch altitude, m
a_1 = servicer launch altitude, m
B = Brownian process with unit volatility and zero mean
BC = ballistic coefficient, kg/m^2
d = discount factor, %/year
F_S = satellites development cost factor
F_0 = servicer vehicle development cost factor
f_p = propulsion dry mass factor
f_{st} = structures mass factor
f_{stC} = structures mass factor for cargo on a servicer
g = gravitational acceleration at Earth's surface, m/s^2
I = integer part function where $I(x)$ is integer and $I(x) \leq x < I(x) + 1$
I_{spS} = specific impulse of customer satellites, s
I_{sp0} = specific impulse of servicer vehicles, s
M_C = cargo mass to be delivered to each satellite, kg
M_{dry}^{base} = satellite dry mass without propulsion system and its structures, kg
M_{dry}^{prop} = dry mass of the spacecraft propulsion system, kg
M_{dry}^{tot} = total spacecraft dry mass, kg
M_{fuel} = total satellite fuel mass
M_{launch} = total satellite mass at launch
M_{struc} = total satellite structural mass
M_0 = servicer minimum dry mass (payload and its structures)
$\mathcal{M}_j$ = number of satisfied users per unit time in state j, year^{-1}
N_{nspp} = number of required satellites per plane
N_p = number of orbital planes
N_{sched} = number of scheduled servicing events
N_{spares} = number of satellite spares per orbital plane
N_{spp} = number of satellites per plane (required plus spares)
N_{sps} = number of satellites per servicer
P_C = probability of a catastrophic failure when attempting servicing
$P_k(t)$ = probability to be in state k at t
r = risk-free interest rate (time value of money), %/year
T_H = time horizon over which missions are evaluated, year
T_0 = orbital period at the altitude a_0, s
V_0 = orbital velocity at the altitude a_0, m/s
v_D = ΔV for the seriver to deorbit normalized by $g I_{sp}$
v_H = ΔV to go from servicer orbit to satellite docking normalized by $g I_{sp}$
v_P = ΔV to go from one satellite the next normalized by $g I_{sp}$
α = expected rate of return of the revenues, %/year
Δi = inclination change required for servicing, deg
ΔT_{max} = maximum time allowed for a maneuver, s
ΔV_D = incremental velocity to deorbit, m/s
ΔV_d = incremental designed to be performed before refueling, m/s
ΔV_f = incremental velocity for fine proximity maneuvers, m/s
ΔV_H = incremental velocity for a Hohmann transfer, m/s
ΔV_{inc} = incremental velocity for a change in inclination, m/s
ΔV_{ph} = incremental velocity for a circular phasing maneuver, m/s
ΔV_{yr} = incremental velocity for orbit maintenance over one year, m/s
ϵ = fuel fraction carried on spacecraft during launch
λ = failure rate, year^{-1}
μ_k = maintenance (replace or refuel or repair) rate from state k, year^{-1}
ϕ = baseline phase between two coplanar satellites, deg

Introduction

SPACE systems are still the only complex engineering systems without maintenance, repair, and upgrade infrastructure. One-of-a-kind, reliable, and expensive spacecraft, designed for the longest possible lifetime, have been the result of this lack of space logistics. On-orbit servicing has long been suggested as having the

Received 18 June 2001; revision received 1 February 2002; accepted for publication 11 February 2002.

*Research Assistant, Department of Aeronautics and Astronautics; currently Associate, McKinsey and Co., Washington, DC 20005.

†Research Assistant, Department of Aeronautics and Astronautics; currently Engineer, Advanced Space Concepts, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91125; elisabeth.lamassoure@alum.mit.edu. Member AIAA.

‡Professor, Aeronautics and Astronautics and Engineering Systems, 33-413, 77 Massachusetts Avenue; hastings@mit.edu. Fellow AIAA.

potential to change the way business is carried out in space. However, its implementation requires a whole new way of designing and managing space systems. In addition, decision makers perceive it as a significant source of mission risk. For on-orbit servicing to be actually deemed worthwhile given the perceived risk, considerable advantages in terms of cost effectiveness must be proven.

Several studies have proposed architectures for autonomous on-orbit servicing of specific space missions and demonstrated potential improvements in terms of cost or cost effectiveness.[1–3] However, no advantages have been proven to date that outweigh the perceived risk and cost uncertainty. In addition, no research has explored in a systematic and quantitative way the advantages of on-orbit servicing for space missions.

The goal of this paper is to propose research directions that can help draw general conclusions about the cost effectiveness of on-orbit servicing. On-orbit servicing will here refer to any on-orbit activity, including refueling, performed after a system has become operational, for the purpose of extending the operational life of the system, or modifying some of its components, or giving it new functionality. Previous work on on-orbit servicing was limited to case studies. A first approach to the problem is to expand on these studies by developing general cost-effectiveness metrics and models and systematically applying them to a large trade space of space missions and servicing infrastructures. The first half of the paper explores this direction and analyzes typical results. This analysis unveils two fundamental limitations. First, the traditional approach is impaired by the lack of cost models for a servicing infrastructure. Second, and more important, this approach does not account for the options that on-orbit servicing gives to decision makers to react to the resolution of uncertainty.

The analysis of these limitations motivates the definition of a new perspective on on-orbit servicing. The second half of this paper proposes to make two fundamental steps forward; First, it argues that the cost of servicing should be separated from its value. Second, it defines a method to quantify the flexibility provided by on-orbit servicing to space systems. Based on this perspective, a framework that accounts for flexibility in modeling the value of on-orbit servicing for space systems is proposed. Several applications of this framework will be analyzed in a companion paper.[4]

Cost-Effectiveness of On-Orbit Servicing: Traditional Approach

In this section, we propose to develop a general model that can systematically apply the approach used by previous case studies to a wide trade space of space missions and servicing infrastructures. The analysis of the results produced by this traditional approach will help us define new directions for research on on-orbit servicing. For this purpose, the limits of the trade space will first be drawn, and general metrics will be defined with which to quantify cost effectiveness. Then a model to estimate systematically cost effectiveness on the trade space will be developed. Finally, analysis of the validity and limits of the model on a typical example will be presented.

Trade Space

We are interested in developing a simplified model of the serviced and servicing missions that captures only the most meaningful cost and performance trends. The trade space is made up of mission types, on the one hand, and maintenance types, on the other hand. For each mission, the goal is to evaluate which maintenance type is most cost effective and under which conditions servicing can be expected to improve cost effectiveness significantly.

Missions

Mission types that are potential customers for on-orbit servicing can be divided into five representative families, designated A–F and summarized in Fig. 1. These families are broadly defined, and a general trade space should have many dimensions within each mission type. The first candidate A is a high value asset, for which replacement costs are prohibitive; an example is the International Space Station. Family B captures new types of space missions that would be made possible by on-orbit servicing, such as highly maneuverable satellites. Family C comprise low-Earth-orbit (LEO) constellations, for which proximity to Earth and economies of scale can be exploited. Whereas family D is one to three geostationary-Earth-orbit (GEO) satellites, family E is made up of the whole GEO ring, exploiting their design similarities and common orbital plane. Finally, the hardest customer to service would be a system comprising several orbital planes at various inclinations and altitudes, family F.

	A. High value asset	B. New Missions	C. LEO constellation	D. GEO constellation	E. Whole GEO ring	F. Multiple planes
1. Without on-orbit servicing						
2. Disposable servicer						
3. Satellites roudtrip to depot						
4. Servicer roudtrips to depot						
5. Refuelable servicer						

Fig. 1 Trade space: missions and infrastructures.

For each mission, a minimum set of parameters is necessary to estimate the baseline cost and performance of the mission. Important factors for cost at the constellation level include the number of orbital planes, number of satellites per plane, altitude, inclination, plus the development and yearly operations costs. At the satellite level, we consider at a minimum failure rate, mass, and production cost. An example of the minimum information for measuring mission performance is the required number of satellites per plane for the mission requirements to be met.

The focus of this paper is on the cost and value of on-orbit servicing rather than the technology necessary to make it feasible. Thus, in the rest of this section, we will assume that the appropriate autonomous rendezvous and capture (AR&C) equipment[5] is added to the satellite payload cost and mass whenever existing satellites are considered.

Infrastructures for Servicing

Changing orbital planes in Earth orbit requires high incremental velocity. If the altitude is low enough and sufficient time is allowed, nodal regression can be used to get around this problem. Otherwise, it is reasonable to restrict the trade space to launching servicing material separately into each serviceable orbital plane.

The description of a maintenance infrastructure comprises parameters that are technology dependent and parameters that can be optimized for a given mission. Technology-dependent parameters include production and launch rates, maximum time allowed for a maneuver, probability of a catastrophic failure when attempting an AR&C maneuver, and minimum servicer dry mass (corresponding to the AR&C payload with the power and structure to support it). The parameters that can be optimized depend on the kind of infrastructure. We define five broad families of maintenance infrastructures as summarized in Fig. 1. Whereas each family represents a different on-orbit servicing concept, there are many possible servicing architectural designs within each category. The first family corresponds to the baseline case: It has no on-orbit servicing. Satellites are replaced if they fail, run out of fuel, or need a hardware upgrade. The second family corresponds to the minimal servicing capability: The orbital replacement units to be delivered are launched onboard a servicer vehicle that is disposed of after it has delivered all its cargo. The third family corresponds to satellites traveling to an orbital depot or station. This could be envisioned if the downtime in mission availability is judged acceptable and if the satellites have not yet failed or run out of fuel; however, it adds significant complexity to the mission design. The fourth family, called servicer roundtrips, corresponds to the case when servicer spacecraft can be reused (but not refueled) by going to a depot/station and loading new cargo, until they have just enough fuel to deorbit. Finally, the fifth family has a refuelable servicer that can also be refueled at the station/depot, to multiply the number of servicing missions it can perform over its life. At this point, no particular assumption is made on the design of the depot/station.

Each of the preceding schemes 1–5 can be carried out either on an on-demand or on a scheduled basis. In the scheduled case, a servicing period is defined. This method is well suited to service components with quasi-deterministic time to failure/depletion, such as propellant for station keeping. In the on-demand case, components are serviced as they fail. This method is well suited to components with a probabilistic time to failure/depletion.

Definition of Cost Effectiveness

To evaluate the cost effectiveness of each of schemes 1–5 for each of the missions A–F, we need to define a general measure of cost effectiveness. The generalized information network analysis (GINA) methodology[6] is a meaningful tool to evaluate space mission cost effectiveness. It relies on the premise that most satellite systems can be represented as information transfer networks. Their quality of service is measured by four capabilities: isolation, rate, integrity, and availability of the information transferred. Different architectures are compared by means of a cost-per-function (CPF) metric. The CPF amortizes the total lifetime cost over all satisfied users of the system during its life. The total lifetime costs include costs to initial operating capability (IOC) as well as operation costs and expected failure compensation costs. The function is the expected total number of times the system will meet the minimum user requirements. These requirements differ from mission to mission, but can always be expressed in terms of the four capability metrics. The CPF is, therefore, a meaningful quantitative measure of cost effectiveness.

A servicing mission is not an information transfer network in itself. In most cases, it is a mass transfer network. However, the final goal of the mass delivery is to enhance the capabilities of the serviced mission. Therefore, on-orbit servicing of an information transfer network is defined as cost effective if it reduces its CPF compared to a traditional approach. The major effects that servicing can have on the CPF are as follows: First is a decrease in the initial cost because there is less need for redundancy, a possible reduction in the bus mass, and/or smaller fuel tanks can be built into each satellite. Second is a change in the failure compensation costs; these costs include replacements and servicing, as well as satellite and servicer replacement after potential catastrophic failures. Third is a change in mission performance, because the probability to meet the requirements depends on the failure compensation scheme, but also because satellites can be upgraded for improved mission performance. Through failure compensation costs and mission performance, risk assessment is embedded in the mission CPF.

Model to Estimate Cost Effectiveness

To estimate the potential improvements in the CPF of a given space system, it is necessary to have a general model that captures the main effects of servicing on both cost and function. The impacts on cost include the design changes as satellites are made serviceable, the cost of the chosen serviceable infrastructure, and the failure compensation costs. The impact on mission performance should take the risk of servicing into account, as well as the impact of possible upgrades. This section proposes such a general model.[7]

Serviceable Spacecraft

Let us first consider the impact of serviceability on spacecraft mass. Mass differences will translate through cost models into cost differences. The first impact is the cost savings from refueling. A refuelable spacecraft has to be designed for the maximum total velocity increment ΔV_d between two refueling operations. The fuel mass it has to carry increases exponentially with its design ΔV through the rocket equation[8]:

$$M_{\text{fuel}} = M_{\text{dry}}^{\text{tot}}[\exp(\Delta V_d/gI_{\text{sp}}) - 1] \tag{1}$$

Increasing the fuel mass also increases the propulsion system's dry mass. For a given type of spacecraft design, it is common to define three constants relating the various masses as follows: a propulsion dry mass factor f_p such that $M_{\text{dry}}^{\text{prop}} = f_p M_{\text{fuel}}$, a structure's mass factor f_{st}, and the fraction ϵ of the total fuel mass that is carried on the spacecraft at launch, so that $M_{\text{struc}} = f_{\text{st}}(M_{\text{dry}}^{\text{tot}} + \epsilon M_{\text{fuel}})$. The total spacecraft dry mass and launch mass then behave as the following functions of the design ΔV:

$$M_{\text{dry}}^{\text{tot}} = M_{\text{dry}}^{\text{base}} \Big/ \{1 - (f_p + f_{\text{st}} f_p + \epsilon f_{\text{st}})[\exp(\Delta V_d/I_{\text{sp}}g) - 1]\} \tag{2}$$

$$M_{\text{launch}} = M_{\text{dry}}^{\text{tot}}\{1 + \epsilon[\exp(\Delta V_d/I_{\text{sp}}g) - 1]\} \tag{3}$$

Another main advantages of a repairing capability is to decrease the initial spacecraft cost by designing space systems for a shorter design life. The cost savings incurred by reducing the required design lifetime have been studied[9] by analyzing the impact of the design lifetime requirement on the various spacecraft subsystems, as well as on the overall redundancy. It was concluded that a cost penalty is incurred by designing systems for a longer design lifetime. A linear fit to the final results gives the typical relative dry mass impact to design for T_D years instead of an arbitrary reference lifetime of three years in the form

$$M(T_D)/M(3) = 1 + \kappa(T_D - 3) \qquad \text{with} \qquad \kappa \approx 2.75\%/\text{year} \tag{4}$$

Thus, a spacecraft designed for 10 years has approximately 20% more dry mass than the same system designed for 3 years.

Finally, maneuver modeling is key to sizing the spacecraft that moves in the servicing process. There are five important types of servicing maneuvers. First, changes in inclination are necessary at least to correct for launch inaccuracies; for an impulsive burn, the incremental velocity is $\Delta V_{\text{inc}} = 2V_0 \sin(\Delta i/2)$. Second, transfers from an orbital altitude a_0 to another $(a_1 = \alpha a_0)$ using a simple Hohmann transfer require

$$\Delta V_H/V_0 = \left|\sqrt{2\alpha/(1+\alpha)} - 1\right| + 1\Big/\sqrt{\alpha}\left|\sqrt{2/(1+\alpha)} - 1\right| \tag{5}$$

A Hohmann transfer can also be used for deorbiting any spacecraft at end of life. Third, nodal regression can be used to transfer a servicer between different orbital planes in the constellation. This is done by changing the semimajor axis of the servicer orbit and waiting for the difference in nodal regression rate to cancel out the plane difference. The impulsive incremental velocity required depends on the time allowed. Fourth, phasing maneuvers within one orbital plane are useful for a servicer to go from a satellite to the next; circular phasing can be performed by temporarily raising the apogee. The impulsive incremental velocity for such a maneuver is to first order inversely proportional to the time allowed:

$$\frac{\Delta V_{\text{ph}}}{V_0} = 2\left|1 - \sqrt{2 - \left(\frac{I(\Delta T_{\max}/T_0 + \phi/2\pi)}{I(\Delta T_{\max}/T_0 + \phi/2\pi) - \phi/2\pi}\right)^{\frac{2}{3}}}\right| \tag{6}$$

$$\frac{\Delta V_{\text{ph}}}{V_0} \approx \frac{|\phi|}{2\pi}\frac{T_0}{\Delta T_{\max}} \qquad \text{for} \qquad T_0 \ll \Delta T_{\max} \tag{7}$$

Finally there are fine maneuvers for the AR&C proximity and final phases: ΔV_f depends on the AR&C control algorithms. A conservative $\Delta V_f = 120$ m/s will be used. (This corresponds to a 30% margin in most cases.)

Servicing Infrastructure

For all infrastructure types but type 3, the servicer's maneuver scheme can be described in terms of three integers, N, L, and K: N is the number of satellites to visit before loading more cargo. L is the number of sets of N satellites to visit before the servicer runs out of fuel. K is the number of times the servicer is refueled to service $N \times L$ more satellites. Thus, the total number of satellites per servicer is $N_{\text{sps}} = NLK$. The servicers fuel mass can then be shown to be

$$M_{\text{fuel}}^{\text{servicer}} = AM_{\text{dry}}^{\text{servicer}} + BM_C$$

where

$$A = \exp[L(N-1)v_P + 2Lv_H + v_D] - 1$$

$$B = \frac{\exp[L(N-1)v_P + 2Lv_H] - 1}{\exp[(N-1)v_P + 2v_H] - 1} \times \left[\frac{\exp(Nv_P) - 1}{\exp(v_P) - 1}\exp(v_H) - N\right]$$

The servicer dry mass is finally a function of its maneuvering scheme and its specific impulse as follows:

$$M_{\rm dry}^{\rm servicer} = \frac{M_0 + f_{\rm stC} N M_C + (f_p + \epsilon f_{\rm st} + f_{\rm st} f_p) B M_C}{1 - A(f_p + \epsilon f_{\rm st} + f_{\rm st} f_p)} \tag{8}$$

Performance

Mission performance, defined as the instantaneous probability to meet the minimum requirements, depends on the failure and maintenance scheme of the constellation and on the level of technology insertion made possible by servicing. Nondeterministic satellite failures are usually modeled by a failure rate λ such that the probabilities P to be in various failed states follow a Markov process: $\dot{P} = A_M P$ (Ref. 8). Similarly, on-demand maintenance can be modeled by a repair rate μ in the Markov matrix A_M. Quasi-deterministic satellite failures, such as fuel consumption for station keeping, can be taken into account by incorporating a change in the Markov model initial conditions at the expected date of failure. To account for scheduled maintenance while keeping a Markov process, each state in the model is divided into $(1 + N_{\rm sched})$ substates, going from never maintained until maintained $N_{\rm sched}$ times. At each scheduled time, the Markov matrix is changed to incorporate new repair rates for transition to the next substate.

The expected number of times that a system has to be maintained (serviced or replaced) from a state k is then

$$N_k = \int_0^{T_H} \mu_k P_k(t)\,\mathrm{d}t \tag{9}$$

The mission's function, that is, its number of satisfied users, is the result of the Markov model on the one side and a market model for the information transfer network on the other side. Let $\mathcal{M}_j(t)$ be the instantaneous number of satisfied users per unit time in the operational state j. Upgrades are taken into account by modeling the effect of improved performance on this market level $\mathcal{M}_j(t)$. Then the function is

$$Fn = \sum_j \int_0^{T_H} P_j(t)\mathcal{M}_j(t)\,\mathrm{d}t \tag{10}$$

Cost Modeling

Spacecraft economic models are not adequate for the servicing infrastructure, which differs from the historical data used to derive them. Combining several cost models can, to some extent, mitigate this problem. Three models, based on cost estimation relationships (CERs), are publicly available: the unmanned spacecraft cost model,[8,9] the small satellite cost model,[10] and a rule-of-thumb industry model.[11] For the latter, the theoretical first unit cost (TFU) is simply proportional to dry mass ($C_{\rm TFU}/M_{\rm dry}$ = \$77,000/kg) and the development costs scale with $C_{\rm TFU}$ by a technological factor F. Cost uncertainties are estimated by combining standard deviations of the CERs.

The model is now ready to be applied throughout the trade space matrix (Fig. 1). The model as developed here can be applied to refueling, repairing, or upgrading. However, in this paper we have chosen to focus on refueling as this is the simplest case to analyze. One only has to consider life extension, which can be valued within the framework of the original mission.

Analysis of Typical Refueling Case and Conclusion

Problem Statement

Let us consider refueling a LEO communication mission, taking mission parameters on the example of the Iridium constellation,[12] hereafter called LEO66. For a communications constellation, a good metric of performance is the number of billable minutes provided by the system, where a billable minute is a minute of communications with the required isolation, rate, integrity, and availability.[13] As seen by decision makers before the launch of the mission, the market $\mathcal{M}_{th}$ was expected to grow quasi linearly as described in Ref. 13 and shown in Fig. 2.

Table 1 LEO66 constellation assumptions

Parameter description	Name	Value	Source
Constellation altitude	a_0	780 km	Ref. 12
Satellite launch altitude	a_{00}	650 km	Ref. 12
Number of orbital planes	N_p	6	Ref. 12
Number of satellites per plane	$N_{\rm spp}$	12	Ref. 12
Satellites per plane to meet the requirements	$N_{\rm nspp}$	11	Ref. 12
Number of on-orbit spares	$N_{\rm spares}$	1	Ref. 12
Satellite specific impulse	$I_{\rm spS}$	320 s	Chemical
Satellite mean time to failure	$1/\lambda$	9.18 yr	Ref. 12
Time to replace one satellite after failure	$1/\mu$	3 mo	Ground spare available
Satellite development cost factor	F_S	4	State of the art
Discount factor	d	7%/yr	Observed
Ballistic coefficient	BC	50 kg/m^2	Assumed

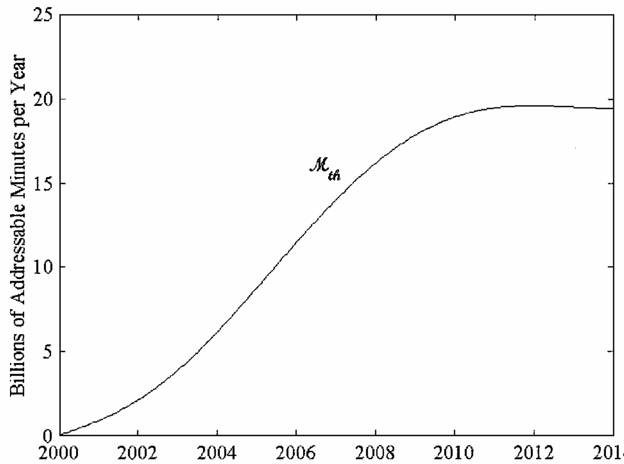

Fig. 2 LEO66 market forecast, adapted from Ref. 12: $\mathcal{M}_{th}$ is the theoretical market level in terms of minutes of communication that can be offered with the required isolation, rate, integrity, and availability each year.

For this case study, we will compare three design alternatives, all aimed at achieving an effective mission lifetime of 16 years. The first alternative is to keep the spacecraft required lifetime of eight years and replace the satellites after eight years. The second is to keep the spacecraft required lifetime of eight years and refuel the satellites after eight years. In this case, the most natural choice of infrastructure is type 2, where servicers are launched with their cargo into each orbital plane. The refueling can be scheduled because fuel consumption is quasi deterministic. The main servicing design parameter is the number of satellites to be refueled by each servicer, which creates six subcases. The third alternative is to design the spacecraft propulsion system for 16 years of operation. This increases the cost to IOC, but requires no scheduled maintenance. In all cases, satellites are replaced as they undergo random failures with a failure rate λ = 1/mean mission duration. Tables 1 and 2 summarize the numerical assumptions.

Results in the Baseline Case

In the baseline case, the satellites' propulsion requirements are dominated by the fuel to deorbit, which accounts for almost one-third of their total mass. Servicing makes it possible to do away with this mass by refueling at end of life. The CPF is here the lifetime cost per billable minute. Figure 3 is a comparison of several servicing schemes with the two traditional alternatives; total costs are indicated in place of CPF because the function adds up to almost

Table 2 LEO66 servicing assumptions

Parameter description	Name	Value	Source
Probability to crash at AR&C	P_C	0.1%	Free parameter
Servicer specific impulse	I_{sp0}	320 s	Chemical
Servicer launch altitude	a_1	700 km	Assumed
Error in servicer inclination	Δi	1 deg	Conservative
Time allowed for approach	ΔT_{max}	2 days	Short while allowing low ΔV
Final approach ΔV	ΔV_f	120 m/s	Conservative
Attitude control ΔV	ΔV_{yr}	30 m/s · yr	Typical LEO
Servicer development cost factor	F_0	5	New technology
Structures mass factor	f_{st}	0.15	Aggressive servicer design assumed
Propulsion dry mass factor	f_p	0.1	Optimistic
Structures mass factor for cargo	f_{stC}	0.2	Estimated

Fig. 3 LEO66 baseline case costs for three design options: 1) replace satellites every 8 years, 2) refuel the satellites, with various possible number of servicers per satellite (minimum cost obtained for six satellites per servicer), and 3) design satellites for 16 years of station keeping. Vertical are uncertainty brackets. Without servicing, designing for 16 years is optimal for a total lifetime cost of $4 billion; if servicing were free (no refuel cost), it would reduce the total lifetime cost to $3.6 billion and, thus, become the optimal solution.

the same amount for all cases. Within the cases with on-orbit servicing, the minimum lifetime cost is obtained by using one servicer for every six satellites, that is, two servicers per orbital plane. First suppose that we want to extend the mission life after eight years as a response to unexpected market growth; then refueling after eight years is hardly more cost effective than replacing the whole constellation. If 16 years was the initial desired lifetime, refueling is even less interesting. The satellite's ΔV requirements are low enough that they can carry fuel for 16 years without significantly increasing lifetime costs.

Sensitivity studies[7] show that refueling becomes more cost effective when the failure rate is very low (four times the baseline reliability), so that running out of fuel becomes the major cause of failure, or when the altitude is very low (400 km), so that carrying fuel for life becomes too expensive. However, in all cases, the cost advantages remain smaller than the cost uncertainty. The cost effectiveness of refueling is also sensitive to the probability P_C of a catastrophic failure when attempting AR&C. Increasing P_C both decreases the mission performance and increases its failure compensation costs because both the satellite and the servicer have to be replaced after the failure. Figure 4 shows how, for P_C higher than 1%, the mission performance (probability to have the minimum number of satellites to meet the requirements) drops below 70% just after each refueling event. This is an unacceptable risk when market capture is at sake.

Conclusion on the Traditional Evaluation of Servicing

This example is typical of the results that can be obtained with the traditional approach to on-orbit servicing. Although situations can be found for which servicing proves cost effective,[2] the cost advantage most often remains smaller than the cost uncertainty.

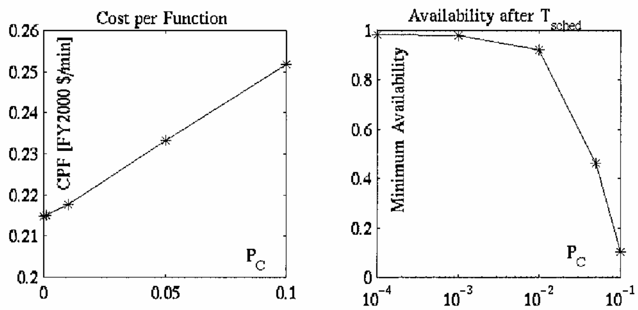

Fig. 4 LEO66: sensitivity to P_C, performance and cost per billable minute where T_{sched} is time of scheduled refuel.

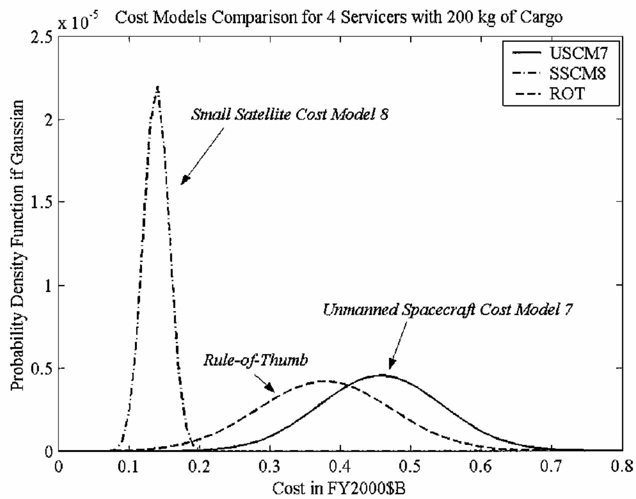

Fig. 5 Inadequacy of cost models for servicer, example (design for this servicer vehicle based on a scaled version of Lincoln Laboratory experiment).

This uncertainty is made up of two parts: the uncertainty in the constellation cost and the uncertainty in the servicing cost.

The uncertainty in the constellation cost arises from the standard deviation in the CERs used. CERs are designed from historical data to model the way costs depend on various design parameters. Their standard deviation reflects an uncertainty in absolute cost, but the relative cost difference between various designs are usually well captured. Thus, cost uncertainty is only a minor limiting factor for the conclusions in terms of relative constellation costs. However, a servicer spacecraft would be very different from historical satellites, so that cost models would not be directly applicable. The different subsystems for the servicer spacecraft herein modeled actually fall into the applicability ranges for different cost models, making the use of these models unpractical and doubtful, as shown by Fig. 5.

Cost models are useful in a relative sense because they accurately capture the design parameters that drive the cost of a satellite. This reasoning is not valid any more for a servicing infrastructure, for which the design and cost drivers are likely to be different. Even if a cost model valid in a relative sense for servicing architectures was developed, one could not add those costs to the cost of a satellite constellation unless their absolute scaling was accurately captured. Therefore, the uncertainty in the servicing price is here a limiting factor even in a relative sense. This makes any definitive conclusion about the cost effectiveness of servicing impossible. This has also been one of the major problems faced by previous work on on-orbit servicing.

However, note that at least one conclusion can be drawn by considering Fig. 3. If servicing were free, the lifetime costs for a serviceable design would be 10% lower than for a 16-year design. Whatever the price of servicing, the potential cost savings for this case cannot exceed this limit. This is an indication of the intrinsic value of servicing. It suggests a new perspective on on-orbit servicing, which we will now explore in more detail.

General Framework for the Value of Servicing Under Uncertainty

Separating Servicing Value from Servicing Cost

With the traditional approach considered in the preceding section, whether on-orbit servicing proves more interesting than traditional methods is the result of a tradeoff between two main effects that can be summarized as the cost savings from servicing vs the price space missions pay for servicing. These concepts are shown in Fig. 6.

The cost savings from servicing depend mainly on the satellite design and the elements to be serviced. For any given space mission, they can be estimated with reasonably good accuracy on a large trade space of designs, as shown, for example, in Ref. 11.

On the other hand, the price to pay for servicing depends on two factors, both of which are highly uncertain. First, the cost of servicing depends on the assumptions, design choices, and cost models for the servicing architecture and the satellite design changes. Moreover, cost models are not adequate for such an infrastructure, so that even these limited conclusions bear a very high uncertainty. Furthermore, the price of servicing is not necessarily equal to its cost. The price of servicing is the amount of money that will be charged to a customer's space mission for a given service, whereas the cost of servicing is the amount of money that the servicing infrastructure has to spend to provide this service. The price of servicing depends not only on the cost, but also on the infrastructure development policy: The cost of the whole architecture could be paid by one space mission, or shared among several missions. Even better, an infrastructure could be developed by a governmental organization, so that only the marginal cost of servicing would be charged to individual space missions.

Estimating the cost savings from servicing separately from the servicing price, therefore, presents several major advantages. The estimate can serve as a good indicator of the maximum price that a space mission would be willing to pay for servicing; missions for which servicing could make sense would be the ones for which its value is significant compared to total mission value. In addition, separating value from cost would significantly reduce the uncertainty in the conclusions drawn.

Because no infrastructure for on-orbit servicing exists yet, results that would give some guidance as to what types of technologies to develop, what space missions to target, and what cost cap not to exceed would be very valuable. We propose to seek such results by studying the value of servicing separately from its cost.

Accounting for the Flexibility Provided by On-Orbit Servicing

Separating the value of servicing from its cost can be an important step forward for on-orbit servicing research. Here we propose to go even further by accounting for flexibility in estimating this value. Contrary to what has been implicitly assumed by the traditional approach, the value of servicing is not limited to potential cost savings. Having the ability to access space assets represents a source of flexibility. We define flexibility of a design as a property of a system that allows it to respond to changes in its initial objectives and requirements (both in terms of capabilities and attributes) occurring after the system has been fielded, that is, is in operation, in a timely and cost effective way.[14] On-orbit servicing can provide decision makers with the options to refuel and repair spacecraft for life extension, upgrade them for improved performance, or change their payload to perform a new mission. The decision can depend on the resolution of parameters that were uncertain at the time of launch. Market demand and technology levels, which often vary on timescales shorter than a space mission lifetime, are two typical examples of such a parameter.

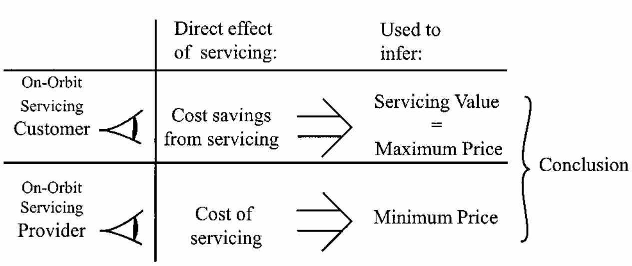

Fig. 6 Cost, cost savings, and price associated with on-orbit servicing.

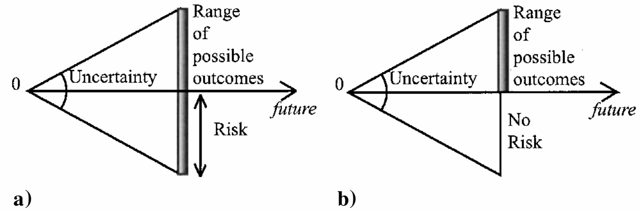

Fig. 7 Cone of uncertainty: a) without options: uncertainty means risk and b) with options: uncertainty means higher expected outcome.

A good way to capture conceptually the importance of having options is the notion of cone of uncertainty proposed by real options theory[15] to represent the uncertainty in revenues (Fig. 7). Decision makers consider the future as seen from the apex of the cone (present). The further they look into the future, the more uncertainty there is associated with their forecast, as is represented by the cone with its opening angle being a measure of the level of uncertainty. If no option is available, then an increasing uncertainty translates into an increasing probability of a negative outcome; thus, uncertainty is directly mapped into risk. However, if options are available to react to uncertainty, then negative outcomes can be avoided, and a higher uncertainty translates into a higher expected outcome.

Thus, for flexible designs, uncertainty is not a synonym for risk any more; it can even become a source of value. Only by accounting for the flexibility of decision makers in reacting to the resolution of uncertainty, for example, avoiding negative outcomes, can this important advantage of on-orbit servicing be evaluated. This is what we propose to do in the rest of this paper.

Example of Options Available to Space Missions

Five examples of options are 1) the option to abandon an on-orbit asset if operational costs are deemed too high compared to mission benefits, which is available to all space missions and has been exercised in the past; 2) the option to replace, for life extension or after a random failure, also available to all space missions; 3) the option to service for life extension, which would only be available to a serviceable mission; 4) the option to upgrade, which would be available to serviceable spacecraft to avoid technological obsolescence and to improve their performance with respect to their baseline requirements; and 5) the option to modify, which would make it possible for serviceable spacecraft to respond to changes in their requirements. In each of the five cases, the decision to exercise the option, for example, to abandon the mission, or to service it, does not have to be set before launch, but will rather depend on the resolution of one or several uncertain parameter(s). In the future, decision makers should be expected to take the optimal decision given the state of the environment they observe at a given time. Traditional space mission valuation, which relied on net present value calculations, underestimates the value of servicing in the presence of uncertainty by assuming a set, and therefore, suboptimal, sequence of events would occur. The rest of this paper proposes to go one step further by developing a framework to quantify the value of the flexibility offered by on-orbit servicing. To deal with probabilistic sequence of mission events, this framework builds on and expands ideas developed by decision tree analysis and real options theory. These two frameworks capture the value of managerial flexibility associated with the availability of options to make decisions in the future.[16]

Basic Elements of the Framework

This subsection defines the basic elements required to describe fully a situation where options are available to adapt to the resolution of uncertainty. These elements are the building blocks of the framework, which are common to any option valuation. Their particular values and behaviors must be defined for each case under study as the first step in the valuation process.

Uncertain Parameter X

In a world of certainty, options have no value. Therefore, the cornerstone of any option valuation is modeling the uncertainty in the future states of nature. At least one uncertain parameter is required, which can be modeled as an instantaneous stochastic process X. This is a generalization of decision tree analysis to account for continuity in the possible states of nature. If several parameters are uncertain, X can be taken to have several dimensions: $X = (X_1, \ldots, X_n)$.

In the following description, we will always assume that the uncertain parameter follows a Markov process: The distribution of X at time $t > t_0$ knowing the path $X([0; t_0])$ is only a function of $X(t_0)$. In other words, only the latest information about X is relevant. We will define $p_t(x|x_0)$ as the probability density function of X at t knowing $X(t_0) = x_0$. This assumption greatly simplifies the valuation process without sacrificing any important effect of flexibility value. It is reasonably valid for sources of uncertainty such as market dynamics, military contingency location, or random failures.

Example: In the case of commercial missions with uncertain revenues, X can be defined as the ratio of the actual revenues per unit time $\mathcal{M}(t)$ to the revenues rate given by the market forecast $\mathcal{M}_{th}(t)$. A typical assumption used by real options theory is the geometric random walk process, which is a good description of the behavior of stocks values. Under this assumption, if $X(t)$ is known, then $x = X(t + \tau)/X(t)$ has a log-normal probability density function with mean $e^{\alpha\tau}$ and variance $\sigma\sqrt{\tau}$:

$$p_{t+\tau}(x|x_t) = p_\tau(x) = \frac{1}{x}\frac{1}{\sqrt{2\pi}\,\sigma\sqrt{\tau}}\exp\left\{-\frac{[\ln(x) - (\alpha - \sigma^2/2)\tau]^2}{2\sigma^2\tau}\right\} \tag{11}$$

Mission Horizon T_H

To evaluate alternative mission scenarios on a fair basis, their costs and benefits should be compared over the same elapsed time $[0; T_H]$, where T_H will be called the mission horizon. In the most general case, T_H can be taken to be infinite.

Decision Points T_k

The valuation of options relies on the existence of decision points, which are instances in the future when decision makers will have the option to choose between several alternative decisions.

Modes of Operation m

At each decision point, alternative decisions can be represented as several possible modes of operation.[16] Typical examples of modes of operation that could be available to space missions are abandoned, operational in its initial design, or operational with a particular design modification. We will mark the value of any variable Y in mode of operation (m) by a superscript: $Y^{(m)}$, the value of any variable linked to a switch from mode (n) to mode (m) by $Y^{(n\to m)}$, and any variable linked to a history of successive modes of operation $(m_1, m_2, \ldots, m_n)$ by $Y^{(m_1, m_2, \ldots, m_n)}$.

Utility Metric U

The utility metric is a generalization of the notion of revenues for commercial missions. For $t_1 < t_2$, here $U([t_1; t_2])$ is a measure of the aggregated benefits from the mission over the time interval $[t_1; t_2]$. These benefits are not necessarily monetary.

In most noncommercial cases, there can be several choices for the utility metric. For example, the benefits from a space-based radar mission could be the total number of square kilometers protected, or the total time a given area has been protected. The right choice should be the one that most describes what is of importance to decision makers. The utility metric must be such that, among several architectural alternatives with the same cost, decision makers would choose the one offering the highest utility function.

Utility Rate u

For any meaningful model, mission benefits will be an increasing function of time. In any given state of the system, there will be an instantaneous utility rate u such that $u = \mathrm{d}U/\mathrm{d}t$. If this utility rate depends only on the present state of the system, then we can say that utility is time additive, that is, total utility is the sum of past and future utility.

Example: For a commercial mission with uncertain revenues, the utility rate is directly proportional to the uncertain parameter: $u(t) = \mathcal{M}(t) = \mathcal{M}_{th}(t)\,X(t)$.

Cost Metric C

The cost metric is the sum of all of the expenses associated with the mission and its options. For $t_1 < t_2$, here $C([t_1; t_2])$ is the present value of the aggregated costs of the mission over the time interval $[t_1; t_2]$. Two types of discount rates for costs and revenues are necessary for the valuation of nonfinancial options such as space systems:

1) Amounts that are not subject to uncertainty, or that have a twin-security on the stock exchange, can be discounted at the risk-free interest rate r. A twin security is defined as a security that is traded on the stock exchange and whose behavior mimics the value of the underlying amount considered.

2) Amounts that are uncertain and not linked to a twin security can have an internal rate of return $\alpha_p \neq r$ (Ref. 17). These amounts must be discounted at a rate $r + p$, where p is a risk premium that depends on the level of risk in the project.[18] It is then convenient to define $\alpha = \alpha_p - p = r + \delta$ and to describe uncertain amounts as increasing with a rate of return α. Uncertain amounts can then be discounted at the risk-free interest rate r.

In a real situation, the risk-free interest rate r, the project's internal rate of return and the risk premium are all functions of time, in a fashion that cannot easily be predicted. A convenient approximation is to assume that they remain constant. This is already an improvement over traditional valuation, which uses only one discount rate.

Value Metric V

The value of a mission is a tradeoff between its utility and costs:

$$V([t_1; t_2]) = f\{U([t_1; t_2]), C([t_1; t_2])\} \tag{12}$$

It should be chosen such that among several alternatives decision makers would systematically choose the one that maximizes the future expected value. Typical examples of possible value functions are given in Table 3.

Example: For a commercial mission, the value metric is a linear function of the cost and utility:

$$V = U - C \tag{13}$$

This linearity has the property of making value time additive. This means in particular that maximizing future value is equivalent to maximizing lifetime value.

Decision Model

The decision model describes how the decision should be taken at a decision point T_k as a function of the current mode of operation and the current state of the uncertain parameter $x_k = X(T_k)$. If the value metric has been correctly defined, the decision will be to chose the mode of operation that maximizes future mission value. Let us call $EV^{(m)}_{\geq k}(x_k)$ the expected value of the mission after T_k knowing the current state of the uncertain parameter $x_k = X(T_k)$ and assuming that mode of operation m is chosen:

$$EV^{(m)}_{\geq k}(x_k) = E\left\{V^{(m)}([T_k; T_H])\,|\,X(T_k) = x_k\right\} \tag{14}$$

Table 3 Examples of possible value functions

Value functions	Commercial	Military	Scientific
Utility U	Revenues	Utility function	GINA Function
Mission value V	$U - C$	U/C	$U/C = 1/\mathrm{CPF}$
	$(U - C)/C$	maximum(U; $C <$ cost cap)	maximum(U; $C <$ cost cap)
		minimum(C; $U > U_{\min}$)	minimum(C; $U > U_{\min}$)
Source of uncertainty	Revenues	Theater location	Scientific return

The total mission value is $V = EV^{(0)}_{\geq 0}(x_0)$. The optimal decision mode $\hat{m}_k(x_k)$ at T_k is given by

$$\hat{m}_k(x_k) : \max_m \left[EV^{(m)}_{\geq k}(x_k) \right] \tag{15}$$

This optimal choice is a function of x_k, which is unknown at the start of the mission. Thus, for each mode of operation m, this decision process defines the set $I^{(m)}_k$ of values of x_k of the uncertain parameter for which the mode m should be chosen.

Black–Scholes Equation[19] Example

Before describing the valuation process in its general setting, it is useful to consider the simplest possible example: the simple option on life extension.

Black–Scholes Equation[19]

Consider a space mission designed for $T = 10$ year with the option to be serviced at T, thus increasing its lifetime up to $T_H = 20$ years. At T, decision makers will choose between two modes of operation: 1) not operational or 2) operational. Choosing operation 1 would incur the cost $C^{(2\rightarrow 1)} = 0$ and choosing operation 2 would incur the cost $C^{(2\rightarrow 2)} = C_S + C_0 = E = \$100$ million, where C_S is the cost of servicing and C_0 the cost to operate the system from $t = T$ until $t = T_H$. We call this cost E because it is similar to the exercise price of a stock option: Whereas a stock trader can buy an option on a stock, here the decision maker can buy an option on a life extension.

This option could, for example, be interesting for a commercial mission, for which the revenues after time T are uncertain at the time of launch; let us call these S because they are similar to the stock price for a stock option. S is a stochastic process, and at time $t = 0$ its value is observed to be $S_0 = \$125$ million. Staying operational after T is interesting only if it incurs more revenues than expenses, that is, only if $S > E$.

If the market behaves as a stock, its rate of change can be described as a diffusion process (geometric random walk) with volatility σ as commonly used by real options theory:

$$\frac{\mathrm{d}S}{S} = \alpha\,\mathrm{d}t + \sigma\,\mathrm{d}B \tag{16}$$

With these assumptions, the revenues at T follow the log-normal probability density function:

$$p(S) = \frac{1}{\sqrt{2\pi}\,\sigma\sqrt{T}}\frac{1}{S}\exp\left\{-\frac{\left[\ln(S/S_0) - (\alpha - \sigma^2/2)T\right]^2}{2\sigma^2 T}\right\} \tag{17}$$

The utility function is the revenues, and the value is simply $V = U - C$. This value being time additive, the value of the option at $t = 0$ is simply the expected present value incurred after time T, which for risk-neutral investors is

$$E\{V\} = \int_0^E 0 \times p(S)\,\mathrm{d}S + \int_E^\infty e^{-rT}(S - E)p(S)\,\mathrm{d}S \tag{18}$$

With the change of variables

$$x = \left[\ln(S/S_0) - (\alpha - \sigma^2/2)T\right]\Big/\sigma\sqrt{T}$$

and $y = x - \sigma\sqrt{T}$, this yields

$$V = e^{(\alpha - r)T} S_0 N(d_1) - e^{-rT} E N(d_2) \tag{19}$$

where N is the cumulative normal distribution function

$$N(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}}\exp\left(\frac{t^2}{2}\right)\mathrm{d}t = 1 - N(-x)$$

$$d_1 = \frac{\ln(S_0/E) + (\alpha + \sigma^2/2)T}{\sigma\sqrt{T}}, \qquad d_2 = d_1 - \sigma\sqrt{T}$$

With $\alpha = r$, Eq. (19) becomes identical to the Black–Scholes equation, which was a key result in the foundation of options pricing theory in 1973.[19] The generalization of this equation for $\alpha \neq r$ is useful for cases when the underlying option does not behave as a financial asset.[17]

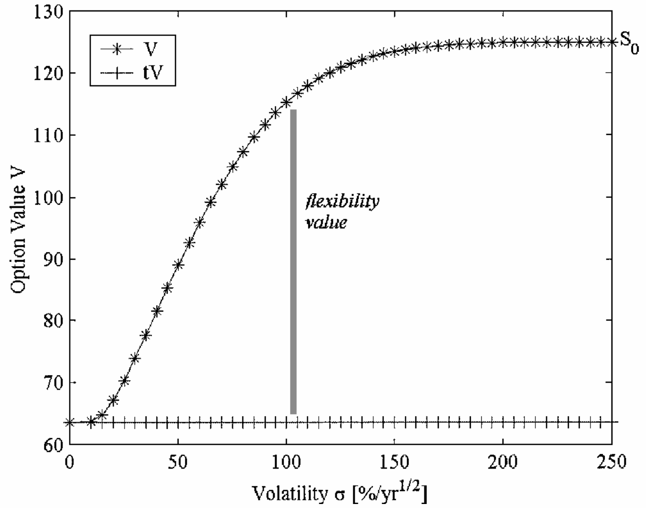

Fig. 8 Simple option on life extension: value from the Black–Scholes equation[19] for E = \$100 million, S_0 = \$125 million, T = 10 years, r = 5%/ year, and $\delta = 0$.

Quantifying Flexibility

This example is particularly relevant to illustrate the limitations of traditional valuation methods and to demonstrate that, in some instances, flexibility can be quantified.

A net present value calculation would not take into account the existence of an option. For expected revenues smaller than the expected expense $[e^{(\alpha - r)T} S_0 < e^{-rT} E]$, it would consider that the spacecraft will never be serviced and, therefore, that there is no value in the servicing option. This corresponds to neglecting the probability that the mission be more successful than expected, thus being overpessimistic. For expected revenues higher than the expected expense $[e^{(\alpha - r)} S_0 > e^{-rT} E]$, it would assume that the spacecraft will be serviced whatever the market level S. Therefore, it would be overoptimistic; the net present value incurred after time T would underestimate the value of serviceability by an amount that we can define as the flexibility value:

$$F = V - tV = \int_0^E e^{-rT}(E - S)p(S)\,\mathrm{d}S > 0 \tag{20}$$

Figure 8 shows the value of the option to be serviced for life extension and compares it with the net present value after time T, as a function of the volatility σ. Two main general trends, common to any option valuation, are worth noticing. The option value increases with uncertainty. Thus, uncertainty, which has traditionally been considered as a source of risk, is turned into a source of value by the presence of flexibility. Furthermore, the option value is greater than 0 and has as an upper bound the present value of the expected revenues, $e^{(\alpha - r)T} S_0$. This amount is the worst possible difference between actual revenues and forecast revenues. This is, therefore, a statement that the value of an option cannot exceed the value of the potential losses it helps prevent.

Option Value for Space Missions: General Case

Let us now consider the general case with discrete decision points $0 < T_1 < T_2 < \cdots < T_N < T_H$. Choosing a certain mode of operation at time T_k gives the option to choose other modes of operation at time T_{k+1}. For example, a spacecraft that has been serviced once has bought not only a life extension, but also the option to be serviced once again. The value of this future option must be taken into account in estimating the value of the first option. This situation is called a compound option. For each decision point T_k let us call the next period and note τ_k the time to the next decision point, $\tau_k = T_{k+1} - T_k$, with the convention $\tau_n = T_H - T_N$. For any variable Y, variable Y_k will be the quantity of Y incurred during the kth period: $Y_k = Y([T_k; T_{k+1}])$.

Matrix of Switching Costs

Consider a mission entering a decision point T_k in a specific mode of operation, m. There will be a cost $C_k^{(m\to n)}$ to switch from mode m to any mode n. This matrix of switching costs $C_k = (C_k^{(m\to n)})_{m,n}$ bears the index k because it can depend on time. That a certain mode l may not be accessible from mode m is taken into account by setting $C_k^{(m\to l)} = \infty$, for example, a mission cannot be reinitiated once its spacecraft have been deorbited.

Last Decision

The mission value after time T_N will depend on the mode of operation chosen at T_N. This choice will depend on the current mode of operation n as well as the current state of the uncertain parameter $x = X(T_N)$. For example, a mission should be serviced if its market booms, but abandoned if its market shrinks. If the mode of operation m is chosen, then the expected future value of the mission after T_N, given $X(T_N) = x$, will be

$$EV_N^{(n\to m)}(x) = f\left(U_N^{(m)}, C_N^{(n\to m)}\right) \tag{21}$$

At $t = 0$, both the uncertain parameter $X(T_N)$ and the mode of operation n at T_N are unknown. For each possible entering n, the decision model determines the sets $I_N^{(n\to m)}$ of values of the uncertain parameter for which a switch to mode m at T_N will maximize future value:

$$x \in I_N^{(n\to m)} \iff \max_l\left\{EV_N^{(n\to l)}(x)\right\} = EV_N^{(n\to m)}(x) \tag{22}$$

Induction Relations

Now consider decision point T_{N-1}. A future mode of operation must be chosen on the basis of the current mode of operation l and the current value of the uncertain parameter $x_{N-1} = X(T_{N-1})$. In making the decision to switch to a mode n, two things must be traded off as shown in Fig. 9: 1) the cost $C_{N-1}^{(l\to n)}$ and utility $U_{N-1}^{(n)}$ incurred during $[T_{N-1}, T_N]$ and 2) the costs and utility incurred after T_N, given that the decision point T_N will be entered in mode n and that the uncertain parameter $X(T_N)$ will follow the density probability function $p_N(x|x_{N-1})$. Thus, the value to maximize is

$$EV_{N-1}^{(l\to n)}(x_{N-1}) = \sum_m \int_{I_N^{(m)}} f\left[U_{N-1}^{(n)} + U_N^{(m)}(x), C_{N-1}^{(l\to n)} + C_N^{(n\to m)}(x)\right] p_N(x|x_{N-1})\,\mathrm{d}x \tag{23}$$

This decision model will determine the sets $I_{N-1}^{(l\to n)}$ of values of the uncertain parameter at T_{N-1} for which a switch to mode l maximizes future value.

At decision point T_{N-2}, the decision model will in turn maximize, for each entering mode k and each value of the uncertain parameter $x_{N-2} = X(T_{N-2})$, the future value

$$\begin{aligned}EV_{N-2}^{(k\to l)}(x_{N-2}) &= \sum_{n,m} \int_{I_{N-1}^{(n)}} p_{N-1}(x|x_{N-2})\,\mathrm{d}x \int_{I_N^{(m)}} p_N(y|x)\,\mathrm{d}y \cdots \\ &\cdots f\left[U_{N-2}^{(l)}(x_{N-2}) + U_{N-1}^{(n)}(x) + U_N^{(m)}(y), C_{N-2}^{(k\to l)}(x_{N-2}) \right. \\ &\left. + C_{N-1}^{(l\to n)}(x) + C_N^{(n\to m)}(y)\right]\end{aligned} \tag{24}$$

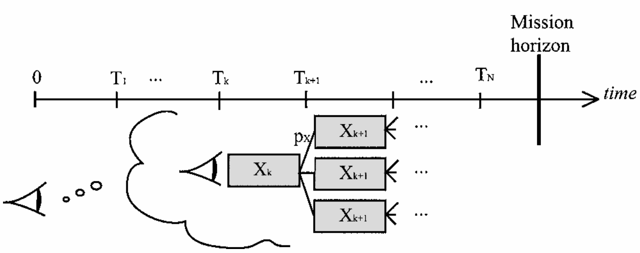

Fig. 9 Schematic of decision situation at decision point *k*.

The same principle can be applied according to a backwards iterative process back to $T_0 = 0$, where it gives the total mission value as seen from the initial point, $V = EV_{\geq 0}^{(0)}$. This process is similar to working backward in a decision tree that would have an infinite number of branches.

Commercial Case

The case of a commercial mission with uncertain revenues is worth mentioning because the linearity of the value metric greatly simplifies the valuation process. The induction relation indeed becomes

$$EV_{\geq k}^{(n)}(x) = U_k^{[\hat{m}_k(n,x)]}(x) - C_k^{[n\to \hat{m}_k(n,x)]} + e^{-r\tau_k}\int EV_{\geq k+1}^{[\hat{m}_k(n,x)]}(y)p_k(y|x)\,\mathrm{d}y \tag{25}$$

Traditional Value, Flexibility Value

Now let us define the traditional value as tV, the value that would be obtained if the existence of options had not been taken into account. A traditional valuation such as a net present value method would have considered that all decisions must be made at $t = 0$ and determined a set sequence of modes of operation $(m_0, \ldots, m_N)$ from the observation of $x_0 = X(T_0)$. Thus, the traditional mission value would have been

$$tV = \int f\left[\sum_{k=0}^{N} U_k^{(m_k)}(x_k), \sum_{k=0}^{N} C_k^{(m_{k-1}\to m_k)}(x_k)\right] \times p[(x_1, \ldots, x_N)|x_0]\,\mathrm{d}x_1, \ldots, \mathrm{d}x_N \tag{26}$$

If the value function has been chosen to be compatible with time additivity, then the value that takes flexibility into account will always be greater than the traditional value $V \geq tV$. This leads a natural definition of the value of flexibility:

$$F = V - tV \tag{27}$$

Determination of a Threshold Servicing Price

The framework and valuation process defined earlier can be used to derive three types of information about the market base for servicing:

1) The cost penalty that a space mission would be willing to pay to design for serviceability is directly given by the value of the serviceability option; an example is the Black–Scholes equation[19] for the simple option on life extension seen earlier.

2) The value of a serviceable mission is a function of the price of servicing. Comparing this function with the value of a nonserviceable mission will give the maximum servicing price that would make a space mission choose a serviceable design:

$$\text{serviceable} > \text{traditional} \iff C_{\text{service}} < \Upsilon \tag{28}$$

3) Determining the relative flexibility value with respect to traditional value, $(V - tV)/tV$, will indicate by how much traditional valuation underestimates mission value and, therefore, in what cases this framework is most interesting.

Limitations of the Framework

This paper represents a first attempt at defining a general framework for the valuation of options for space missions faced with uncertainty. Although it can account for many practical cases of space options, some simplifying assumptions had to be made, which limit its generality.

Discrete Decision Times

For the sake of clarity, a finite number of set decision points were implicitly assumed. When the period is set to be infinitesimal, this can easily be generalized to continuous decision points. However, it is clear that an iterative backward process with a cascade of integrals becomes impossible as the number of decision points tends to infinity. A solution to this problem is to alter the definition of value, defining $V' = f(E\{U\}, E\{C\})$ instead of $V = E\{f(U, C)\}$.

This new definition makes no difference for linear valuation functions such as the one used for commercial missions. It does for utility-per-cost metrics, and studying this effect would be an interesting problem for future research.

Finite Number of Modes of Operation

For the sake of clarity, the possible modes of operation were described as a finite set. This assumption is not fundamental, and the same framework can directly be generalized to the case of a continuous range of possible modes, such as a whole interval of possible orbital altitudes, simply replacing a summation by an integral. Although the process can be directly generalized to account for the two former types of continuity, some basic assumptions have been made that constitute fundamental limits to the framework:

Exogenous Uncertainty

The framework applies for cases when the uncertainty is exogenous, that is, the source of uncertainty is external to the mission and cannot be affected by decisions taken after the system has been fielded. The presence of options reduces risk not by reducing uncertainty, but rather by affecting the consequences that uncertainty has on the mission.

This assumption can be an adequate description of the option to service to react to random failures, to market and technology dynamics, or to changing requirements. It cannot be used to describe the interactions between the source of uncertainty and the mission's decision, such as the dynamics of competitive markets.

Describing the Uncertainty

The proposed valuation process relies on the availability of the information necessary to define the building blocks of the framework. For many practical cases, this information is not observable. In particular, the probability density function $p_k(x|x_0)$ of the uncertain parameter in the future, which describes the uncertainty in the parameter's forecast, is usually very uncertain itself. Assumptions have to be made, and the sensitivity of the results to the assumed distribution must be estimated. There is likely to exist a threshold uncertainty over which the conclusions change: No conclusion can be drawn in situations where the uncertainty is estimated to lie close to this threshold.

Forms of Flexibility

This framework describes flexibility as a known set of possible modes of operation available to decision makers. Therefore, this approach cannot account for the most general form of flexibility, which lies in the ability to define new, unpredictable modes of operation to respond to unknown sources of uncertainty.

Conclusions

The traditional approach for estimating the cost effectiveness of on-orbit servicing of space systems has two major flaws. First, traditional cost models are not suitable for a servicing infrastructure, which generally leads to a cost uncertainty outweighing any cost difference. Second, traditional valuation assumes a set sequence of actions, which does not take into account the flexibility provided by on-orbit servicing to space systems.

This paper proposed a new approach to on-orbit servicing. The value of servicing, defined as the maximum price that a space mission would be willing to pay for servicing, should be studied independent of its cost. The paper developed a framework to account for flexibility in estimating this value. The framework relies on the definition of a few building blocks, the more important being a model of the exogenous uncertainty, a set of reachable operational modes, a sequence of decision points, and a definition of mission value. Once these building blocks have been properly defined for a given servicing situation, a backward valuation process similar to a decision tree with infinite number of branches is proposed. This process embeds the value of flexibility by considering only the optimal decision following the gradual resolution of uncertainty with time.

This framework is valid for any space mission having a known set of options to react to the resolution of any source of uncertainty that behaves as an exogenous Markov process. Its application to on-orbit servicing can help draw a map in a requirements/uncertainty space of the maximum servicing price that would make a servicing infrastructure interesting. Such information would give directions for future on-orbit servicing technology development as to what missions to target and at what servicing price.

Acknowledgments

This research was supported by the Defense and Advanced Research Project Agency: Grand Challenges in Space ASTRO/Orbital Contract F29601-97-K-0010. Daniel E. Hastings would like to acknowledge useful discussions with David Whelan (currently at The Boeing Company).

References

[1]Waltz, D. M., *On-Orbit Servicing of Space Systems*, Krieger, Malabar, FL, 1993.

[2]Davinic, N. D., Chappie, S., Arkus, A., and Greenberg, J., "Spacecraft Modular Architecture Design Study: Cost Benefit Analysis of On-Orbit Satellite Servicing," International Academy of Astronautics, IAA Paper 97-1.4.07, Oct. 1997.

[3]Leisman, G., Wallen, A., Kramer, S., and Murdock, W., "Analysis and Preliminary Design of On-Orbit Servicing Architectures for the GPS Constellation," AIAA Paper 99-4425, Sept. 1999.

[4]Lamassoure, E., Saleh, J. H., and Hastings, D. E., "Space Systems Flexibility Provided by On-Orbit Servicing: Part 2," *Journal of Spacecraft and Rockets*, Vol. 39, No. 4, 2002, pp. 561–570.

[5]Polites, M. E., "Technology of Automated Rendezvous and Capture in Space," *Journal of Spacecraft and Rockets*, Vol. 36, No. 2, 1999, pp. 280–291.

[6]Shaw, G., Miller, D. W., and Hastings, D. E., "The Generalized Information Network Analysis for Distributed Satellite Systems," *Journal of Spacecraft and Rockets*, Vol. 36, No. 2, 1999, pp. 280–291.

[7]Lamassoure, E., "A Framework to Account for Flexibility in Modeling the Value of On-Orbit Servicing for Space Systems," M.S. Thesis, Dept. of Aeronautics and Astronautics, Massachusetts Inst. of Technology, Cambridge, MA, June 2001.

[8]Wertz, J. R., and Larson, W. J., *Space Mission Analysis and Design*, 3rd ed., Space Technology Series, Microcosm and Kluwer Academic, Norwell, MA, 1999.

[9]Saleh, J. H., Hastings, D. E., and Newman, D. J., " Spacecraft Design Lifetime," *Journal and Spacecraft and Rockets*, Vol. 39, No. 2, 2002, pp. 244–257.

[10]Wertz, J. R., and Larson, W. J., *Reducing Space Mission Cost*, Space Technology Series, Microcosm and Kluwer Academic, Norwell, MA, 1996.

[11]Gumbert, C. C., Violet, M. D., Hastings, D. E., Hollister, W. M., and Lovell, R. R., "Cost per Billable Minute Metric for Comparing Satellite Systems," *Journal of Spacecraft and Rockets*, Vol. 34, No. 6, 1997, pp. 837–846.

[12]Fosa, C. E., Raines, R. A., Gunsch, G. H., and Temple, M. A., "An Overview of the Iridium Low Earth Orbit Satellites System," *Proceedings of the IEEE*, 1998.

[13]Shaw, G., "The Generalized Information Network Analysis for Distributed Satellite Systems," Ph.D. Dissertation, Dept. of Aeronautics and Astronautics, Massachusetts Inst. of Technology, Cambridge, MA, June 1999.

[14]Saleh, J. H., Hastings, D. E., and Newman, D. J., "Extracting the Essence of Flexibility in System Design," NASA/Dept. of Defense Workshop on Evolvable Hardware, July 2001.

[15]Amram, M., and Kulatilaka, N., *Real Options: Managing Strategic Investment in an Uncertain World*, Harvard Business School Press, Cambridge, MA, 1998.

[16]Trigeorgis, L., *Real Options; Managerial Flexibility and Strategy in Resource Allocation*, MIT Press, Cambridge, MA, 1996.

[17]McDonald, R. L., and Siegel, D. R., "Investment and the Valuation of Firms when There is an Option to Shut Down," *International Economic Review*, Vol. 26, No. 2, 19895, pp. 331–334.

[18]Merton, R. C., "An Intertemporal Capital Asset Pricing Model," *Econometrica*, No. 41, Sept. 1973, pp. 867–887.

[19]Black, F., and Scholes, M., "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, Vol. 81, No. 3, 1973, pp. 637–654.

A. C. Tribble
Associate Editor