A MARTIAN QUARANTINE RISK MODEL

by

S. Eric Steg and Richard G. Cornell

March 8, 1968

Technical Report Number 15

NASA Grant Number NGR-10-004-029

Department of Statistics
Florida State University
Tallahassee, Florida

A MARTIAN QUARANTINE RISK MODEL

I. Introduction.

The specific aims of this report are:

1. to determine a meaningfully structured model utilizing loss functions to determine the risk of employing a given decontamination procedure to all Martian unmanned landing craft, and
2. to evaluate the relative risks of several spacecraft decontamination levels for a variety of loss functions using this model, and finally
3. to re-examine current decontamination policies from the viewpoint of the present work.

An underlying goal of this paper will be to develop a basic new approach for assessing the validity of the choice of the parameter v, the probability of one viable organism aboard the landing capsule at the time of landing, which we will later equate to the mean number of viable microorganisms on the capsule, as has been selected by Sagan and Coleman (1966) and subsequently criticized by Horowitz *et al.*, (1967). Apparently no previous attempt has been made to develop a model involving this parameter which also takes cognizance of the type of risks actually foreseeable and formulates some conceptions of these risks in numerical terms - a primary purpose of this work. The conclusions of this paper will then actually relate to the minimal form of decontamination required to achieve the levels of cleanliness suggested. The required decontamination procedures could range from gas decontamination of exterior and mating surfaces to prolonged heating of the entire spacecraft.

We will consider only the biological ramifications of the lander missions, since these losses must be deemed of higher order than those possessing a political or financial nature. In particular, we are concerned that any life on Mars not be subverted or changed by the introduction of terrestrial biota prior to a full study of that life, or that a condition of non-life unknowingly be supplanted with or modified by organisms from Earth to be subsequently misinterpreted as aboriginal. It is in our interest, on the other hand, to attempt to learn the true nature of the Martian biological condition before the impending invasion of Mars by contaminant-bearing human beings. Such information could add to the growing body of knowledge about the nature and the origin of our existence in a dramatic way. It could also offer clues to the possibility of back-contamination from Mars on the later manned ventures which might actually threaten our existence.

II. Development of the Model.

In our analysis we shall employ state, outcome, and action spaces, a member of each being a tuple of the form:

A. States $\theta = (\cdot,\cdot)$, $\theta \in \boldsymbol{\theta}$

First element: G - a probability that a single viable organism released to the Martian surface will grow and multiply, leading to the contamination of a significant portion of the planet.

Second element: S - a probability triple expressing the respective likelihood of the outcomes "soft landing, crash, and miss of the planet," as accorded by engineering proficiency and by the amenability of the Martian environment to engineering success.

B. Outcomes $\xi = (\cdot,\cdot)$, $\xi \in \Xi$

First element: C - Mars significantly contaminated.

$\bar{C}$ - Mars not significantly contaminated.

Second element: D - data objectives adequately met.

$\bar{D}$ - data obtained insufficient for a declaration of success.

C. Actions $\alpha = (\cdot,\cdot)$, $\alpha \in A$

First element: L - a given level or procedure of decontamination.

Second element: W_L - the number of missions or slots which the lander program would be delayed to achieve decontamination level L.

In this work a "soft landing" implies a landing having sufficiently low impact velocity for the experimental and functional apparatus not to have suffered harm. By "significant contamination" we mean contamination that has diffused across the surface of Mars embodying sufficient density to have a high likelihood of biasing any later experiments. A "slot" is an optimal launch period for one or more missions characterized by a near approach of Mars to Earth. These periods occur about 25 months apart. Further, the element W_L depends directly on the decontamination procedure decided upon since the time of the initiation of the program depends upon the time required to perfect that procedure to a satisfactory degree to allow a mission to be flown. We assume that once a level is decided upon it will be used for all the missions, for the present. Thus no delays will be incurred after the initial waiting time for the start of the unmanned lander program. As an important basic premise, we assume that a politically or financially oriented decision will determine when the first Martian manned venture will be undertaken, and thus, effectively, when the unmanned program will cease. If the unmanned program is delayed, therefore, this decision will truncate the last missions from that program.

We could approach the ensuing analysis in three different manners: with a model centered on individual experiments, on missions, or on slots. A total lack of knowledge about the number and type of experiments to be carried out over the complete program suggests abandonment of that approach in favor of the latter two strategies.

The advantage of the slot approach rests in its effective handling of the problem in which a mission landing during the same slot as an earlier contaminating mission will not necessarily suffer the effects of that contamination. We would be assuming then that contaminating organisms require more than the width of a slot to become generally dispersed--if growth and dispersion are indeed possible. The total number of slots, N, utilized by the unmanned lander program will probably be small, i.e., between five and eight. We may visualize multiple numbers of lander missions and even single-bus, multiple-craft missions in each slot.

If we are willing to make the more conservative assumption that any contamination released by one mission will adversely affect the next and all subsequent missions, then the mission approach has an advantage in that with it we need not approximate how many missions will be flown in each individual slot, but only the total number of missions to be flown.
In either approach the analysis will be the same when the proper definitions are given to the parameters. We will refer to slots or missions as "dates" in the following work.

We must now specify the nature of our losses. At each date u we should learn a certain amount. Failure to do so results in an informational loss having value f(u). Making no missions on a date, or making them but obtaining little or no data both incur the same implied loss f(u). A further loss for the outright act of contaminating also seems appropriate. This loss we call g(u). Its magnitude will be chosen to reflect both the value of worthwhile data which could be obtained by early manned missions if the planet has not suffered a prior prolonged period of spreading contamination, and the increased danger from back-contamination resulting from biased knowledge. Further, it could include a term accounting for the new hazard of back-contamination by familiar organisms having a changed nature due to their breeding in a new environment (possibly as a result of genetic changes caused by increased solar radiation). Our date u loss table is simply

	D	$\bar{D}$
$\bar{C}$	0	f(u)
C	g(u)	f(u) + g(u)

where the element (C,D) implies that the released contamination has not biased the data of that date.

Let us define w to be the number of date outcomes ($\bar{C}$,D), and x to be those of ($\bar{C}$,$\bar{D}$) among those of the N total dates. The information loss for the first unused dates is $\sum_{u=1}^{W_L} f(u)$. Thus the risk function (expected loss) for a fixed state θ has the form

$$R(\theta,\alpha) = \sum_{u=1}^{W_L} f(u) + \sum_{D\in\mathcal{D}} \ell_\theta(\xi_{W_L+1},\ldots,\xi_N;\alpha)\cdot p_\theta(\xi_{W_L+1},\ldots,\xi_N;\alpha)$$

with $\alpha\in A$, the action space; $\xi_k\in\Xi$, the outcome space (ξ_k is the outcome at date k); $D\in\mathcal{D}$, the $(N-W_L)$-fold product space of possible sequences of outcomes; and where $\ell_\theta(\cdot)$ is the loss for the sequence of outcomes while $p_\theta(\cdot)$ is the probability of the occurrence of the sequence. An average risk $\bar{R}$ may then be found for an action α by averaging over the possible states as

$$\bar{R}(\alpha) = \sum_{\theta\in\Theta} R(\theta,\alpha)m(\theta)$$

where $m(\theta)$ is a probability measure on the state space Θ.

The probabilities $P(\cdot)$ of the possible outcomes given within the parentheses for date u will be denoted by

$$p_1(u) = P_\theta(\bar{C},D;u) \quad p_2(u) = P_\theta(\bar{C},\bar{D};u) \quad p_3(u) = P_\theta(C,D;u) \quad p_4(u) = P_\theta(C,\bar{D};u).$$

For the present development we shall make the assumption that data taken on dates after the first date yielding contamination, B, is invalid information due to biasing. Thus we let B be the numerical index of the first date for which contamination of the planet occurs. Up to this time each outcome ξ_k is either $\xi_k = (\bar{C},D)$ with a loss 0 and probability $p_1(k)/(p_1(k)+p_2(k))$, or $\xi_k = (\bar{C},\bar{D})$ with a loss f(k) and probability $p_2(k)/(p_1(k)+p_2(k))$. In addition to this loss, during the contaminating mission's date we expect to lose $[p_4(B)/(p_3(B)+p_4(B))]f(B) + g(B)$, while by the conservative assumption previously made, all information garnered after date B is assumed invalid, whence we encounter a further loss equal to $\sum_{u=B+1}^{N} f(u)$. Summing over the possible first times to contamination, B, the risk function becomes

$$R(\theta,\alpha) = \sum_{u=1}^{W_L} f(u) + \sum_{B=W_L+1}^{N+1} \left[\sum_{u=W_L+1}^{B-1} \frac{f(u)\cdot p_2(u)}{p_1(u)+p_2(u)} + \frac{f(B)\cdot p_4(B)}{p_3(B)+p_4(B)} + g(B) + \sum_{u=B+1}^{W} f(u) \right] \cdot p(B)$$

where $p(B) = (p_3(B)+p_4(B))\cdot \prod_{u=W_L+1}^{B-1} (p_1(u)+p_2(u))$ if $B \neq N+1$,

$$p(N+1) = \prod_{u=W_L+1}^{N} (p_1(u)+p_2(u)),$$

and by definition $B = N + 1$ implies that the unmanned program ends without contaminating Mars, so that $f(N+1) = g(N+1) = 0$. Combining terms inside the brackets to allow the term $\sum_{u=W_L+1}^{N} f(u)$ to be brought outside the summation over B results in the simplification of the risk to

$$R(\theta,\alpha) = \sum_{u=1}^{N} f(u) - \sum_{B=W_L+1}^{N+1} \left[\sum_{u=W_L+1}^{B-1} \frac{f(u)\cdot p_1(u)}{p_1(u)+p_2(u)} + \frac{f(B)\cdot p_3(B)}{p_3(B)+p_4(B)} - g(B) \right] p(B) .$$

We now need to define the probabilities $p_i(u)$ more precisely. We will do this first for the mission-centered approach, and then extend these definitions in the next section to the case of slots having specified numbers of missions per slot. Here we assume $p_i(u) = p_i$, $W_L+1 \leq u \leq N$. Thus for a single mission we compute

$p_1 = P(\bar{C},D) = P(D|\bar{C})\cdot P(\bar{C}) \doteq P(D)\cdot P(\bar{C}) = P(D)\cdot(1-P(C))$ when $P(C)$ is small since $P(D) = P(D|C)\cdot P(C) + P(D|\bar{C})\cdot P(\bar{C}) \doteq P(D|\bar{C})$ when $P(C)$ is small,

$p_2 = P(\bar{C},\bar{D}) = P(\bar{D}|\bar{C})\cdot P(\bar{C}) = (1-P(D|\bar{C}))\cdot P(\bar{C}) \doteq (1-P(D))(1-P(C))$,

$p_3 = P(C,D) = P(D|C)\cdot P(C)$,

$p_4 = P(C,\bar{D}) = (1-P(D|C))\cdot P(C)$, where in particular,

$$P(C) = P(\text{viable organism aboard})\cdot P(\text{release}|\text{aboard})\cdot P(C|\text{release, aboard}),$$

$$P(\text{release}|\text{aboard}) = P(\text{soft land})\cdot P(\text{release}|\text{soft land, aboard}) + P(\text{crash land})\cdot P(\text{release}|\text{crash, aboard}),$$

$$P(D) = P(D|\text{soft land})\cdot P(\text{soft land}), \text{ and}$$

$$P(D|C) = P(D|C, \text{soft land})\cdot P(\text{soft land}).$$

III. Application with the Sagan-Coleman Parameters.

The Sagan-Coleman analysis of the planetary quarantine problem had a strong influence on the recommendations set forth by the international committee known as COSPAR. It therefore becomes important in reviewing current United States policy to review it in terms of the work and parameter evaluations which instigated this policy through the COSPAR agreements. Consequently, in this section we will use the Sagan-Coleman mission values wherever theirs fit into the framework of our model. For missions, we employ their probability of significant contamination of Mars given the release of a single organism of $P(C|\text{release, aboard}) = 10^{-2}$, where the interpretation of this value varies for the mission-centered approach and the slot-centered approach. Their release probability $P(\text{release}|\text{aboard}) = 1$ and their $P(\text{crash}) = 0.1$ are also accepted. We then set $P(\text{soft land}) = 0.9$ and $P(\text{miss}) = 0.0$. These five probabilities may be combined to form a definition for a state θ. We next take $P(D|\text{soft land})$ as 0.9 to indicate a high probability of successful data collection after a soft landing, and $P(D|C, \text{soft land})$ as 10^{-2} to reflect the belief that data collected on a contaminating mission will likely be biased. Finally letting $v = P(\text{viable organism aboard})$, we obtain the mission parameters

$p_1 = 0.81(1-v\cdot 10^{-2})$, $p_2 = 0.19(1-v\cdot 10^{-2})$, $p_3 = 0.009v\cdot 10^{-2}$, and $p_4 = 0.991v\cdot 10^{-2}$.

Sagan and Coleman (1966) give the estimate of 30 missions, or three missions per opportunity, to be flown during the United States unmanned lander program. It will be appropriate for us to only consider United States policy both because decisions are made to delay or not delay the program at the national level as related to an individual country's ability to achieve prescribed decontamination levels, and because we must presume a lack of information from Russia regarding its plans and accomplishments.

To form the corresponding slot parameters, we note that the Sagan-Coleman estimate of an average of three missions per slot and 30 total missions implies an estimate of 10 slots available for the unmanned program (some of which may go unused). For the purposes of this analysis we will assume that each slot contains exactly three missions. If we require over one-half of the missions in a given slot to yield significant data to warrant a declaration of slot success, we will then require two fruitful missions in each slot. To transform from the single mission probabilities p_i to the slot probabilities p_i', we form

$$p_1' = p_1^3 + \binom{3}{1} p_1^2 p_2 = p_1^3 + 3p_1^2 p_2 ,$$

for each slot u, since under our assumptions $p_i'(u) = p_i'$ for all u. In a similar fashion we find, upon disregarding terms with multiple contamination probability factors as insignificantly small,

$$p_2' = p_2^3 + 3p_2^2 p_1,$$
$$p_3' = 3p_1^2 p_3 + 3p_1^2 p_4 + 6p_1\cdot p_2\cdot p_3,$$
$$p_4' = 3p_2^2 p_3 + 3p_2^2 p_4 + 6p_1\cdot p_2\cdot p_4 .$$

Regarding v as the expected number of viable organisms aboard a lander when v exceeds one, the values for p and p′ for v ranging from 10 down to 10^{-6} are given in Table 1.

Table 1

Values of p for various values of v.

v	p_1	p_2	p_3	p_4
10	.72900000000	.17100000000	.00090000000	.09910000000
1	.80190000000	.18810000000	.00009000000	.00991000000
10^{-1}	.80919000000	.18981000000	.00000900000	.00099100000
10^{-2}	.80991900000	.18998100000	.00000090000	.00009910000
10^{-3}	.80999190000	.18999810000	.00000009000	.00000991000
10^{-4}	.80999919000	.18999981000	.00000000900	.00000099100
10^{-5}	.80999991900	.18999998100	.00000000090	.00000009910
10^{-6}	.80999999190	.18999999810	.00000000009	.00000000991

Values of p′ for various values of v.

v	p_1'	p_2'	p_3'	p_4'
10	.66004972200	.06895027800	.16010545860	.08289454140
1	.87852617998	.09177282002	.01937276049	.01003023951
10^{-1}	.90270446135	.09429853765	.00197265936	.00102134364
10^{-2}	.90514640176	.09455362824	.00019762153	.00010231847
10^{-3}	.90539083773	.09457916257	.00001976571	.00001023369
10^{-4}	.90541528375	.09458171625	.00000197661	.00000102339
10^{-5}	.90541772837	.09458197163	.00000019766	.00000010234
10^{-6}	.90541797284	.09458199716	.00000001977	.00000001023

Finally, in choosing loss functions, we take three different forms for the informational loss:

$$f_1(u) = a \cdot u,$$
$$f_2(u) = b,$$
$$f_3(u) = c/u,$$

where a, b, and c are constants taken such that $\sum_{u=1}^{N} f(u) = 100$, j = 1, 2, and 3. For N = 30 in the mission approach, a = 0.21505376, b = 3.33333333, and c = 25.0313696. For the slot approach with N = 10 slots, a = 1.818182, b = 10.000000, and c = 34.141715. These f(u) functions represent different possible forms of sequential data loss and they are increasing, constant, and decreasing, or alternatively their cumulatives are respectively concave, linear, and convex in u. We stress that the numerical values assigned the f(u) functions are dimensionless, with the choice of $\sum_{u=1}^{N} f(u) = 100$ being arbitrary and indicating only a loss value relative to the contamination loss value g(u). It is solely the relative size of the two losses which is of importance, and without considering their comparative size meaningful losses could not be assigned.

We also adopt a g(B) function having one of the following forms:

$g_1(B) = 100.0,$

$g_2(B) = 100.0 + 0.10752688(930.0-B(B-1.0))$ for missions, or

$= 100.0 + 0.909091(110.0-B(B-1.0))$ for slots,

$g_3(B) = 1000.0$.

Taking g = 100 gives the informational and contaminating losses equal importance, whereas g = 1000 stresses the loss due to planetary contamination. The variable g(B) depending on B is, in either form, large for a small B and decreasing with increasing B, reflecting the conviction that the impulse of loss to the risk function $R(\theta,\alpha)$ should be greater for exposing the planet to contamination for a more prolonged period prior to the inception of manned landings. The largest possible loss would be 200, and the smallest 100, if the earliest contamination occurred during the first or last mission, respectively.

Using the equation for $R(\theta,\alpha)$, values for both the mission and the slot approach were computed for the risk. Examples of these values are given in Table 2 for the parameter selections given above, and for the variable g(B) function.

Table 2

Risk values for the Sagan-Coleman parameter values.

A. Using Missions

v	f_1	f_2	f_3	f_1	f_2	f_3
	No Delay (W_L=0)			One Mission Delay (W_L=1)		
10	266.653	256.389	233.147	263.122	254.247	246.019
1	78.087	74.712	68.957	76.019	75.136	86.901
10^{-1}	25.596	25.211	24.571	25.496	27.636	44.571
10^{-2}	19.667	19.628	19.563	19.813	22.300	39.811
10^{-3}	19.067	19.063	19.056	19.238	21.760	39.329
10^{-4}	19.007	19.006	19.006	19.181	21.706	39.281
10^{-5}	19.001	19.001	19.001	19.175	21.701	39.276
10^{-6}	19.000	19.000	19.000	19.174	21.700	39.275
	Three Mission Delay (W_L=3)			Six Mission Delay (W_L=6)		
10	255.481	249.320	251.500	242.608	240.326	248.431
1	72.355	76.007	99.583	68.091	77.402	106.165
10^{-1}	25.818	32.494	60.967	27.616	39.805	72.772
10^{-2}	20.628	27.645	56.656	23.158	35.665	69.088
10^{-3}	20.104	27.155	56.220	22.708	35.246	68.716
10^{-4}	20.051	27.106	56.176	22.663	35.205	68.679
10^{-5}	20.046	27.101	56.172	22.659	35.200	68.675
10^{-6}	20.045	27.100	56.172	22.658	35.200	68.675

B. Using Slots

v	f_1	f_2	f_3	f_1	f_2	f_3
	No Delay (W_L=0)			One Slot Delay (W_L=1)		
10	245.812	234.929	215.825	235.235	228.790	227.618
1	69.310	65.792	60.784	63.989	67.916	85.043
10^{-1}	16.174	15.772	15.213	16.971	23.992	45.336
10^{-2}	10.138	10.097	10.040	11.697	19.066	40.873
10^{-3}	9.526	9.522	9.516	11.164	18.568	40.421
10^{-4}	9.465	9.465	9.464	11.110	18.518	40.376
10^{-5}	9.459	9.459	9.459	11.105	18.513	40.371
10^{-6}	9.458	9.458	9.458	11.104	18.512	40.371

We see that for fixed W_L, the risk values necessarily decrease with decreasing v. For a fixed v, several entries in the table tell us to unconditionally delay the lander program, implying that the chance of contamination under these large v values exhibiting this behavior is too great compared to the value of the information which might be collected. Numerous "no delay" values are dominated by, that is have a larger risk than, "delay" values for smaller magnitude v values. The "no delay" risks, under an f_i, however, are smaller than any of the "delay" values under the same f_i in the mission table for $v \leq 10^{-3}$ under f_1, $v \leq 10^{-2}$ under f_2, and $v \leq 10^{-1}$ under f_3; and in the slot table for $v \leq 10^{-2}$, $v \leq 10^{-1}$, and $v \leq 10^{-1}$, respectively. Thus it is best to initiate the landing program at the first planned launch opportunity if we can obtain a value of v satisfying the inequality above which corresponds to the f function and the approach finally adopted. We should delay the start if we cannot obtain the prescribed level at that time, but feel that with the delay additional improvements in design can be made which will allow the achievement of a sufficiently lower level of spacecraft contamination to give a smaller risk, while, of course, maintaining a high level of spacecraft reliability.

IV. Other Parameter Values.

It might be reasonable rather than to employ the 10 usable slots available when Sagan and Coleman first did their work, to consider only the 6 planned slots between 1973 and 1984. Then the 30 missions average out to 5 missions per slot. With the p_i' and the f and g functions appropriately modified by a change in coefficients, we have computed the risk tables for the same p_i parameter values as before. These results are summarized later in Table 3. On the other hand, it might be a better approximation to use the "average of three missions per slot" criterion for the six likely slots, thus obtaining a total of only 18 missions. With the f and g functions again appropriately modified, we have once more tabled the risk values, both for mission and slot approaches. Table 3 also contains a résumé of these findings.

Again using 18 missions in 6 slots, we might assume that there will be varying numbers of missions per slot, say 1, 2, 3, 4, 4, and 4 missions in the six successive slots, over one-half of which must be data-wise successful for a determination of slot success. Proper evaluation of the risk function using the Sagan-Coleman p_i values and the proper f and g functions effects a further risk table also abstracted in Table 3.

Lastly, modified p_i evaluations have been carried out and the consequent risks found. We now take parameter values which could currently be regarded as better than those given by Sagan and Coleman. To begin, as stated by the Spacecraft Sterilization Advisory Committee (AIBS, 1967), even $P(C|\text{release}) = 10^{-3}$ may be thought of as a conservative estimate, and we use this value now. We next adopt both the Space Science Board and Horowitz-Davies release probability estimates (AIBS, 1967), applying the former to $P(\text{release}|\text{crash}) = 10^{-1}$ and the latter to $P(\text{release}|\text{soft land}) = 10^{-3}$. The landing outcome triple is taken to be (0.85, 0.10, 0.05), slightly more cautious than before. We shall take a more conservative estimate than that used in the last section by specifying 0.75 for the probability of successful data collection given a soft landing. Again taking $P(D|C, \text{soft land}) = 0.01$ completes the parameter specifications and gives single mission p_i values of $p_1 = 0.6375(1-1.085v \cdot 10^{-5})$, $p_2 = 0.3625(1-1.085v \cdot 10^{-5})$, $p_3 = 0.0085(1.085v \cdot 10^{-5})$, and $p_4 = 0.9915(1.085v \cdot 10^{-5})$. A risk table has been worked out for this new set of parameter values and it is summarized along with the other approaches in Table 3.

Table 3 gives the v values representing the conservative approximate upper bounds for the contamination level to justify not delaying the unmanned lander program under the alternative specifications which have been presented in this section. If the level cited cannot be achieved by the first launch opportunity, but a lower level which will give smaller risk can be procured by delaying, then the appropriate action is to delay. Generally the level required to give a smaller risk after a delay is the same as upper bound value listed.

Table 3

Decontamination levels required for no delay (v values).

A. Mission Approach

N	Form of g	Form of p	vs. 1 mission delay f_1	f_2	f_3	vs. 3 mission delay f_1	f_2	f_3
30	g_1	S-C	10^{-3}	10^{-2}	10^{-1}	10^{-2}	10^{-1}	1
	g_2		10^{-3}	10^{-2}	10^{-1}	10^{-2}	10^{-1}	10^{-1}
	g_3		10^{-4}	10^{-3}	10^{-2}	10^{-3}	10^{-2}	10^{-1}
18	g_1	S-C	10^{-2}	10^{-1}	1	10^{-1}	10^{-1}	1
	g_2		10^{-2}	10^{-1}	10^{-1}	10^{-2}	10^{-1}	1
	g_3		10^{-3}	10^{-2}	10^{-1}	10^{-2}	10^{-2}	1
18	g_1	Modified	10	10^{2}	10^{2}	10	10^{2}	10^{3}
	g_2		1	10	10^{2}	10	10^{2}	10^{2}
	g_3		1	10	10	10	10	10^{2}

B. Slot Approach

N	Form of g	Form of p	Missions per slot	Slots	vs. 1 slot delay f_1	f_2	f_3
30	g_1	S-C	3	10	10^{-2}	10^{-1}	10^{-1}
	g_2				10^{-2}	10^{-1}	10^{-1}
	g_3				10^{-3}	10^{-2}	10^{-1}
30	g_1	S-C	5	6	10^{-2}	10^{-1}	1
	g_2				10^{-2}	10^{-1}	10^{-1}
	g_3				10^{-2}	10^{-2}	10^{-1}
18	g_1	S-C	3	6	10^{-1}	10^{-1}	1
	g_2				10^{-1}	10^{-1}	1
	g_3				10^{-2}	10^{-2}	10^{-1}
18	g_1	S-C	Variable	6	10^{-1}	10^{-1}	1
	g_2				10^{-1}	10^{-1}	1
	g_3				10^{-2}	10^{-2}	10^{-1}
18	g_1	Modified	Variable	6	10^{2}	10^{2}	10^{3}
	g_2				10	10^{2}	10^{2}
	g_3				10	10	10^{2}

The variability of the results for differing values of g leads to the question of which constant g value would give equal risks for not delaying at a level $v = 10^m$ and delaying one mission or one slot to obtain level $v^* = 10^{m-1}$. The appropriate expression for g is

$$g = \left\{ p_3/(p_3+p_4)\cdot \sum_{B=W_L+1}^{N} f(B)\cdot(p(B)-p^*(B)) + p_1/(p_1+p_2)\cdot \sum_{B=W_L+2}^{N+1} \left[\sum_{u=W_L+1}^{B-1} f(u)\cdot(p(B)-p^*(B)) \right] + f(W_L+1)\cdot p_1/(p_1+p_2) \right\} \Bigg/ \sum_{B=W_L+1}^{N} (p(B)-p^*(B))$$

where p(B) is the probability of B given $v = 10^m$ and p*(B) is that given $v^* = 10^{m-1}$ with, in particular, $p^*(W_L+1) = 0$. We have given this expression for g only when $p_i(u) = p_i$ for all i. For 30 missions Table 4 gives the constant g function value which would give equal risk to the situations described above. Table 4 also gives the g values for the six-slot, three missions per slot problem. Since rather large values of g are thought to be inconsistent with the importance of obtaining data on Mars before manned landings, that is, with $\sum_{u=1}^{N} f(u) = 100$, the mission table indicates, for instance, that under f_3 a contamination loss of such magnitude as to make the informational loss almost insignificant would be required to give equal risk to waiting to obtain a level $v^* = 10^{-3}$ and not delaying at level $v = 10^{-2}$. In the slot table the same comment can be made for waiting for $v^* = 10^{-3}$ and not delaying under $v = 10^{-2}$. In either case the correct action would be to not delay in beginning the unmanned program, if a level of 10^{-2} can be attained.

Table 4

Table of constant g losses yielding equal risks.

A. 30 missions, one mission delay.				
v	v*	f_1	f_2	f_3
10	1	-76.152	-62.842	-13.171
1	10^{-1}	-56.392	-32.591	65.338
10^{-1}	10^{-2}	-48.652	59.037	739.511
10^{-2}	10^{-3}	9.392	955.908	7473.121
10^{-3}	10^{-4}	587.914	9922.717	74808.436
10^{-4}	10^{-5}	6372.942	99590.616	748161.516
10^{-5}	10^{-6}	64223.320	996269.582	7481692.310

B. 6 slots, 3 missions per slot, one slot delay.				
v	v*	f_1	f_2	f_3
10	1	-77.563	-47.703	5.863
1	10^{-1}	-28.647	56.287	219.260
10^{-1}	10^{-2}	208.655	881.426	2236.227
10^{-2}	10^{-3}	2559.669	9114.329	22395.871
10^{-3}	10^{-4}	26067.635	91441.537	223991.314
10^{-4}	10^{-5}	261147.074	914713.433	2239945.650
10^{-5}	10^{-6}	2611941.450	9147432.380	22399489.000

From these various tables it becomes clear that the current quarantine requirement of $v = 10^{-3}$ implies that the possible data loss is small compared to the contamination loss. Moreover these results are almost certainly conservative, since the release of contaminating organisms, as well as their growth and spread to the point of representing significant planetary contamination appears to be much less likely to occur than we have assumed. In addition, the decreasing f_3 type of function would seem the most appropriate since this corresponds to assigning larger losses to the failure to gain data early in the program than to later missions. Such an early failure would require truncation of the last, more complex missions due to a lack of preliminary Martian data acquisition. Under f_3 the decontamination requirements, as indicated by Tables 2 and 3, are the least rigid. Thus it would be appropriate for United States policy to be reviewed with interest focused on the forms and relative sizes of data collection failure losses and contamination losses to see if a lowering of current requirements might not be in order.

V. Extensions for the Risk Model.

The model previously given is flexible enough to easily include numerous extensions, each of which might aid it in becoming a more precise decision-making tool.

Including additional elements in the θ state space is straightforward and requires no modification of the risk formulation. Our action space A will not need expansion for the unmanned lander program. However, rather than the current Ξ outcome space containing "significant" contamination, we may wish to consider the amount of spread and the mean density of contamination. In addition, it could prove useful to consider partial collection of mission data. We will break this discussion up into three parts:

(A) Local and global contamination,

(B) Density of contamination, and

(C) Partial data collection.

In considering (A) and the spread of contamination in terms of the portion of the planet experiencing it, unfortunately g(B) will vary in value depending on where successive landers actually land with respect to where the contamination is located. Since we cannot know the total extent of contamination or whether a later probe will ever land amidst it, we are forced to consider extending only to the two concepts of local and global contamination. The former biases only the mission which brought it, that is, the contamination is presumed to remain in the close vicinity of its inception, and the latter biases all subsequent missions to the planet as a result of spreading across the full face of that body.

Let

$\bar{C}$ = no contamination,

C_0 = local contamination, and

C_1 = global contamination.

Under C_0 the loss $\sum_{u=B+1}^{N} f(u)$ will not occur, but we may experience multiple losses $f(B)\cdot p_3(B)/(p_3(B)+p_4(B))$. Thus under C_0 the risk becomes

$$R_0(\theta,\alpha) = \sum_{u=1}^{W_L} f(u) - \sum_{k=W_L+1}^{N} f(k)\cdot(p_2(k)+p_4(k))$$

where k is a mission possibly causing local contamination. The slot approach will not readily admit this extension. Suppose we feel that C_0 is the type of contamination that would occur with probability δ when an organism is released and C_1 with probability $1 - \delta$. Then our weighted risk would have the form:

$$R^*(\theta,\alpha) = \delta\cdot R_0(\theta,\alpha) + (1-\delta)\cdot R\,(\theta,\alpha) .$$

Concerning (B), we may regard the mean density of organisms as relating directly to the probability γ that successive missions are biased, that is, given a density factor $r(\gamma)$, we add in only $\gamma \cdot \sum_{u=B+1}^{N} f(u)$, $0 \leq \gamma \leq 1$, to our new risk function. Also we possibly incur a smaller $g(B)$, say of the form $h(\gamma) \cdot g(B)$, where h is some function (monotone increasing in γ).

In extending to the case of partial data (C), we first consider the amount of information collected to be representable on a continuous scale and that we can place a continuous distribution on the amount or proportion ρ_D of data collected. Then the terms $f(u) \cdot p_1(u)/(p_1(u)+p_2(u))$ and $f(B) \cdot p_3(B)/(p_3(B)+p_4(B))$ from the original risk function are respectively replaced by

$$f(u) \cdot \int_0^1 \rho_D \, h_u(\rho_D|\bar{C}) d\rho_D = f(u) \cdot E_u(\rho_D|\bar{C}) \qquad \text{and}$$

$$f(B) \cdot \int_0^1 \rho_D \, h^*(\rho_D|C) d\rho_D = f(B) \cdot E_B(\rho_D|C),$$

where h is the distribution of ρ_D given $\bar{C}$ and h* is that given C. If we are interested in discretizing the amount of information by dividing it into an integer m number of equal parts, the integrals above become sums. Numerical risk calculations have not been carried out with these extensions, partly due to a present lack of knowledge as to appropriate values and distributions for them. However, they do illustrate the flexible potential of this approach to developing a planetary quarantine model. It is clear that the incorporation of these extensions would liberalize the conclusions, that is, would call for a further lowering of the current high standards.

REFERENCES

AIBS Spacecraft Sterilization Advisory Committee (1967). Minutes of June Meeting at Marina del Ray, California.

Horowitz, N. H.; Sharp, R. P.; and Davies, R. W. (1967). Planetary Contamination I: The Problem and the Agreements. Science, 155, 1501-1505.

Sagan, C. and Coleman, S. (1966). Standards for Spacecraft Sterilization. Biology and the Exploration of Mars. National Academy of Sciences, Washington, D. C., 470-481.

Program Appendix

To obtain the results cited in the tables in the text, numerous computer programs were run. We include three different programs in this section to illustrate their form:

1. a program to compute $R(\theta,\alpha)$ for the mission approach with 30 total missions,

2. a program to compute $R(\theta,\alpha)$ for the slot approach with 18 total missions in 6 slots with variable numbers of missions in each slot, and

3. a program to compute a constant g value which will equate "no delay" and "delay at the next lower v level" for the slot approach with 3 missions per slot and 6 total slots.

These programs can be readily modified to accommodate the other parameterizations of the unmanned lander quarantine problem.

```
THIS IS THE MISSION PROGRAM FOR SAGAN-COLEMAN PARAMETERS WITH A TOTAL OF 30
MISSIONS IN THE UNMANNED LANDER PROGRAM
      COMMON IA,IT,IB,N,P1,P2,P3,P4,JJ,KP
C     N IS THE NUMBER OF MISSIONS
      N=30
C     LL IS THE NUMBER OF V LEVELS, WHERE V=10.**(2-LL)
      DO 66 LL=1,8
      JA=2-LL
C     Q,R,S,AND T ARE THE SAGAN-COLEMAN MISSION PARAMETERS
      Q=.81*(1.-10.**(-2+JA))
      R=.19*(1.-10.**(-2+JA))
      S=.009*10.**(-2+JA)
      T=.991*10.**(-2+JA)
      P1=Q
      P2=R
      P3=S
      P4=T
      WRITE(6,52) P1,P2,P3,P4
   52 FORMAT(4E18.10)
      KY=0
C     JJ, THE NUMBER OF THE FIRST USED MISSION DATE, EQUALS WL + 1
      JJ=1
      WRITE(6,1000) JJ
 1000 FORMAT(1H ,2X,14HDELAY VALUE OF,I2)
   67 KY=KY+1
C     KM INDEXES THE G FUNCTIONS
      DO 101 KM=1,3
      WRITE(6,1001) KM
 1001 FORMAT(1H ,4X,1HG,I1,1X,8HFUNCTION)
C     KP INDEXES THE F FUNCTIONS
      KP=1
   21 ST=0.0
      NP=N+1
C     THE FOLLOWING STEPS COMPUTE THE RISK FUNCTION R
      DO 2 IB=JJ,NP
      TIB=IB
      PFA=0.0
      FA=0.0
      IF(IB.EQ.JJ) GO TO 10
      IF(P1.EQ.0.0) GO TO 10
      IC=IB-1
      DO 1 I=JJ,IC
      FA=FA+F(I)
    1 CONTINUE
      PFA=FA*P1/(P1+P2)
   10 CPFB=0.0
      G=0.0
      IF(IB.EQ.NP) GO TO 20
   74 CPFB=F(IB)*P3/(P3+P4)
      IF(KM-2)  102,103,104
  102 G=100.0
      GO TO 20
  103 G=  100.+.10752688*(930.-TIB*(TIB-1.0))
      GO TO 20
  104 G=1000.0
   20 TN=(PFA+CPFB-G)*P(IB)
      ST=ST+TN
    2 CONTINUE
C     TT IS THE TOTAL RISK R
      TT=100.0-ST
      WRITE(6,30) TT
   30 FORMAT(1H0,10X,E16.8)
      IF(KP.EQ.3)  GO TO 22
      KP=KP+1
      GO TO 21
```

```
   22 CONTINUE
  101 CONTINUE
      JJ=JJ+KY
      IF(JJ.GT.7)  GO TO 66
      GO TO 67
   66 CONTINUE
      CALL EXIT
      END

      FUNCTION F(I)
      COMMON IA,IT,IB,N,P1,P2,P3,P4,JJ,KP
      X=I
      IF(KP.EQ.1) GO TO 31
      IF(KP.EQ.2)  GO TO 32
      IF(KP.EQ.3)  GO TO 33
C     THESE ARE F1,F2,AND F3, RESPECTIVELY
   31 F=X*.21505376
      GO TO 34
   32 F=3.33333333
      GO TO 34
   33 F=25.0313696/X
   34 CONTINUE
      RETURN
      END

      FUNCTION P(IB)
      COMMON IA,IT,IB,N,P1,P2,P3,P4,JJ,KP
C     THIS PART COMPUTES P(B)
      IF(IB.EQ.N+1) GO TO 11
      P=(P1+P2)**(IB-JJ)*(P3+P4)
      GO TO 12
   11 P=(P1+P2)**(N-JJ+1)
   12 CONTINUE
      RETURN
      END
```

```
THIS IS THE SLOT PROGRAM FOR VARIABLE NUMBERS OF MISSIONS PER SLOT AND MODIFIED
MISSIONS PARAMETERS FOR A TOTAL OF 18 MISSIONS IN THE UNMANNED LANDER PROGRAM
      DIMENSION P1(10),P2(10),P3(10),P4(10)
      COMMON IA,IT,IB,N,P1,P2,P3,P4,JJ,KP
C     N IS THE NUMBER OF SLOTS
      N=6
C     LL IS THE NUMBER OF V LEVELS, WHERE V=10.**(5-LL)
      DO 66 LL=1,12
      WRITE(6,999) LL
  999 FORMAT(1H0,3HLL=,I2)
      JA=7-LL
C     Q,R,S,AND T ARE THE MODIFIED MISSION PARAMETERS
      Q=.6375*(1.0-.00001085*10.**(-2+JA))
      R=.3625*(1.0-.00001085*10.**(-2+JA))
      S=.0085*.00001085*10.**(-2+JA)
      T=.9915*.00001085*10.**(-2+JA)
C     PI VALUES ARE FORMED ACCORDING TO THE CONVENTION OF OVER HALF MISSION
C     SUCCESSES FOR A SLOT SUCCESS
      P1(1)=Q
      P2(1)=R
      P3(1)=S
      P4(1)=T
      P1(2)=Q**2
      P2(2)=R**2+2.*Q*R
      P3(2)=2.*S*Q
      P4(2)=2.*T*R+2.*T*Q+2.*S*R
      P1(3)=Q**3+3.*Q**2*R
      P2(3)=R**3+3.*Q*R**2
      P3(3)=3.*S*(Q**2+2.*Q*R)+3.*T*Q**2
      P4(3)=3.*T*(R**2+2.*Q*R)+3.*S*R**2
      P1(4)=Q**4+4.*Q**3*R
      P2(4)=R**4+4.*Q*R**3+6.*Q**2*R**2
      P3(4)=S*(4.*Q**3+12.*Q**2*R)+T*(4.*Q**3)
      P4(4)=T*(4.*R**3+12.*Q*R**2+12.*Q**2*R)+S*(4.*R**3+12.*Q*R**2)
      P1(5)=P1(4)
      P2(5)=P2(4)
      P3(5)=P3(4)
      P4(5)=P4(4)
      P1(6)=P1(4)
      P2(6)=P2(4)
      P3(6)=P3(4)
      P4(6)=P4(4)
C     JJ, THE NUMBER OF THE FIRST USED MISSION DATE, EQUALS WL + 1
      DO 67 JJ=1,2
      WRITE(6,1000) JJ
 1000 FORMAT(1H ,2X,14HDELAY VALUE OF,I2)
C     KM INDEXES THE G FUNCTIONS
      DO 101 KM=1,3
      WRITE(6,1001) KM
 1001 FORMAT(1H ,4X,1HG,I1,1X,8HFUNCTION)
C     KP INDEXES THE F FUNCTIONS
      KP=1
   21 ST=0.0
      NP=N+1
C     THE FOLLOWING STEPS COMPUTE THE RISK FUNCTION R
      DO 2 IB=JJ,NP
      TIB=IB
      FA=0.0
      IF(IB.EQ.JJ) GO TO 10
      IC=IB-1
      DO 1 I=JJ,IC
      FA=FA+F(I)*P1(I)/(P1(I)+P2(I))
    1 CONTINUE
   10 CPFB=0.0
```

```
      G=0.0
      IF(IB.EQ.NP) GO TO 20
   74 CPFB=F(IB)*P3(IB)/(P3(IB)+P4(IB))
      IF(KM-2)  102,103,104
  102 G=100.0
      GO TO 20
  103 G=  100.+2.38*(42.0-TIB*(TIB-1.0))
      GO TO 20
  104 G=1000.0
   20 TN=(FA+CPFB-G)*P(IB)
      ST=ST+TN
    2 CONTINUE
C     TT IS THE TOTAL RISK R
      TT=100.0-ST
      WRITE(6,30) TT
   30 FORMAT(1H ,10X,E16.8)
      IF(KP.EQ.3)  GO TO 22
      KP=KP+1
      GO TO 21
   22 CONTINUE
  101 CONTINUE
   67 CONTINUE
   66 CONTINUE
      CALL EXIT
      END

      FUNCTION F(I)
      DIMENSION P1(10),P2(10),P3(10),P4(10)
      COMMON IA,IT,IB,N,P1,P2,P3,P4,JJ,KP
      X=I
      IF(KP.EQ.1) GO TO 31
      IF(KP.EQ.2)  GO TO 32
      IF(KP.EQ.3)  GO TO 33
C     THESE ARE F1,F2, AND F3, RESPECTIVELY
   31 F=X*4.76
      GO TO 34
   32 F=16.67
      GO TO 34
   33 F=(1./X)*40.82
   34 CONTINUE
      RETURN
      END

      FUNCTION P(IB)
      DIMENSION P1(10),P2(10),P3(10),P4(10)
      COMMON IA,IT,IB,N,P1,P2,P3,P4,JJ,KP
C     THIS PART COMPUTES P(B)
      IF(IB.EQ.N+1) GO TO 11
      ZP=P3(IB)+P4(IB)
      IF(IB.EQ.JJ)  GO TO 15
      IC=IB-1
      DO 14 K=JJ,IC
      ZP=ZP*(P1(K)+P2(K))
   14 CONTINUE
   15 P=ZP
      GO TO 12
   11 ZP=1.0
      DO 16 K=JJ,N
      ZP=ZP*(P1(K)+P2(K))
   16 CONTINUE
      P=ZP
   12 CONTINUE
      RETURN
      END
```

```
THIS IS THE SLOT PROGRAM FOR THREE MISSIONS PER SLOT AND SAGAN-COLEMAN MISSION
PARAMETERS FOR A TOTAL OF 18 MISSIONS TO FIND AN EQUALIZING G CONSTANT
      DIMENSION Z1(10),Z2(10),Z3(10),Z4(10)
      COMMON KP,P1,P2,P3,P4,PP1,PP2,PP3,PP4,N
C     N IS THE NUMBER OF SLOTS
      N=6
C     LL IS THE NUMBER OF V LEVELS, WHERE V=10.**(2-LL)
      DO 66 LL=1,8
      JA=2-LL
C     Q,R,S,AND T ARE THE SAGAN-COLEMAN MISSION PARAMETERS
      Q=.81*(1.-10.**(-2+JA))
      R=.19*(1.-10.**(-2+JA))
      S=.009*10.**(-2+JA)
      T=.991*10.**(-2+JA)
C     PI VALUES ARE FORMED ACCORDING TO THE CONVENTION OF TWO OF THE THREE
C     MISSIONS IN A SLOT SUCCESSES FOR A SLOT SUCCESS
      P1=Q**3+3.*Q**2*R
      P2=R**3+3.*R**2*Q
      P3=3.*Q*(Q*S+Q*T+2.*R*S)
      P4=3.*R*(R*S+R*T+2.*Q*T)
      WRITE(6,52) P1,P2,P3,P4
   52 FORMAT(4E18.10)
      Z1(LL)=P1
      Z2(LL)=P2
      Z3(LL)=P3
      Z4(LL)=P4
   66 CONTINUE
      NP=N+1
      DO 7 K=1,NP
      KP=0
      KK=K+1
      P1=Z1(K)
      P2=Z2(K)
      P3=Z3(K)
      P4=Z4(K)
C     PPI ARE THE PI* PARAMETERS
      PP1=Z1(KK)
      PP2=Z2(KK)
      PP3=Z3(KK)
      PP4=Z4(KK)
    6 KP=KP+1
      QS=0.0
      PS=0.0
C     THE FOLLOWING STEPS COMPUTE THE CONSTANT G FUNCTION
      DO 11 IB=1,N
      Z=P(IB)-PP(IB)
      QS=QS+F(IB)*Z
      PS=PS+Z
   11 CONTINUE
      QF=QS*P3/(P3+P4)
      TS=0.0
      DO 12 IB=2,NP
      IN=IB-1
      T=0.0
      DO 3 IU=1,IN
      T=T+F(IU)
    3 CONTINUE
      TS=T*(P(IB)-PP(IB))+TS
   12 CONTINUE
      TT=TS*P1/(P1+P2)
C     G IS THE VALUE SOUGHT
      G=(QF+TT+(P1/(P1+P2))*F(1))/PS
      WRITE(6,9) K
    9 FORMAT(1H0,I2)
```

```
      WRITE(6,8) G
    8 FORMAT(1H0,E16.8)
      IF(KP.LE.2) GO TO 6
    7 CONTINUE
      CALL EXIT
      END

      FUNCTION F(I)
      COMMON KP,P1,P2,P3,P4,PP1,PP2,PP3,PP4,N
      X=I
      IF(KP.EQ.1) GO TO 31
      IF(KP.EQ.2)  GO TO 32
      IF(KP.EQ.3)  GO TO 33
C     THESE ARE F1,F2,AND F3, RESPECTIVELY
   31 F=X*4.76
      GO TO 34
   32 F=16.67
      GO TO 34
   33 F=(1./X)*40.82
   34 CONTINUE
      RETURN
      END

      FUNCTION P(IB)
      COMMON KP,P1,P2,P3,P4,PP1,PP2,PP3,PP4,N
C     THIS PART COMPUTES P(B)
      IF(IB.EQ.N+1) GO TO 11
      P=(P1+P2)**(IB-1)*(P3+P4)
      GO TO 12
   11 P=(P1+P2)**N
   12 CONTINUE
      RETURN
      END

      FUNCTION PP(IB)
      COMMON KP,P1,P2,P3,P4,PP1,PP2,PP3,PP4,N
C     THIS PART COMPUTES P*(B)
      IF(IB.EQ.N+1) GO TO 111
      IF(IB.EQ.1) GO TO 110
      PP=(PP1+PP2)**(IB-2)*(PP3+PP4)
      GO TO 112
  110 PP=0.0
      GO TO 112
  111 PP=(PP1+PP2)**(N-1)
  112 CONTINUE
      RETURN
      END
```