N 70 42727
CR 110902

PLANETARY QUARANTINE

CASE FILE
COPY

SC-RR-70-561
August 1970

A STOCHASTIC APPROACH TO BIOBURDEN ESTIMATION AND PREDICTION-- A PRELIMINARY REPORT

A. L. Roark
Planetary Quarantine
Systems Studies Division 1741

SANDIA LABORATORIES

OPERATED FOR THE UNITED STATES ATOMIC ENERGY COMMISSION BY SANDIA CORPORATION | ALBUQUERQUE, NEW MEXICO, LIVERMORE, CALIFORNIA

LEGAL NOTICE

This report was prepared as an account of Government sponsored work. Neither the United States, nor the Commission, nor any person acting on behalf of the Commission:

A. Makes any warranty or representation, expressed or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately owned rights; or

B. Assumes any liabilities with respect to the use of, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this report.

As used in the above, "person acting on behalf of the Commission" includes any employee or contractor of the Commission, or employee of such contractor, to the extent that such employee or contractor of the Commission, or employee of such contractor prepares, disseminates, or provides access to, any information pursuant to his employment or contract with the Commission, or his employment with such contractor.

SC-RR-70-561

A STOCHASTIC APPROACH TO BIOBURDEN ESTIMATION AND PREDICTION-- A PRELIMINARY REPORT*

A. L. Roark
Planetary Quarantine
Systems Studies Division 1741
Sandia Laboratories, Albuquerque

August 1970

ABSTRACT

This report motivates and describes an approach to modeling bioburdens that has many of the properties generally recognized as desirable for such models. A probability distribution for the bioburden on a surface, at any time t, is derived in closed form. This distribution depends upon the way in which organisms are "clumped" and the deposition rate and removal percent of these clumps.

Examples are given of how this model may be used to estimate and predict bioburdens and specify confidence limits about these estimates and predictions. An indication is also given of how the model may be used to establish sampling protocols.

*This work was conducted under Contract No. W-12-853 Bioscience Division, Office of Space Science and Application, NASA Headquarters, Washington, D. C.

CONTENTS

		Page
I.	Introduction	5
II.	Discussion of the Physical Problem	8
III.	The Approach to Modeling	11
IV.	Estimation Model	14
V.	Prediction Model	19
VI.	Interaction Between Models	21
VII.	Future Areas of Interest	30
APPENDIX A - Mathematical Derivation of the Basic Approach		33
APPENDIX B - Solutions, $P_{ij}(\tau, t)$, and Their Properties		39
APPENDIX C - The Model and Its Implementation		51
REFERENCES		57

FIGURES

		Page
1.	The Physical Problem	11
2.	An Illustration of the Use of the Prediction Model	21
3.	Stainless Strip Data from Cape Kennedy	23
4.	Prediction Based on Three Samples	24
5.	Prediction Based on Four Samples	25
6.	Prediction Based on Five Samples	25
7.	Prediction Based on Six Samples	26
8.	Prediction Based on Seven Samples	26
9.	Prediction in "Uniform" Environment, Three Samples	27
10.	Prediction in "Uniform" Environment, Four Samples	28
11.	Prediction in "Uniform" Environment, Five Samples	28
12.	Prediction in "Uniform" Environment, Six Samples	29
13.	Nonconstant Environmental Data	29

A STOCHASTIC APPROACH TO BIOBURDEN ESTIMATION AND PREDICTION--A PRELIMINARY REPORT

I. Introduction

In almost any environmental situation viable microorganisms can be found on surfaces. These microbes are placed on surfaces by fallout from the surrounding atmosphere, contact from humans, or contact with other surfaces. Recently concern has risen over this fact, due to NASA's commitment to a Planetary Quarantine Program. Part of this commitment is reflected in the requirement of a high probability that any piece of planetary space hardware due to land on certain planets should have a low probability of contaminating the planet with microorganisms.[1] This commitment currently implies that the hardware must have a high probability of being sterile. In order to establish a sterilization procedure, it is necessary to know the initial bioburden on the hardware at the start of sterilization.[2, 3, 4] Thus, several studies have been made to determine the numbers and kinds of microorganisms which are on surfaces in various environments.[5, 6, 7, 8] Most of these studies are based on microbiological assays of surfaces. Since it is not feasible to assay an entire spacecraft, sampling techniques must be employed. These techniques utilize direct assays of small portions of the surface of the spacecraft and a characterization of the environment around the spacecraft. This fact necessitates a mathematical model which yields a probability distribution for the bioburden. Such a model has become known as an <u>estimation or direct assay microbial model</u>.

Actual samples, in most cases, also cannot be taken on all surfaces completely up to the actual start of sterilization. It is then necessary to extrapolate forward in time from previous estimates (based on surface assay data) using environmental data. A model which has this ability is known as a <u>prediction model</u>.

The specific type of model which is needed is dictated by how such a model must be used. This in turn is dictated by the responsibilities imposed upon the user by NASA policy. Let us for a moment consider the Planetary Quarantine Provisions for a Martian lander.[9]

The first place where a model should be capable of assisting in meeting the Planetary Quarantine Provision requirements is in the various planning documents. The contractor is required to submit a microbiological assay and monitoring plan. Among other things, this document is required to describe the number of samples to be taken, the location from which these specimens are taken, and the time schedules for the sampling. In order to assist in this, the bioburden estimation and prediction models must possess a form which lend themselves readily to the planning of sampling protocol and hypothesis testing. This will enable the contractor to plan the sampling so that enough samples are taken from the necessary areas at the right times to obtain a confident estimate of the bioburden. This will also prevent him from taking more samples than are necessary. As a constraint on his plans, the NASA Project Office will carry on an independent sampling program. This second set of samples is taken to verify the numbers provided by the contractor. The contractor must allow for this in his planning. Since the project office is only attempting to monitor the sampling done by the contractor, they should require fewer samples. This would again imply that the models that are developed should be capable of being used in this type of hypothesis testing.

The contractor must also provide a Decontamination Plan. This plan may contain the justification and an analysis of the need for decontaminating various pieces of hardware. Bioburden models should be ones which could be used to determine this need and to determine the optimal time for decontamination and amount of decontamination required at these times.

Other planetary quarantine provisions for Martian landers which impose properties on bioburden models are related to the accuracy of choosing the parameters in the models. The parameters must be capable of being demonstrated as valid.

The contractor must also verify that the levels of numbers of microorganisms on the spacecraft are within specified limits at the beginning of sterilization. This,

together with the certification of the level of sterility required from the Planetary Quarantine Officer prior to launch, implies that the bioburden models used be of a form so that some level of confidence can be attached to their predictions and estimates.

In addition to the above document, there is another recent publication which should be kept in mind while developing a bioburden model. This document is the review of the JPL-Martin report by the Planetary Quarantine Advisory Committee.[10] A summary of some of their more important suggestions are as follows:

1. The committee recommended that "direct assay techniques and a model capable of integrating direct assay measurements into a bioload estimate be developed for the design of sterilization processes, and a prediction model, validated and verified on a continuing basis by direct assay methods, be used to complement direct assay estimation where the latter is not applicable because of constraints on accessibility in terms of spacecraft locations or in time."

2. The models should serve as management tools in pointing out activities and ranking them by priorities with a view to controlling the microbial contamination on the spacecraft before assembly. They should also be capable of being used during the actual assembly process to facilitate decisions concerning control measures which might be necessary to keep the bioload within specified limits at sterilization time.

3. The model should take into account the clumping of microorganisms and allowance should be made for the size distribution of these clumps.

4. The model should be capable of being verified by setting up an experimental situation in which all the parameters in the model are measured (or controlled).

5. There should be a sensitivity analysis performed on the model.

6. The parameters should be determined on the basis of "well-defined" procedures and some information be provided regarding the "goodness" of the parameter values.

7. If the direct assay and the predicted value are inconsistent, there should be some strategy for deciding what to do.

8. The assay model and the prediction model should be "viewed as a single entity."

The purpose of this document is to develop a bioburden estimation and prediction model which is capable of meeting these objectives.

II. Discussion of the Physical Problem

While the problem of predicting and estimating bioburdens on surfaces can be viewed as an extremely complex one, there are a few basic principles that are readily observed.

The first principle is reflected by the fact that several investigators have observed that if a surface is left in an environment for a reasonable period of time an equilibrium is reached in the bioburden on this surface.[11, 12] This phenomena has become known as the "plateau" effect. If there were no loss of organisms from surfaces, one would expect the bioburden to increase steadily in a given environment. Since this does not happen, and a plateau is observed, there must be some loss of viable organisms from the surface. This is undoubtedly explained in part by the death of organisms on the surface, but since the plateau has been observed when only bacterial spores are present one suspects that there are removal mechanisms other than death. These mechanisms are termed physical removal and include blow-off, contact removal, and so forth. This discussion leads naturally to the first basic principle associated with bioburden estimation and prediction, namely, that viable particles are removed from surfaces. Any model should take into account both modes of removal of microorganisms from surfaces--removal through a loss of viability or through physical removal. Both of these modes depend upon many factors. Among these are[13, 14, 15, 16]

1. Surface characteristics such as material, angle, roughness, contaminants, conductivity, etc.

2. Environmental gas properties such as composition, temperature, humidity, particle size, etc.

3. External force fields.

In spite of these many factors and our lack of understanding of them, we will be able to make a statement regarding the result of the combined effects of all of

them. Before we can discuss it, however, we need to consider the second principle involved in predicting and estimating bioburdens.

In the Introduction we referred to the fact that microorganisms in environments and on surfaces tend to occur in clumps. This is the second important principle in our analysis. There is both direct and indirect evidence to support this observation. The indirect evidence is supplied by the fact that there is physical removal of microorganisms from surfaces. So-called "naked" organisms (those unattached to other particles or organisms) are extremely difficult to remove physically due to their very small size. When attached to larger particles their removal is much more easily effected, but larger particles such as skin flakes, lint, and dust particles are likely to have upon them more than one organism. This gives us our indirect evidence. Direct experimental evidence indicates that the expected number of organisms in such an organism-bearing clump is a small number between 1 and 10 depending upon the environment. Unpublished Public Health Service data indicate a mean number of organisms per clumps of about 4 to 10 in three environments.[17] More sophisticated experimentation is now being done by the same group. Similar experimentation at Sandia indicates a mean number between 1 and 2 in research areas.

Because of these observations, clumps of organisms (whether attached to an ambient particle or not) may be thought of as particles. They will be termed <u>viable particles</u>.

Let us return again for a moment to the removal of microorganisms from surfaces. If we assume that microorganisms are removed as viable particles, then the number removed should be directly proportional to the number of viable particles on the surface. The proportionally constant, $\mu(t)$, is the result of both modes of removal. It will depend on all of the factors we have discussed.

Even though this dependence is not well understood, it would currently appear that the proper level at which to describe removal of viable particles from surfaces is using only the notion of a removal fraction of the number on the surface as a function of time and incorporate into this concept <u>all</u> forms of removal. This is precisely the approach taken in this document.

Our last basic principle relating to bioburden modeling is that viable particles are deposited on a surface. This again can be from two basic sources, the environment and contact. The number deposited on a surface in either of these modes again depends upon many factors. Among these are:

1. Collection efficiency of surface
2. Concentration of viable particles in surrounding environment
3. Areas of contacts
4. Types of contacts
5. External force fields.

Furthermore, as with removal, the effects of these many factors are not well understood. In the case of deposition, the basic observation that may be made is that, independent of the mode of deposition and all of the factors listed above, there is some rate, $\lambda(t)$, at which viable particles are deposited on a surface. This rate will obviously change as a function of time, and, itself depend (in an unknown way) upon the above factors. Nevertheless, a deposition rate, representing deposition from all sources, does exist. Because of the lack of understanding of the dependence of the deposition rate on the many factors which must affect it, the best level at which to describe deposition of viable particles currently appears to be through the use of a function, $\lambda(t)$, incorporating deposition from all sources. Again this is the approach taken in this document.

In summary, Figure 1 describes the problem pictorially. Here it is shown that organisms occur in clumps that become deposited on the surface from the environment or through contact at a rate $\lambda(t)$ particles per unit time at time t and are being removed, through death or physically, at some rate $\mu(t)N(t)$ per unit time at time t. The function $N(t)$ represents the number of viable particles present at time t. Here the proportionality factor, $\mu(t)$, is the removal fraction.

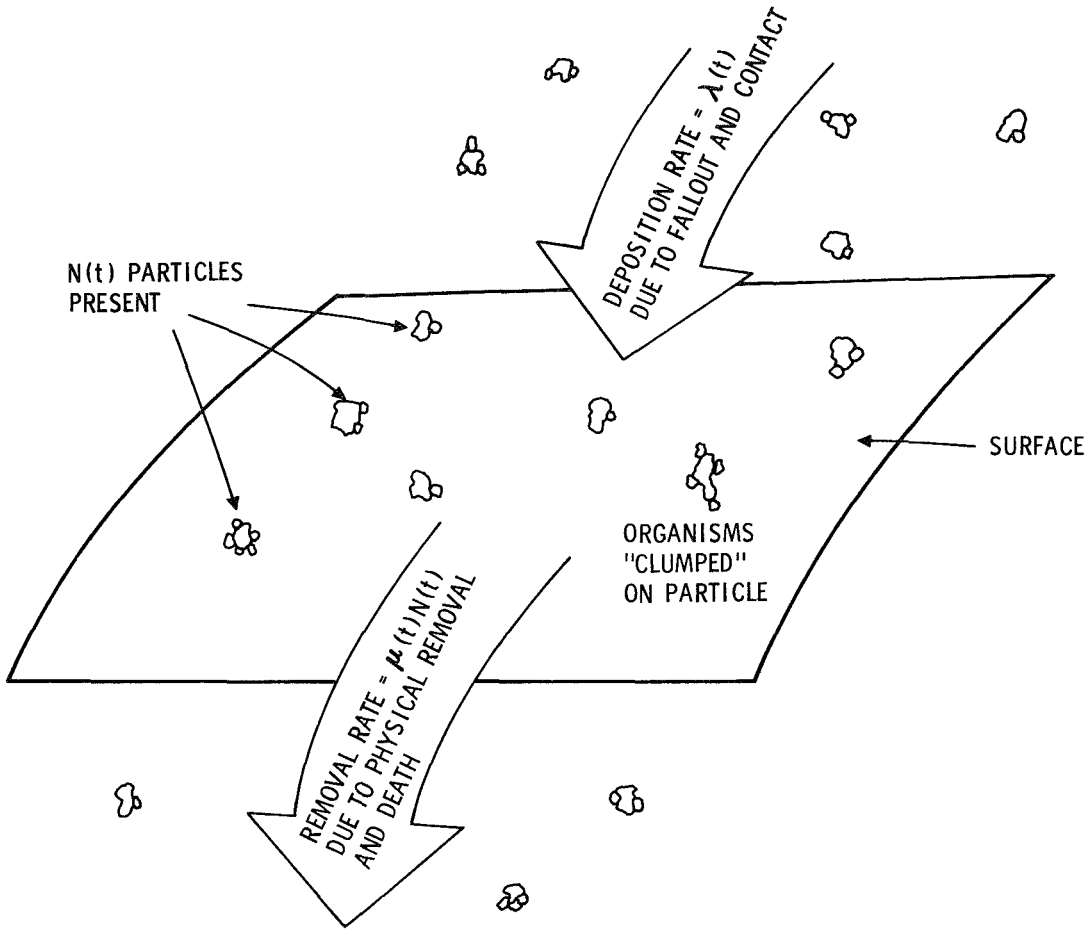


Figure 1. The Physical Problem

III. The Approach to Modeling

In the previous section we attempted to show that there are three basic entities that determine the number of organisms on a surface at a given time. First, organisms should be viewed as occurring in clumps, termed viable particles. Roughly, the number of organisms per clump is one of these basic entities. Then the rate of deposition of these clumps (from all sources) onto the surface and the rate of their removel (in any way) are the other two. Intuitively, knowledge of these three entities over a long period of time should be sufficient to allow one to estimate or predict the bioburden on a surface at any given time. The primary purpose, then, of a bioburden model is to describe the relationships which exist between these three

basic entities and the bioburden on the surface. There are secondary purposes related primarily to sampling which should become evident later in this document.

The form the model takes is dictated by two conditions. The first is the problem itself. The deposition and removal of microorganisms on surfaces is a random process. Thus the model should be stochastic rather than deterministic in nature.

The second condition is dictated by the requirements outlined in the Introduction. In order to meet the requirements, a bioburden model should first be capable of yielding estimates and predictions of the expected or mean number of organisms on a surface for various times t. It is possible, however, that the probability that the true bioburden exceeds this predicted mean may be large. This implies then that a person needs more information than just the mean. Since the need for a bioburden model was brought about by the sterilization requirements for spacecraft, the type of additional information required should be dictated by these goals. In establishing sterilization cycles one of two approaches can be taken.

The first requires only an estimate of the actual bioburden. Since we desire to have a very low probability that the spacecraft is not sterile at the end of the sterilization cycle, we should have a low probability that the actual burden exceeds the value we use for the initial population assumed for sterilization. We would like to determine the probability that the actual bioburden exceeds any given value. Then if the probability that the actual burden exceeds, say ten times the mean of our distribution, is very low, one may sterilize for a burden of ten times the mean knowing it is very unlikely that he is incorrect. Using just the estimated mean value for the sterilization cycle could give incorrect assessments of the degree of sterility a large portion of the time! In order to obtain information describing the probability of having a bioburden greater than various multiples of the mean, it is necessary to know the probabilities of having various numbers of organisms on the surface at any given time in closed form.

The second method of determining sterilization cycles, proposed by Fredrickson,[18] actually uses this probability distribution as its input information regarding the initial sterilization population--but, in any event, it too requires knowledge of the whole probability distribution of the bioburden.

If we therefore let

$P_j(t)$ = the probability that there are j organisms on the surface of concern at time t

a bioburden model should, because of the above reasoning, be capable of representing these quantities as functions of our three basic entities.

In this document we will derive such a model. There are two approaches to the derivation which lead to the same model. The heuristic approach is used to derive the estimation and prediction models in Sections IV and V. The mathematical approach to the derivation, which is based on first principles, is given in Appendices A, B, and C. The basic expressions that form the model are

$$P_j(t) = \sum_{k=0}^{\infty} \frac{(M(t)/\gamma)^k e^{-M(t)/\gamma}}{k!} Q(j, k) \tag{1}$$

where

$M(t)$ is the expected number of organisms on the surface at time t,

γ is the expected number of organisms per viable particle,

$Q(j, k)$ is the probability that k viable particles contain exactly j organisms.

The latter two of these can be derived from a knowledge of the manner in which organisms are clumped. The deposition rate and removal fraction of particles become important since $M(t)$ satisfies the differential equation

$$M'(t) = \gamma \cdot \lambda(t) - \mu(t)M(t)$$

which can be solved in closed form for $M(t)$ as a function of γ, $\lambda(t)$ and $\mu(t)$. This function may then be substituted into Equation (1) above yielding the probabilities $P_j(t)$ as functions of γ, $Q(j, k)$, $\lambda(t)$ and $\mu(t)$ only--as desired.

IV. Estimation Model

Since a spacecraft is a large, complex piece of equipment, it is not surprising that different subsections of the spacecraft have different sterilization properties. This is due in part to such things as different heat conduction properties of materials, the mating of surfaces, shielding from radiation, and different decontamination properties. Let us call a subsection of a spacecraft's surface a microbial zone if its entire area has the same sterilization properties. In doing this conceptual partitioning of the surface we will also wish to take into account degrees of variation in the bioburden from one section to another. This means that in defining our microbial zones we will want to require that the entire area of each zone or subsurface sees the same microbial environment and has the same microbial characteristics. By the terms microbial environment and microbial characteristics we mean all of the characteristics of the surrounding environment and of the surface itself which effect the bioburden on the surface of the zone.

With this in mind, let us consider the problem of estimating the number of microorganisms on a microbial zone of a spacecraft's surface at some fixed time. One of the ways to approach this is to take some number of samples of α square inches each from this zone. Call this number of samples n. Let us assume that we do this and process the samples according to the NASA Standard Procedures for the Microbiological Examination of Space Hardware.[19] Let us assume that the kth sample contains x_k microorganisms. Calculating the sample mean we get

$$\bar{y} = \frac{1}{n} \sum_{k=1}^{n} x_k \ .$$

The question now arises: What are we estimating? We know that there is always some sampling error. Let us assume that there is a uniform sampling and assay efficiency of ϵ. Then if our microbial zone contains A square inches, we see that

$$\bar{x} = \frac{\bar{y}A}{\alpha\epsilon} \tag{2}$$

is an approximation to the mean of the distribution of the number of microorganisms in the zone. The reason we must consider a distribution for the numbers

of microorganisms in a zone rather than the actual number is because of the variability in the samples. This arises because the processes by which microorganisms get onto surfaces or are removed from surfaces are random phenomena.

Given $\bar{x}$, defined by Equation (2), as an approximation to the mean, we now ask: What is the distribution which describes the number of microorganisms in the zone? Many people have assumed that the microorganisms are spread over the surface in some uniform manner and thus $\bar{y}$ is an estimate of some representative α square inches.[20] This leads to a Poisson distribution for the number of microorganisms on the surface. If we assume that microorganisms are clumped around ambient particles, this distribution does not hold. The clumps may be distributed in some uniform manner, but the fact that when we find one microorganism we increase our probability of finding a second keeps the Poisson distribution from describing the number of microorganisms. This is the same thing which we would have if we were sampling to find the number of head of cattle in a section of the country since cattle exist in herds. If we sample a certain area and find one head we have a higher than normal probability of finding a second head. Thus, even though the herds may in some sense be uniformly distributed, we cannot count the head of cattle and simply extrapolate. One field we choose to count might not have any cattle in it. Thus, our variance (or spread) of our "clumped" distribution must be greater than that of a Poisson distribution[21] (see Appendix B for a rigorous discussion of this). This has a direct effect on the probability of being various distances from the mean.

Let us assume we know the distribution for "clumping" and let γ be the mean number of microorganisms per clump. Then we know if H is the mean number of clumps in the zone, then $H\gamma$ is the mean number of microorganisms in the zone. Thus, using Equation (2), we have that, approximately,

$$\bar{x} = H\gamma \ . \tag{3}$$

Since the clumps are assumed to be distributed in some uniform manner, we know that[22] the probability of having k clumps in the zone is given by

$$\text{Prob}\left\{k \text{ clumps in the zone}\right\} = \frac{H^k e^{-H}}{k!} \ , \quad k = 0, 1, 2, \ldots . \tag{4}$$

Let us suppose that our estimate of the mean, $\bar{x} = \ell$. If we know only that there are ℓ microorganisms in the zone we do not know the number of clumps. The ℓ microorganisms may arise from 1, 2, 3, or 5000 clumps. Therefore, we have

$$\text{Prob}\left\{\ell \text{ microorganisms in zone}\right\} = P_\ell =$$

$$\sum_{k=1}^{\infty} \text{Prob}\left\{k \text{ clumps in zone}\right\} \text{Prob}\left\{k \text{ clumps have } \ell \text{ microorganisms}\right\} \tag{5}$$

For notational convenience, define (see Appendix B)

$$Q(\ell, k) = \text{Prob}\left\{k \text{ clumps have } \ell \text{ microorganisms}\right\} . \tag{6}$$

Combining Equations (4), (5), and (5) we have

$$P_\ell = \sum_{k=0}^{\infty} \frac{H^k e^{-H}}{k!} Q(\ell, k) . \tag{7}$$

This distribution can be arrived at rigorously from first principles. See Appendix A for further details.

Analyzing this distribution, we see that indeed the mean number of organisms, M, satisfies

$$M = H\gamma \tag{8}$$

and, in addition, that the variance of the distribution, V, is given by

$$V = H\,(v + \gamma^2) \tag{9}$$

where v is the variance of the distribution of the number of microorganims per clump. (This is done in Appendix B.) It is also worth noting that if $v \neq 0$ and $\gamma \neq 0$, then

$$M < V \ ,$$

whereas in a Poisson distribution, these are equal.

Since we approximate the mean number of organisms, M, by our sampling mean $\bar{x}$, it is worthwhile to rewrite Equation (7) as

$$P_\ell = \sum_{k=0}^{\infty} \frac{(M/\gamma)^k e^{-M/\gamma}}{k!} Q(\ell, k) \ . \tag{10}$$

Equation (10) is therefore our estimation model in the form previously discussed in Equation (1).

In addition to estimating bioburdens, one of the main applications of this estimation model is as a management tool in establishing sampling protocol. Let us consider in the remainder of this section an example of this use.

We have already seen that changes in the distribution which describes the "clumping" of the microorganisms will result in changes in the variance of our distribution of microorganisms on surfaces even if the mean is kept constant. This should be reflected in our sampling procedure. Let us consider the types of changes which this can cause by doing a parametric analysis on the mean and the variance of the "clumping" distribution.

Using the type analysis used in Reference 23, we see that if we wish to require that

$$\text{Prob} \left[\frac{(\bar{x} - M)^2}{M} < \beta \right] \geq \theta$$

then we must know

$$\chi^2 \leq \frac{\beta M n}{V}$$

where n = number of samples, and where the parameter χ is defined by the equation $\phi(\chi) - \phi(-\chi) = \theta$ where ϕ is the standard normal distribution. Tables of values of χ are tabulated.

Using Equation (9) we then know that

$$n \geq \frac{\chi^2\left((v/\gamma) + \gamma\right)}{\beta} \quad .$$

We determine β by the requirement that $|\bar{x}-M| < \eta$. This would imply that

$$\beta > \frac{\eta^2}{M_{max}}$$

where M_{max} is the maximum mean number of organisms deemed possible.

For our example, let us assume we are considering the case where

$$\epsilon = 1$$

$$\theta = 0.90$$

$$\eta = 10^3$$

$$M_{max} = 10^7 \quad .$$

Table I then gives the number of square inches $n\alpha$ which must be sampled for various values of v and γ to obtain the accuracy specified by η and θ. We see that these factors are important in establishing the sampling protocol. The case where $\gamma = 1$, v = 0 would be the Poisson case. In a later document the entire area of sampling protocols will be explored in more detail, including how one tests hypotheses regarding the correctness of mean estimates for use in monitoring or verification activities such as those of the Planetary Quarantine Office or any Project Office.[24]

TABLE I

γ/v	0	1	2	3	4	5	10	20
1	27	54	81	100	135	162		
2	54	68	81	95	108	122		
3	81	90	99	108	117	126		
4	108	115	122	128	135	142		
5	135	140	146	151	156	162		
10							296	323
20							552	565

V. Prediction Model

A prediction model is a model which extrapolates the bioburden forward (or backward) in time. We would like, if possible, for the estimation model and the prediction model to be capable of being viewed as a single entity. Let us consider our estimation model. This model is given by Equation (7). This equation expresses the probability of ℓ microorganisms on our surface at some fixed time. If we consider this probability as varying with time and assume that the distribution of the number of microorganisms per clump is constant in time, then the only parameter which can vary is the mean number of clumps on the surface. This is consistent with our previous assumption that the number of microorganisms on the surface at some time is dependent on the number of clumps on the surface at that time. Using this we then see that the probability that ℓ microorganisms are on the surface of the zone at any time t should have the form

$$P_\ell(t) = \sum_{k=0}^{\infty} \frac{\left(H(t)\right)^k e^{-H(t)}}{k!} Q(\ell, k) \quad . \tag{11}$$

This provides heuristic motivation for the use of Equation (1) as a prediction model. In Appendix B we see that the model, derived from first principles, actually assumes

this form. All that remains to complete the development of our prediction model is to discuss the determination of the function $H(t)$.

The function $H(t)$ varies with time. Therefore, it is reasonable for us to first consider the rate of change of $H(t)$. The variation of $H(t)$ in time is due to the fact that new particles are deposited on the surface and other particles removed from the surface at various points in time. Obviously, if we are going to derive an expression for the rate of change of the mean number of particles on a surface, we must consider the deposition rate and the removal rate. Therefore, let $\lambda(t)$ be the deposition rate of particles onto our surface. The removal rate should depend on the number of particles on the surface at various instances in time as discussed earlier. We let $\mu(t)H(t)$ be the removal rate of particles from our surface. The net change in number of particles on the surface therefore should be the difference of these two functions. Hence, we have

$$H'(t) = \lambda(t) - \mu(t)H(t) \quad . \tag{12}$$

Appendix B derives this equation in a precise way which shows its consistency with our model.

Thus Equations (11) and (12) will serve as our prediction model. Let us consider in the remainder of this section an application of this prediction model.

Let us suppose that we are considering surfaces which are sterile at time $t = 0$. Assume those surfaces are inserted into a uniform environment such that 150 microorganisms per hour are deposited on an agar-covered surface of the same size as our surface. Suppose that we know the death rate for the microorganisms in this environment is 15 percent of the population per hour and that 6 percent are removed by physical means. Assume that the microorganisms are clumped according to a Poisson distribution with a mean of 2 microorganisms per clump.

In order to apply our model to this case we let

$$\lambda(t) = 150$$

$$\mu(t) = 0.21$$

in Equation (12). Using $\gamma = 2$ in Equation (8) we can obtain our predicted mean number of organisms, M(t), by solving Equation (12). Equation (9) would then enable us to calculate the variance of our predictions and thus establish confidence limits for various size samples. Figure 2 illustrates this. Here the curve is the predicted mean number of organisms, M(t), and the solid vertical lines are the 95-percent confidence limits for a sample size of 10 such samples.

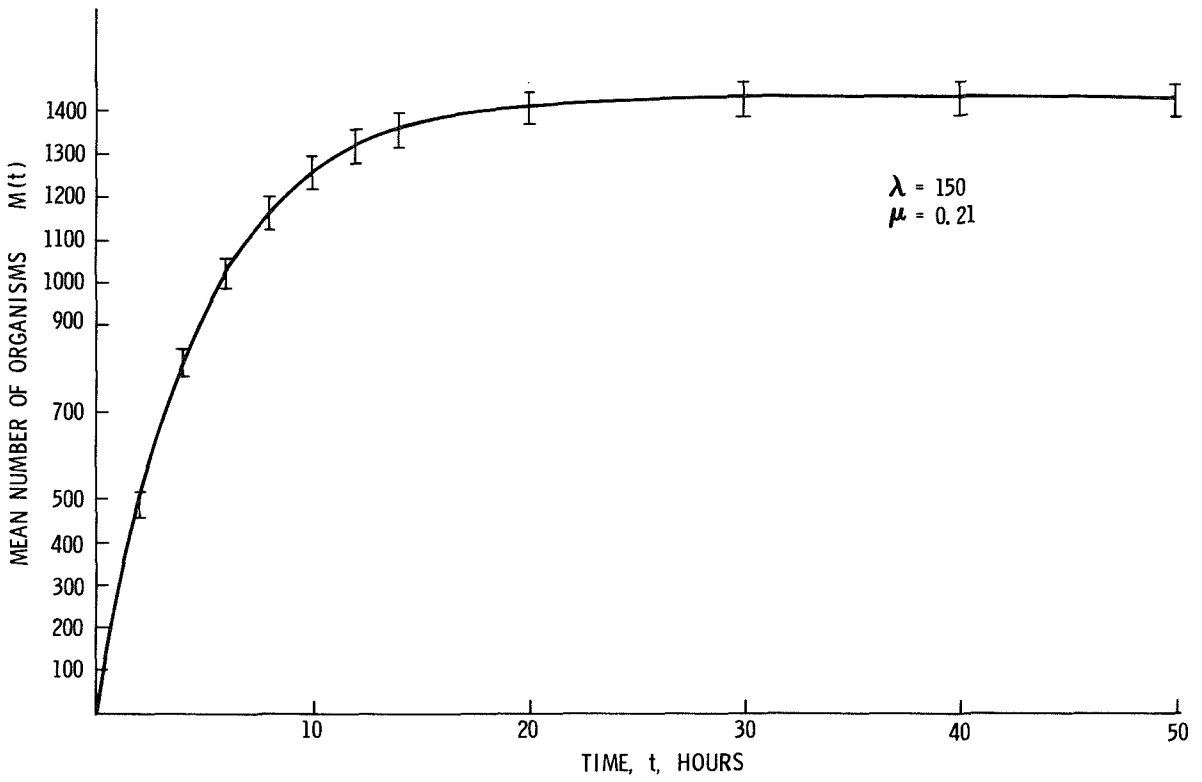


Figure 2. An Illustration of the Use of the Prediction Model

VI. Interaction Between Models

The fact that our estimation model has the same form as our prediction model at any given point in time allows us to use these models together to obtain information regarding the parameters in each model.

Let us proceed as we did in Section IV. Assume that we are sampling at times $t_1, t_2, \ldots, t_m$ and that we obtain the sample means at these times. Let us denote

these by $\bar{y}(t_1), \ldots,$ and $\bar{y}(t_m)$. By allowing for sampling error and the area of the zone we can obtain $\bar{x}(t_1), \ldots,$ and $\bar{x}(t_m)$ as in Equation (2). These are then approximations to the means of the distributions which describe our estimation model at time $t_1, t_2, \ldots, t_m$. In fact, as the number of square inches increases, these quantities converge to the mean of the distribution at each sample time. Given the form of this distribution and the mean as well as the number of samples, we can then establish confidence limits on our estimates at each of these times (see Appendix C).

Knowing the distribution, the mean, and the variance at each of a series of times, it would seem likely that we could then use this information to deduce the deposition rate, $\lambda(t)$, and the removal fraction, $\mu(t)$, during the time when the samples are taken. Indeed, using the techniques developed in Appendix C this can be done. Therefore, if the microbial environment around the surface in the future (beyond t_m) is the same (or similar) to that between t_1 and t_m, we can predict the burden into the future using the deposition and removal rates determined in the period t_1 through t_m. Again we can construct the confidence intervals about these predictions to give us some information regarding the accuracy of our predictions. Thus, it is possible to obtain our prediction using only some information about the number of microorganisms per particle in the environment and surface sampling data at previous times. If the environment in the future is not the same as that between t_1 and t_m we can then use this model to determine the relationships between the deposition rate, the removal fraction, and the properties of the environment discussed in Section II. We can then use these relations to extrapolate into the future. The first of these possibilities is illustrated below.

The examples we wish to discuss were chosen to illustrate the use of the prediction model and to demonstrate its feasibility. The data we shall use were gathered by United States Public Health Service at Cape Kennedy. It was collected using 1- by 2-inch stainless steel strips placed in hangar AO at the Air Force Eastern Test Range. The facility is a laminar downflow clean room which may be used for planetary missions. The particular data we shall use in this first example were gathered during the period from January 17, 1967, through March 30, 1967, from the location designated as "Tray 2." At each sampling period, six strips were processed according to the NASA Standard Procedures for the Microbiological

Examination of Space Hardware[19] and the mean number of microbes per strip is recorded. These means are shown in Figure 3.

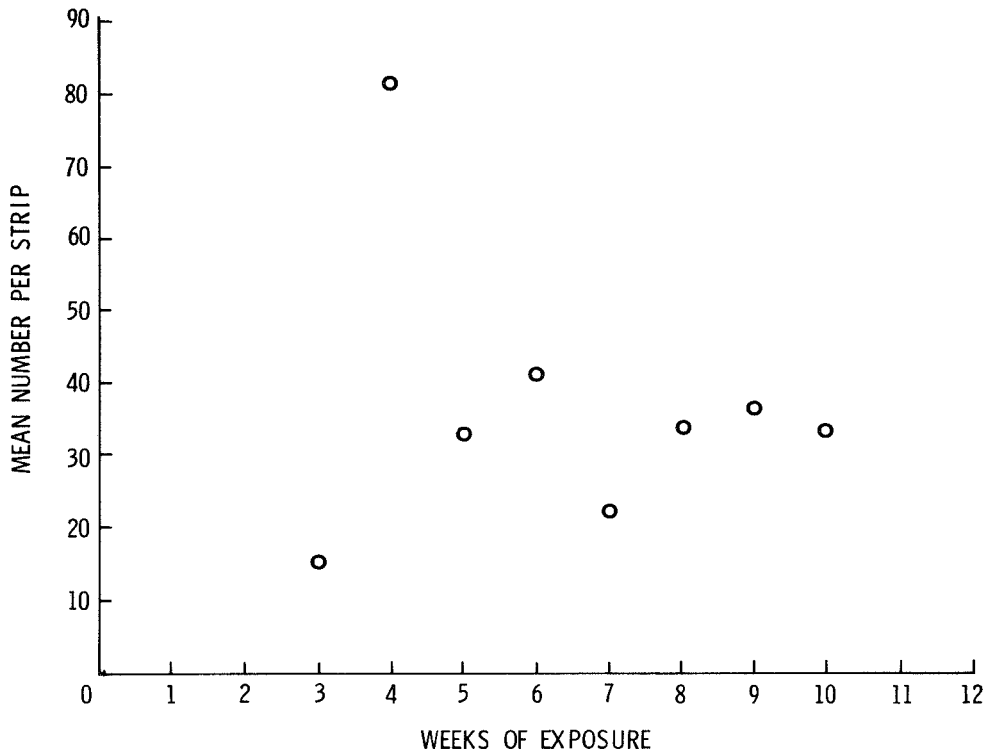


Figure 3. Stainless Strip Data from Cape Kennedy

For this example, let us assume that a microbial zone consists of a 1- by 2-inch stainless steel strip in Hangar AO. Thus our sample would be 100-percent assay of the entire zone. Since this is true, we shall assume that the sampling efficiency is one. This is probably not exactly correct, but it should be a good approximation.

We will restrict our consideration to the class of viable particles in the environment. This is helpful since data is available regarding the number of microbes per viable particle.[17] Let us take γ to be 4.0 and let the distribution describing the number of microorganisms per particle be Poisson (or at least have equal mean and variance). This is consistent with the data presented by the Public Health Service[17] and with the distribution used in References 23 and 25.

We will attempt to choose the deposition rate and removal fraction so that the value our model assumes at various times comes "close" to the data.* The easiest case for us to work with would be a uniform environment. Let us suppose that the microbial environment in hangar AO is uniform during the period of our samples. This means that $\lambda(t)$ and $\mu(t)$ are assumed to be constant. Figure 4 gives our results with these assumptions using the data at 3, 4, and 5 weeks as our sampling periods and predicting out to 10 weeks. Figures 5 through 8 llustrate what our predictions are as we add in additional sampling period data. The vertical marks give the 99-percent confidence intervals about the mean of our model for the mean of six samples.

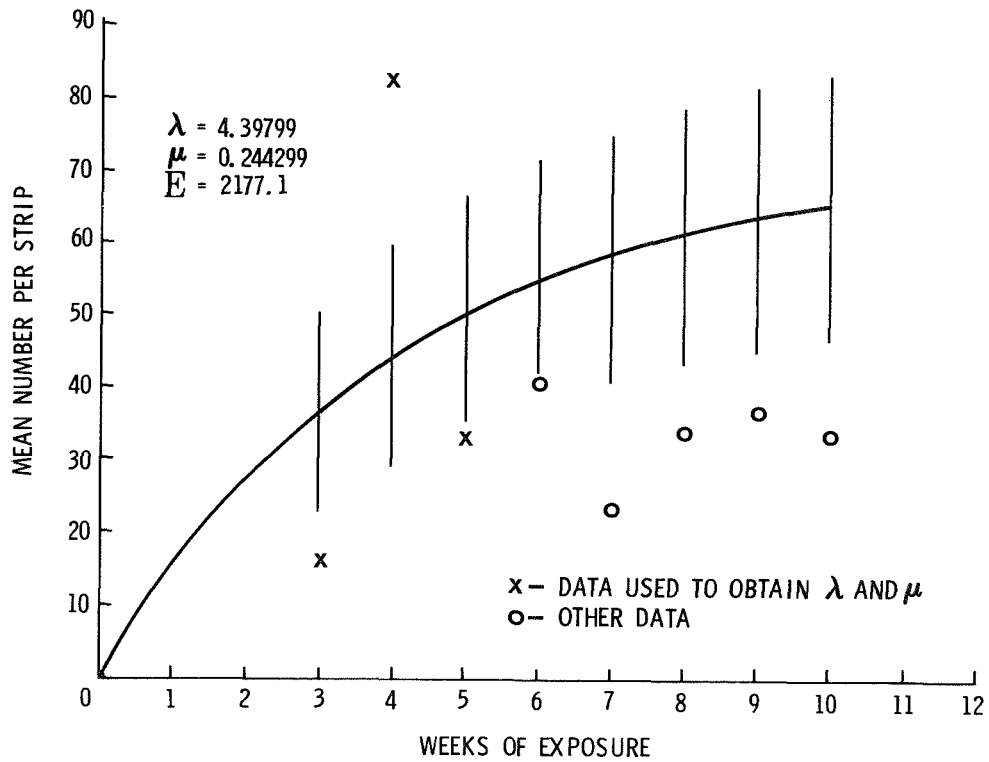


Figure 4. Prediction Based on Three Samples

*This is made precise in Appendix C.

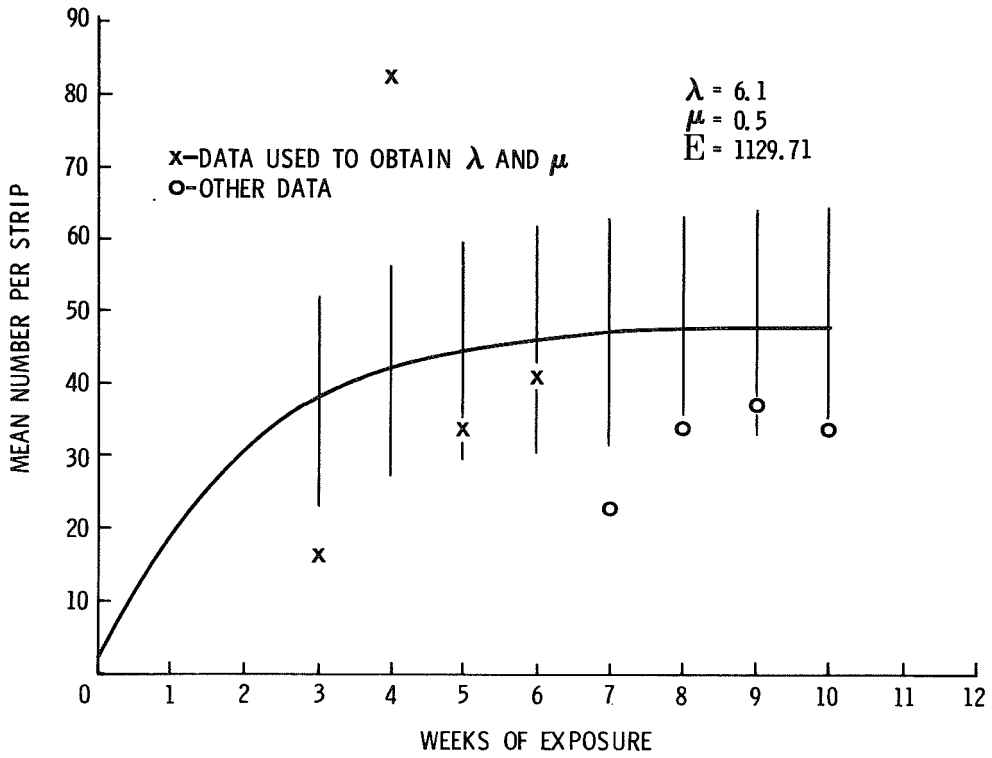


Figure 5. Prediction Based on Four Samples

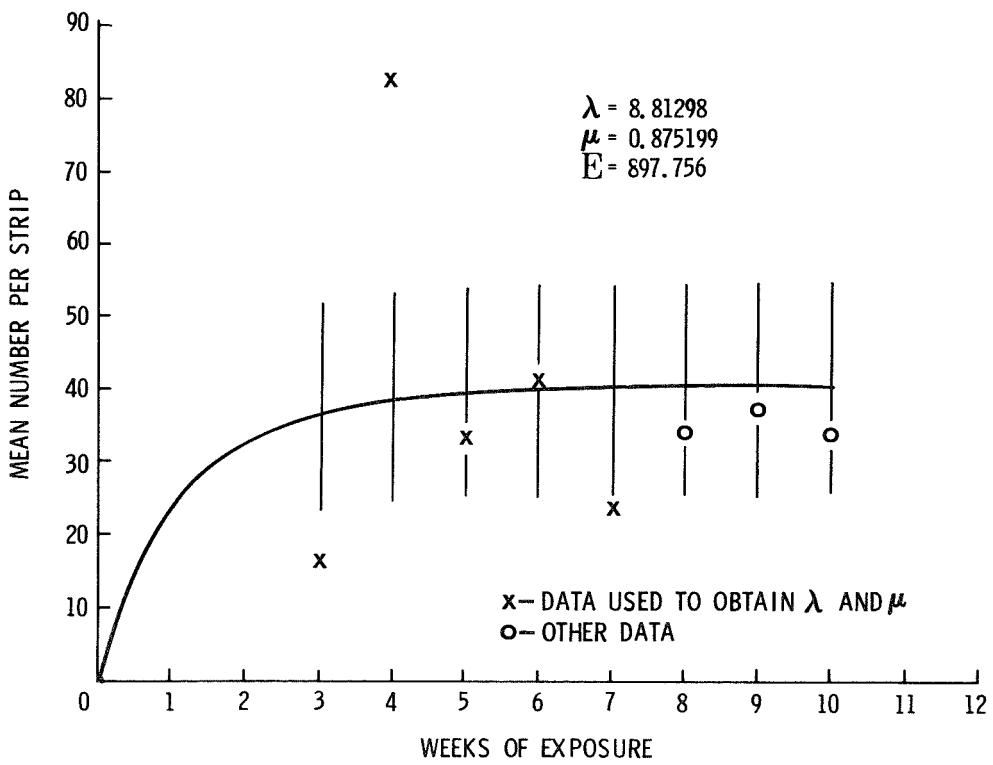


Figure 6. Prediction Based on Five Samples

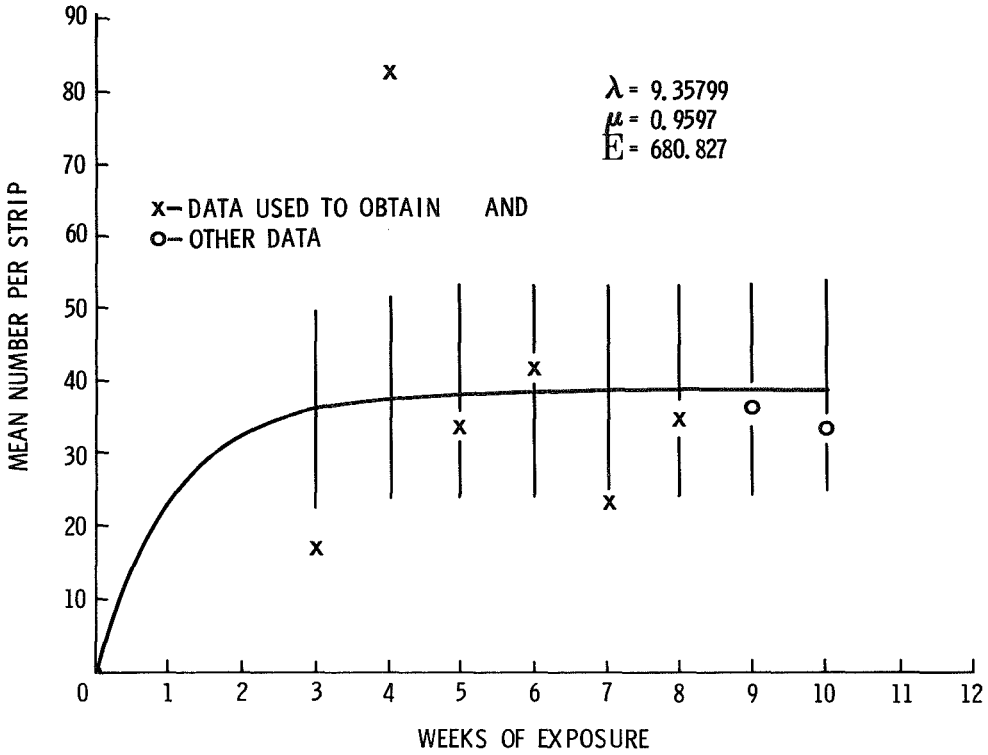


Figure 7. Prediction Based on Six Samples

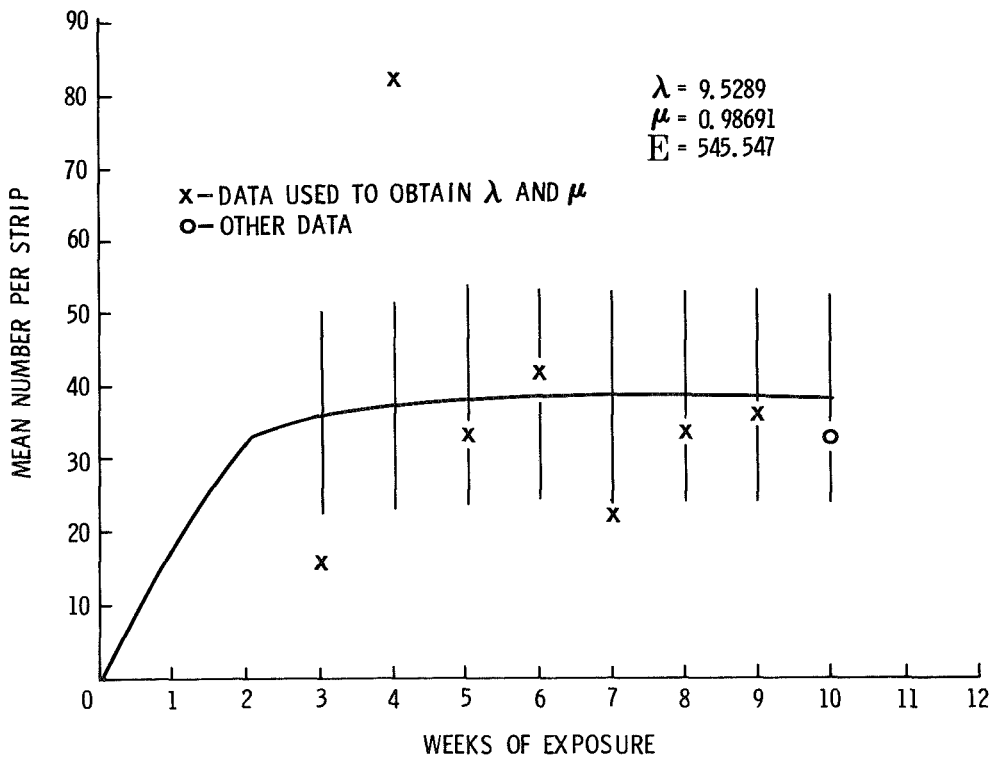


Figure 8. Prediction Based on Seven Samples

There are several observations which are in order. We see that the more data we have the better our predictions are at the tenth week of exposure. Secondly, the values of λ and μ appear to be converging to constant values. The value of the "goodness of fit" parameter represented by E appears to be getting smaller. This parameter is actually the variance between the measurements and our model. This would say the fit of the model to the data improved with additional points.

We observed earlier that the environment is not truly constant. The data point at t = 4 weeks appears to be inconsistent with this assumption. Let us eliminate that point and repeat the process again. Figures 9 through 12 repeat our prediction process beginning with the use of the data points at t = 3, 5, and 6 and adding one point at a time until we teach t = 9. We see that the predictions are much improved. This would say that if we had data from a truly constant environment the model would predict a bioburden behavior very close to the behavior actually observed.

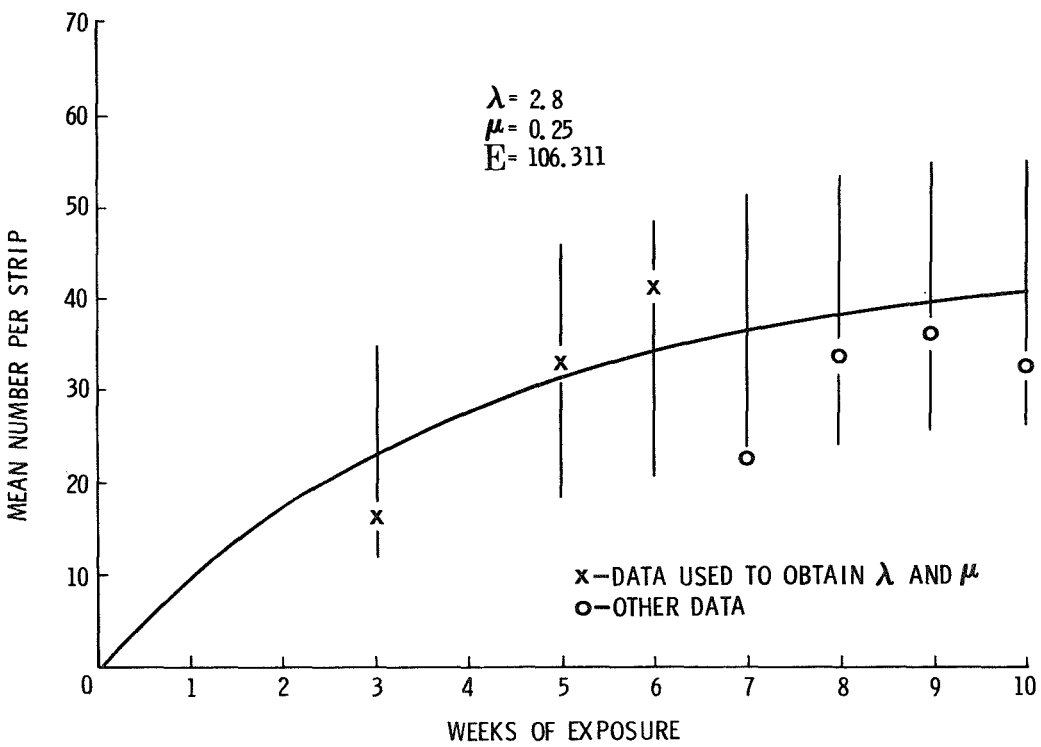


Figure 9. Prediction in "Uniform" Environment: Three Samples

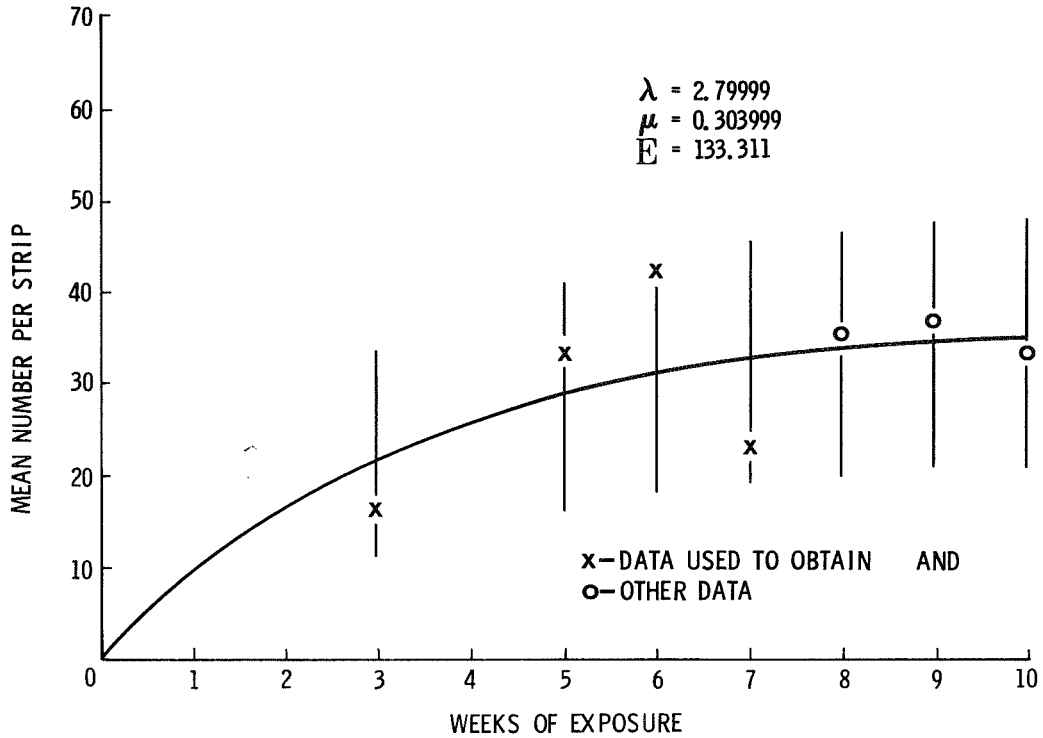


Figure 10. Prediction in "Uniform" Environment: Four Samples

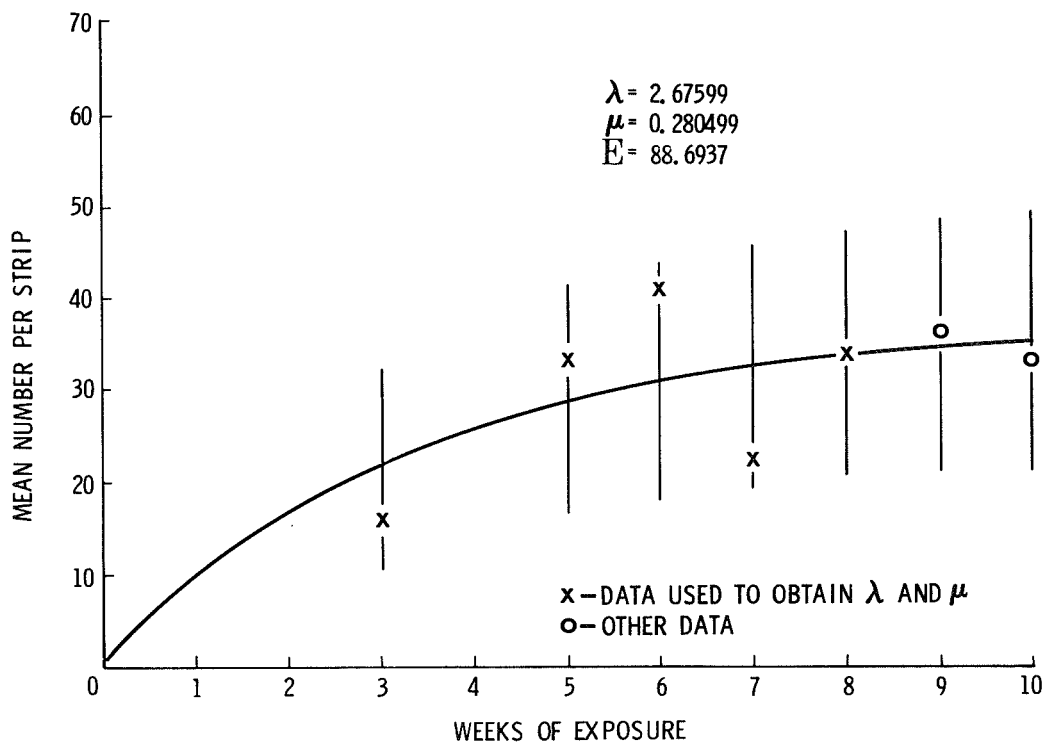


Figure 11. Prediction in "Uniform" Environment: Five Samples

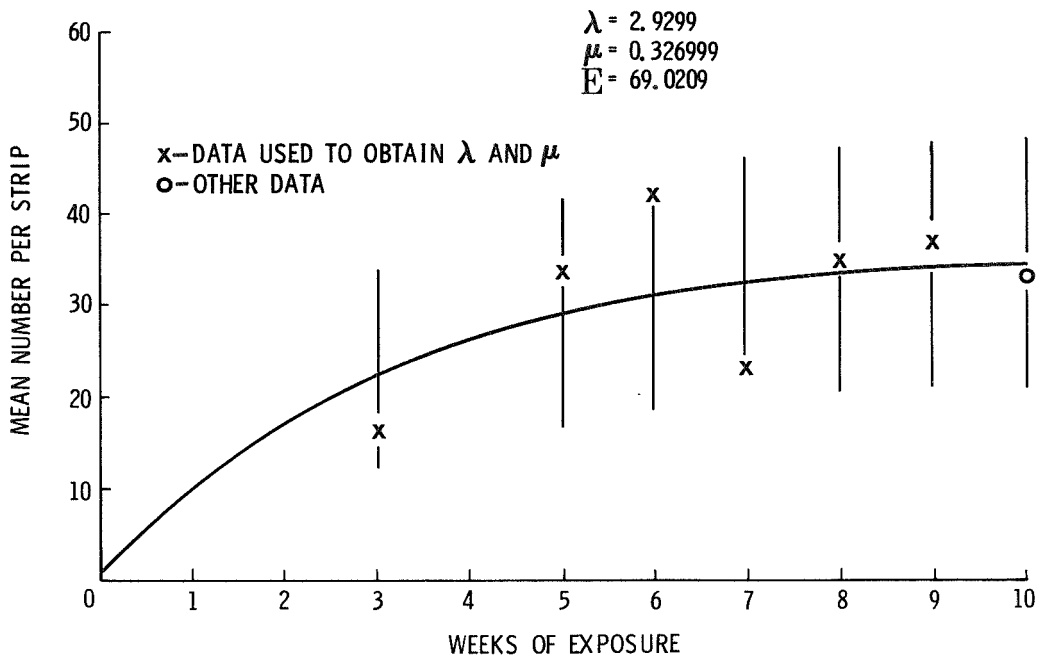


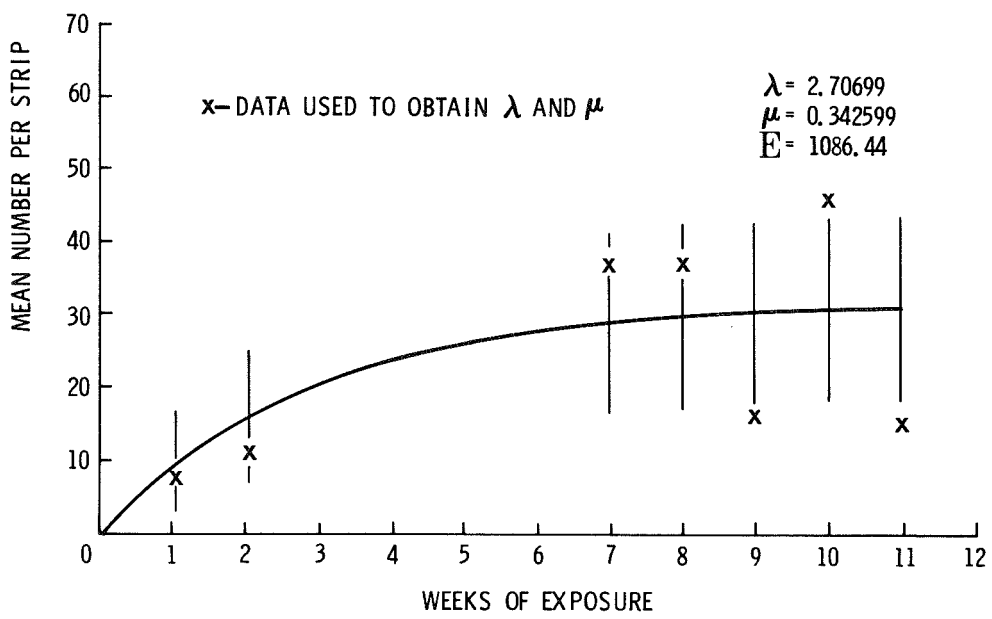
Figure 12.

Prediction in "Uniform" Environment: Six Samples

In order to illustrate the fact that the environment in AO is not constant, let us use the data gathered by the USPHS during the period from December 15, 1966, through February 2, 1967. We will use all 11 weeks of data and try to characterize the environment. Figure 13 illustrates our attempt to do this. We see that after the ninth week, the data does not come from a constant environment.

Figure 13.

Nonconstant Environmental Data



VII. Future Areas of Interest

The models we have presented in this document allow one to study the bioburden problem in more detail. It is intentionally a general approach to the problem so that as more information becomes available it can be added to the existing structure.

The models also have many of the properties we had hoped to achieve. Among these are:

1. A closed form description of the models.
2. Compatibility between the estimation and prediction models.
3. A method for varying the parameters in the model on the basis of surface samples.
4. A way to assign a confidence interval about the predicted value.
5. Clumping is taken into account and the distribution of the number of microorganisms per clump can be specified.
6. The model exhibits the plateau effect.
7. A way to establish sampling protocol.

There are, however, several studies which will be important in the Planetary Quarantine Program. This model should facilitate these studies. Among the more important of these are:

1. A model needs to be developed which will determine the deposition rate and the removal fraction rate as functions of other physical parameters. If we determine these functions by the procedure illustrated in Section VI, then data fitting techniques should help us attain this goal. This, of course, should lead to a more complete understanding of the problem.
2. The entire problem of establishing sampling protocols needs further study. As pointed out in the example of Section IV, the estimation model can be used to establish the number of samples needed at any given sampling period. The other part of establishing a sampling protocol is the establishment of how often the samples are to be taken. If we know an upper bound on the deposition rate and a lower bound on the removal fraction, we can use Equations (8), (11), and (12) to establish how often samples are to be

taken. This can be based on different criteria such as the mean changing by a certain percentage, etc.

We would also expect future studies to take into account other sampling error distributions, rather than being limited to a uniform distribution.

3. Sensitivity studies also need to be performed. These might also result in autocorrelation studies which could have some influence on Item 2. The sensitivity analysis could also lead to some simplification of the studies outlined in Item 1.

4. Controlled experiments where the parameters of the model can be measured independently are required for model validation. These are currently under way.[26]

All of these studies are within the reach of present-day methods and techniques. When they are completed they should lead to a comprehensive and usable bioburden model for Planetary Quarantine as well as for other applications.

APPENDIX A

MATHEMATICAL DERIVATION OF THE BASIC APPROACH

Mathematical Derivation of the Basic Approach

In this appendix we wish to derive a predictive bioburden model in a mathematical manner. Let us assume that we are considering a surface $\mathcal{S}$ of area A. Let us assume that $\mathcal{S}$ has been chosen so that every point of $\mathcal{S}$ is exposed to the same microbial environment and that every point has the same removal and attraction characteristics. Thus, the surface $\mathcal{S}$ could be a microbial zone of a spacecraft as defined earlier.

Since most microorganisms get on surfaces attached to particles (a naked organism will be assumed to be attached to itself), it is appropriate for us to consider particles of a class ψ. This class may be chosen in any of several ways. It may be chosen by deposition or removal characteristics, by size, by clumping characteristics, by the fact that all particles of the class carry microorganisms, or by any other characteristics which can be defined explicitly.

Before proceeding we will need the following definitions:

1. Let $X(t)$ = a random variable representing the number of microorganisms on surface $\mathcal{S}$ at time t due to particles of class ψ.

2. Let $P_{ij}(\tau, t) = \text{Prob}\left\{X(t) = j \,\middle|\, X(\tau) = i\right\}$. (A1)

3. Let Y = a random variable representing the number of microorganisms per particle in class ψ.

4. Let $q_j = \text{Prob}\left\{Y = j\right\}$. (A2)

5. Let $\lambda_m(t)\,h + 0\,(h)$ = probability that the deposition of a particle of class ψ on surface $\mathcal{S}$ occurs between t and t + h given that $X(t) = m$.

6. Let $\mu_m(t)\,h + 0\,(h)$ = probability that the removal from $\mathcal{S}$ of a particle from class ψ takes place between t and t + h given that $X(t) = m$.

These last two definitions require more explanation. The deposition of particles of class ψ onto $\mathcal{S}$ will include every mechanism by which microorganisms can get onto surfaces.

Therefore, the term $\lambda_m(t)\, h + 0\,(h)$ must represent the combined effects of all mechanisms. Logically we would not expect $\lambda_m(t)$ to actually depend on m. Thus, we shall assume $\lambda_m(t) = \lambda(t)$.

The term $\mu_m(t)\, h + 0\,(h)$ represents the combined effects of all removal mechanisms. We would expect $\mu_m(t)$ to be directly proportional to the number of clumps of microbes on $\mathcal{S}$ due to particles in ψ. Thus, we let

$\mu(t)\, nh + 0\,(h) =$ Probability that a removal from $\mathcal{S}$ of a particle from class ψ takes place between t and t + h given that there are n clumps on $\mathcal{S}$ due to ψ at time t.

Therefore, we have

$$\mu_m(t)\, h + 0\,(h) = \sum_{n=0}^{\infty} \text{Prob}\left\{\text{n particles of } \psi \text{ on } \mathcal{S} \text{ at time t given } X(t) = m\right\} \times \text{Prob}\left\{\text{Particle is removed between t and t + h given n particles on } \mathcal{S}\right\}$$

$$= \sum_{n=0}^{\infty} \text{Prob}\left\{\text{n particles of } \psi \text{ on } \mathcal{S} \text{ at time t given } X(t) = m\right\} \times \left(\mu(t)\, nh + 0\,(h)\right)$$

$$= \mu(t)\, h \sum_{n=0}^{\infty} n\, \text{Prob}\left\{\text{n particles of } \psi \text{ on } \mathcal{S} \text{ at time t given } X(t) = m\right\} + 0\,(h)$$

$$= \mu(t)\, hE_m(t) + 0\,(h) \qquad \text{(A3)}$$

where (A3)

$$E_m(t) = \sum_{n=0}^{\infty} n \text{ Prob } \{n \text{ particles of } \psi \text{ on } \mathcal{S} \text{ at time } t \text{ given } X(t) = m\}$$

is the expected number of particles of ψ on $\mathcal{S}$ at time t given $X(t) = m$.

We are now prepared to derive the model. We shall assume that microorganisms due to particles of class ψ are deposited on $\mathcal{S}$ in clumps and that they are removed in clumps. We shall assume Y represents the number of microorganisms on particles in the environment as well as on $\mathcal{S}$. This last assumption is not too restrictive if the number of clumps on $\mathcal{S}$ is large. Observe that the probability that a change in the microbial load on $\mathcal{S}$ takes place between t and t + h is given by

$$1 - P_{jj}(t, t + h) = \left(\lambda(t) + \mu(t)\, E_j(t)\right) h\,(1 - q_0) + 0\,(h)\,,$$

and thus

$$\lim_{h \to 0} \frac{1 - P_{jj}(t,\ t + h)}{h} = \left(\lambda(t) + \mu(t)\, E_j(t)\right)(1 - q_0) \quad . \tag{A4}$$

Similarly we have

$$\lim_{h \to 0} \frac{P_{jk}(t,\ t + h)}{h} = \begin{cases} \lambda(t)\, q_{k-j} & \text{if} \quad k > j \\ \mu(t)\, E_j(t)\, q_{j-k} & \text{if} \quad j > k \ . \end{cases} \tag{A5}$$

Let us assume that the passage to the limit in this equation is uniform.

The Chapman-Kilomogorov equation can be stated as[27]

$$P_{ik}(\tau,\ t + h) = \sum_j P_{ij}(\tau,\ t)\, P_{jk}(t,\ t + h) \quad . \tag{A6}$$

Combining this with Equations (A4) and (A5) and the definition of derivative we obtain

$$\frac{\partial P_{ik}(\tau, t)}{\partial t} = - \left(\lambda(t) + \mu(t)\, E_k(t)\right) (1 - q_0)\, P_{ik}(\tau, t)$$

$$+ \sum_{j=k+1}^{\infty} \mu(t)\, E_j(t)\, P_{ij}(\tau, t)\, q_{j-k} \tag{A7}$$

$$+ \sum_{j=0}^{k-1} \lambda(t)\, P_{ij}(\tau, t)\, q_{k-j} \quad .$$

This would correspond to the Kilomogorov Foreward Equations.[27] This combined with a knowledge of initial conditions form our model. The initial conditions take the form

$$P_{ij}(\tau, t) = \begin{cases} 1 \text{ if } j = i \\ 0 \text{ elsewhere} \end{cases} \quad . \tag{A8}$$

Our interest, then, is in obtaining solutions to Equations (A7) with initial conditions, Equation (A8). These functions $P_{ik}(\tau, t)$ would represent the form of the probability distribution of the number of organisms on the surface $\mathcal{S}$.

APPENDIX B

SOLUTIONS, $P_{ij}(\tau, t)$, AND THEIR PROPERTIES

Solutions, $P_{ij}(\tau, t)$, and Their Properties

Equations (A7) and (A8) of the previous appendix represent our basic model. The first question we should ask about this model concerns the existence of a solution to these equations. Feller[28, 29] has considered this problem for this type of equation. If our deposition and removal rates are bounded, then his results apply and we can conclude that there exists a unique solution to Equations (A7) and (A8).

The next question we should ask concerns the form of the solution to Equations (A7) and (A8). If we assume that the maximum number of microorganisms on $\mathcal{S}$ due to ψ is limited to some maximum, then the classical theory of ordinary differential equation allows us to write down the solution in closed form.[30] This form of the solution is not very practical since it requires the computation of the eigenvalues to certain matrices. Since the maximum must be chosen as a large number, these matrices are very large and, therefore, the computation of the eigenvalues is very difficult and impractical.

The method we shall use to derive the form of the solution to Equations (A7) and (A8) does not require us to choose a maximum. Define $Z(t)$ to be a random variable representing the number of particles of class ψ on $\mathcal{S}$ at time t. Let

$$f(n, t) = \text{Prob}\left\{ Z(t) = n \right\} .$$

Since $X(t)$ represents number of microorganisms on $\mathcal{S}$ at time t due to particles from class ψ, we know that

$$X(t) = \sum_{j=1}^{Z(t)} Y_j \tag{B1}$$

where Y_j, $j = 1, 2, \ldots$ are random variables which have the same distribution as Y. Then we know that under the assumptions we have made[31]

$$P_{ij}(\tau, t) = \sum_{\ell=0}^{\infty} \text{Prob}\left\{Z(t) = \ell\right\} \text{Prob}\left\{\ell \text{ particles contain j microorganisms}\right\}$$

$$= \sum_{\ell=0}^{\infty} f(\ell, t)\, Q\,(j, \ell) \tag{B2}$$

where

$$Q(j, \ell) = \text{Prob}\left\{\ell \text{ particles contain j microorganisms}\right\} .$$

We are now prepared to consider each of the terms in Equation (B2) individually. If we let

$$q_k = \begin{cases} 0 \text{ if } k \neq 1 \\ 1 \text{ if } k = 1 , \end{cases}$$

or if we reasoned again physically from first principles, then Equation (A7) would represent the number of <u>particles</u> in ψ on $\mathcal{S}$. These equations become

$$f'(0, t) = -\lambda(t)\, f\,(0, t) + \mu\,(t)\, f\,(1,\ t)$$

$$f'(n, t) = -\left(\lambda(t) + n\mu\,(t)\right) f\,(n, t) + \lambda(t)\, f\,(n - 1, t) \tag{B3}$$

$$+ (n + 1)\,\mu\,(t)\, f\,(n + 1, t),\ n \geq 1 \ .$$

We recognize these as the ordinary equations for a "Birth and Death" process. These could be derived in the same manner as we did Equation (A7). The initial conditions for this set of differential equations can be determined from Equation (A8). Observe that they must depend on i and τ even though this relationship is not explicitly exhibited in Equations (B2) or (B3). These were originally proposed for use in bioburden models in Reference 32. It is easy to find the mean of this distribution. Let

$$H(t) = \sum_{\ell=0}^{\infty} \ell\, f\,(\ell, t) \ .$$

Then in a straightforward manner we obtain

$$H'(t) = \lambda(t) - \mu(t)\, H(t) \quad . \tag{B4}$$

We are now prepared to state our first main result of the analysis of Equation (A7).

<u>Theorem 1</u>: If $H(0) = 0$ then

$$f(\ell, t) = \frac{\left(H(t)\right)^{\ell} e^{-H(t)}}{\ell!} \quad . \tag{B5}$$

<u>Proof</u>: This can be seen by a simple substitution of Equation (B5) into Equation (B4).

Therefore, if we consider only the case where $H(0) = 0$, we would have $i = 0$ at $\tau = 0$. This would imply that Equation (B2) represents a compound Poisson distribution.[22]

In analyzing $Q(j, \ell)$ we must make use of the fact that

$$Q(j, \ell) = \text{Prob}\left\{\sum_{m=1}^{\ell} Y_m = j\right\} \tag{B6}$$

and therefore since $\text{Prob}\left\{Y_m = p\right\} = q_p$ we have

$$Q(j, \ell) = \sum_{(\ell)} \frac{\ell!}{m_0!\, m_1! \ldots m_j!}\, q_0^{m_0} q_1^{m_1} \ldots q_j^{m_j} \tag{B7}$$

where the summation $\sum_{(\ell)}$ means over all sets of nonnegative integers $(m_0, m_1, m_2, \ldots, m_j)$ which satisfy the relations

$$m_0 + m_1 + m_2 + \ldots + m_j = \ell$$

$$m_1 + 2m_2 + \ldots + jm_j = j \quad .$$

We are now prepared to prove:

Theorem 2: Equation (B2) is a solution to Equation (A7).

Proof: Considering Equation (A3) we see that in order to find $E_j(t)$ we must know the probability of having ℓ particles given that there are m microbes at time t. Using the Bayes relation[31] we have

$$\begin{aligned}
&\text{Prob}\left\{\text{n particles in } \psi \text{ on } \mathscr{S} \text{ at time t given } X(t) = j\right\} \\
&\quad = \text{Prob}\left\{Z(t) = n \mid X(t) = j\right\} \\
&\quad = \frac{\text{Prob}\left\{\text{j microbes given n particles}\right\}\ \text{Prob}\left\{\text{n particles at time t}\right\}}{\sum_{\ell=0}^{\infty} \text{Prob}\left\{\ell \text{ particles at time t}\right\}\ \text{Prob}\left\{\text{j microbes given } \ell \text{ particles}\right.}
\end{aligned}$$

Thus, using the notation of this section, we have

$$E_j(t) = \frac{\sum_{n=0}^{\infty} n\, Q(j,n)\, f(n,t)}{\sum_{\ell=0}^{\infty} f(\ell,t)\, Q(j,\ell)} .$$

Rewriting this we have

$$E_j(t)\, P_{ij}(\tau, t) = \sum_{n=0}^{\infty} n\, Q(j,n)\, f(n,t) . \tag{B8}$$

We shall need the results of two lemmas in order to complete the proof of this theorem.

Lemma 1:

$$Q(k, \ell+1) = \sum_{j=0}^{k} Q(j,\ell)\, q_{k-j} . \tag{B9}$$

Proof: This follows immediately from our definitions since $Q(j, \ell)$ is the probability that the first ℓ particles chosen have j microorganisms attached while q_{k-j} represents the probability that the $\ell + 1$st particle has $k - j$ microorganisms. If we sum over all possible ways of dividing the particles between the first ℓ particles and the $\ell + 1$st particle we obtain Equation (B9).

Lemma 2:

$$Q(j, n) = \sum_{m=0}^{\infty} Q(m + j, n + 1)\, Q(m, 1) \quad . \tag{B10}$$

Proof: This again is an immediate consequence of our definitions.

The proof of our theorem is then completed by substituting Equation (B2) into Equation (A7) and making use of Lemmas 1 and 2 and Equation (B8).

The two things we shall need to know for our application of this model are the mean and the variance. Define

$$M(t) = \sum_{j=0}^{\infty} j\, P_{ij}(\tau, t) \quad . \tag{B11}$$

Again in order to simplify notation we omit the dependence of $M(t)$ on i and τ. We then have the following theorem.

Theorem 3: If $P_{ij}(\tau, t)$ is given by Equation (B2) then

$$M(t) = H(t)\gamma \quad , \tag{B12}$$

where

$$\gamma = \sum_{j=0}^{\infty} j q_j = \text{mean number of microbes per particle} . \tag{B13}$$

Proof: Let us consider the generating function for $P_{ij}(\tau, t)$ which is given by

$$G_x(z, t) = \sum_{j=0}^{\infty} P_{ij}(\tau, t)\, z^j, \qquad 0 < z \leq 1 \quad .$$

Making use of Equation (B2) we have

$$G_x(z, t) = \sum_{j=0}^{\infty} \left[\sum_{\ell=0}^{\infty} f(\ell, t)\, Q\,(j, \ell) \right] z^j$$

$$= \sum_{\ell=0}^{\infty} f(\ell, t) \sum_{j=0}^{\infty} Q\,(j, \ell)\, z^j$$

If we let

$$g(z) = \sum_{j=0}^{\infty} q_j z^j$$

then the generating function for $Q(j, \ell)$ is given by $[g(z)]^{\ell}$. Therefore

$$G_x(z, t) = \sum_{\ell=0}^{\infty} f\,(\ell, t)\, [g\,(z)]^{\ell} \quad . \tag{B14}$$

Let $G_z(z, t)$ be the generating function for the probability distribution $f\,(\ell, t)$. Then Equation (B14) implies that

$$G_x(z, t) = G_z\big(g\,(z), t\big) \quad .$$

Generating functions have the property that

$$M\,(t) = G_x'\,(1, t) \quad .$$

Using this together with the fact that $\sum_{j=0}^{\infty} q_j = 1$ we have

$$M(t) = \sum_{\ell=0}^{\infty} f(\ell, t)\, \ell g'(1) \quad .$$

But $\gamma = g'(1)$ and $H(t) = \sum_{\ell=\gamma}^{\infty} \ell f(\ell, t)$ so that we have our desired result.

We can obtain an expression for the variance in the same manner. Let

$$V(t) = \mathrm{Var}\left(Z(t)\right)$$

$$\mathcal{E}(t) = \mathrm{Var}\left(X(t)\right)$$

and

$$v = \mathrm{Var}(Y) \quad .$$

Using the definition of variance we see that $V(t)$ must satisfy the equation

$$\left[V(t) + H^2(t)\right]' = 2\lambda(t) H(t) + \lambda(t) - 2\mu(t)\left[V(t) + H^2(t)\right] + \mu(t) H(t) \quad . \qquad \text{(B16)}$$

<u>Theorem 4</u>: If $P_{ij}(\tau, t)$ is given by Equation (B2) we have

$$\mathcal{E}(t) = M(t)\, v + \gamma^2 V(t) \quad . \qquad \text{(B17)}$$

In particular if $i = 0$ when $\tau = 0$ then

$$\mathcal{E}(t) = H(t)\, v + \gamma^2 H(t) \quad . \qquad \text{(B18)}$$

Proof: We know that

$$\mathcal{E}(t) = G_x''(1,t) + G_x'(1,t) - \left[G_x'(1,t)\right]^2 ,$$

A simple substitution of Equation (B15) into this expression yields Equation (B17). Equation (B18) follows from Equation (B5).

At this point, let us stop to verify the "plateau" phenomena and to make some remarks regarding the entire spacecraft.

If there are enough particles of ψ on $\mathcal{S}$ such that the distribution of the number of microorganisms per particle in ψ is the same on $\mathcal{S}$ as it is in the environment then the stochastic process with which we are working is a Markov process. From physical grounds we also know that it is irreducible if we limit our consideration to only a finite number of microorganisms on the surface.[33] If the environment in which we are working leads to a process which is homogeneous then

$$\lim_{t\to\infty} P_{ij}(\tau, t) = \pi_j$$

and π_j is independent of i and τ. This is one possible mathematical interpretation which could be given to the plateau phenomena.

Another way to consider the "plateau" effect is to only consider it as implying the existance of a bound on the mean of the distribution.[34] Observe that Equation (B4) implies that if M (0) = 0 then

$$H(t) = e^{\int_0^t \mu(t)\,dt}\left[\int_0^t \lambda(t)\, e^{\int_0^t \mu(t)\,dt}\, dt\right] . \tag{B19}$$

If we know that $0 \le \lambda(t) \le \lambda_{max}$ and that $0 < \mu_{min} \le \mu(t) \le 1$, then we have from Equation (B19) that

$$H(t) \leq \frac{\lambda_{max}}{\mu_{min}} .$$

Thus would then imply

$$M(t) \leq \frac{\lambda_{max}}{\mu_{min}} \gamma .$$

Let us again assume $M(0) = 0$. If we wish to find the probability distribution for the entire spacecraft, it can be obtained from the probability distribution for each zone. Since our probability distribution at any time t is a compound Poisson distribution, we may, therefore, conclude that the distribution for the entire spacecraft is a compound Poisson distribution and that the mean is obtained by summing the means for the various subsections.

APPENDIX C

THE MODEL AND ITS IMPLEMENTATION

The Model and Its Implementation

If we assume that H (0) = 0 (i.e., the number of particles of ψ on $\mathcal{S}$ at time zero is known to be zero) then combining Equations (B2) and (B5) we have

$$P_{0j}(0, t) = \sum_{\ell=0}^{\infty} \frac{e^{-H(t)}\left(H(t)\right)^{\ell}}{\ell!} Q(j, \ell) \tag{C1}$$

where H (t) is determined by Equation (B4) with the initial condition H (0) = 0.

Equation (C1) would then represent the estimation and prediction model as it was presented early in this report [Equation (1)]. It is interesting to observe that if the distribution of the number of microorganisms per particle (or clump) in ψ is Poisson then Equation (C1) represents a Neyman distribution of Type A. This distribution was the estimation model used to determine the sampling requirements for Apollo.[23] If this distribution is not Poisson then the estimation model which is compatible with this prediction model is a generalization of that used in Reference 23. As an estimation model, the use of Equation (C1) seems clear when H (t) is replaced by M (t)/γ. Its use as a prediction model is the subject of the remainder of this section. Let us assume that the form of the distribution of the number of microorganisms per particle in ψ is known and that we know γ and v.

If we know λ (t) and μ (t), then, after using Equation (B4) to get H (t), we can calculate the variance and mean of the distribution Equation (C1) by using Equations (B12) and (B18). The problem arises that we, in most cases, do not know λ (t) and μ (t). Our problem is to obtain values for λ (t) and μ (t) from samples which are taken from $\mathcal{S}$ or from the environment surrounding $\mathcal{S}$. Since surface samples are the most direct approach to the estimation of bioburdens and since the state of the art is not far enough advanced to relate environmental contamination to surface contamination, we shall in this section discuss how values for λ (t) and μ (t) can be obtained from surface samples. The technique might also give us a way of learning

more about the relationships which exist between surface contamination and environmental contamination. We shall discuss this more fully later in the section.

Let us assume that we sample surface $\mathcal{S}$ for microorganisms associated with particles in ψ at time $t_1 < t_2 < t_3 < \ldots < t_m$ and we wish to extrapolate to some time t_f. Suppose that at time t_k, $k = 1, \ldots, m$ we take n_k samples of α square inches each with a uniform sampling efficiency ϵ_k. This assumption of a uniform sampling error is realistic because of our definition of $\mathcal{S}$. We have required $\mathcal{S}$ to be uniform in most characteristics which would cause the distribution of the sampling error to take another form. Let $x_i(t_k)$ be the number of microorganisms found on $\mathcal{S}$ due to class ψ in the ith sample at time t_k. Since the area of $\mathcal{S}$ is A square inches, then the expression

$$\bar{x}(t_k) = \frac{A}{\epsilon_k n_k \alpha} \sum_{i=1}^{n_k} x_i(t_k)$$

provides us with an unbiased consistent estimator to the mean M (t) of the distribution Equation (C1) at time t_k. Thus, we know that $\bar{x}(t_k)$ is an approximation to $M(t_k)$.

From Equations (B4) and (B12) we know that

$$M(t) = e^{-\int_0^t \mu(t)\,dt} \left[\int_0^t \lambda(t)\, e^{\int_0^t \mu(t)\,dt}\, dt \right]^{\gamma} \tag{C2}$$

where γ is assumed known.

Let us assume that we know that $\lambda(t)$ and $\mu(t)$ lie in a certain class of functions which can be spanned using a finite number of real variables; that is, let us assume

$$\lambda(t) \in C_\lambda\left(\beta_1, \beta_2, \ldots, \beta_r\right)$$

and

$$\mu(t) \in C_\mu \left(\alpha_1, \alpha_2, \ldots, \alpha_p\right) .$$

This implies that as we find $\beta_1, \beta_2, \ldots, \beta_r, \alpha_1, \ldots,$ and α_p then we can completely describe $\lambda(t)$ and $\mu(t)$. The class functions C_λ and C_μ which we choose will be based on our physical knowledge of what deposition and removal are like. Since $\bar{x}(t_k)$ approximates $M(t_k)$, we know that

$$\bar{x}(t_k) = M(t_k) + e(t_k), \quad k = 1, 2, \ldots, m$$

where $M(t_k)$ is determined by Equation (C2) and $e(t_k)$ is a random variable representing the error between the model and the observation.

Consider the quantity

$$R = \sum_{i=1}^{m} \left(\bar{x}(t_i) - M(t_i)\right)^2 = R(\mu, \lambda) . \tag{C3}$$

If we know the true $\lambda(t)$ and $\mu(t)$ lie in C_λ and C_μ respectively then the Gauss-Markoff Theorem tells us that the $\hat{\lambda}(t) \in C_\lambda$ and $\hat{\mu}(t) \in C_\mu$ which minimize R are the best unbiased estimates to $\lambda(t)$ and $\mu(t)$.[35] This implies that if $\lambda(t)$ is determined by $\hat{\beta}_1$, $i = 1, 2, \ldots, r$ and $\hat{\mu}(t)$ is determined by $\hat{\alpha}_1$, $i = 1, 2, \ldots, p$ then

$$\text{Expected value}\left(\hat{\beta}_i\right) = \text{true } \beta_i, \quad i = 1, 2, \ldots, r$$

and

$$\text{Expected value}\left(\hat{\alpha}_i\right) = \text{true } \alpha_i, \quad i = 1, 2, \ldots, p .$$

Also, if we let

$$\hat{R} = R(\hat{\lambda}, \hat{\mu})$$

then the expected value of $e(t_k)$ is zero for $k = 1, 2, \ldots, m$ and

$$E = \text{Expected value } \frac{\hat{R}}{m - r - p} = \text{Var}\left(e(t_k)\right) .$$

Thus, the quantity $\frac{R}{m - r - p}$ is a measure of how well the model fits the data.

The problem remaining is to decide how to choose $\lambda(t)$ and $\mu(t)$ for the period from $t_m < t \leq t_f$. It is clear that $\hat{\lambda}(t)$ and $\hat{\mu}(t)$ represent the deposition and removal of particles in ψ from $\mathscr{A}$ during the period $t_1 \leq t \leq t_m$. If we know, or are willing to assume that the environment $\mathscr{A}$ sees in $\left(t_m, t_f\right]$ is the same as it sees in $\left[T_1, T_2\right] \subseteq \left[t_1, t_m\right]$ then we can use $\lambda(\theta t + \eta)$ and $\mu(\theta t + \eta)$ as the deposition and removal rates where

$$\theta = \frac{T_2 - T_1}{t_f - t_m}$$

and

$$\eta = T_2 - \theta t_f .$$

This will be very useful when various environmental parameters agree in a qualitative way.

The other, and more desirable way, of determining $\lambda(t)$ and $\mu(t)$ for use in the time period $\left(t_m, t_f\right]$ is by realizing that $\hat{\lambda}(t)$ and $\hat{\mu}(t)$ should be capable of determination as functions of various environmental measurements taken during $t_1 \leq t \leq t_m$. Using data fitting techniques $\hat{\lambda}(t)$ and $\hat{\mu}(t)$ have been determined from minimizing Equation (C2). We can then determine what function they are of the environmental measurement and then use these functions to extrapolate to the environment which $\mathscr{A}$ will see during the period $t_m < t \leq t_f$. This will be the subject of a future report.

REFERENCES

1. "Outbound Planetary Biological Contamination Control: Policy and Responsibility," Policy Directive NPD 8020.10, National Aeronautics and Space Administration, September 6, 1967.

2. Rahn, Otto, "The Problem of the Logarithmic Order of Death in Bacteria: A Critical Discussion," Biodynamica 4 (1943), pp. 81-128.

3. Silverman, G. J. and Sinskey, T. J., "The Destruction of Microorganisms by Ionizing Irradiation," Disinfection, Sterilization and Preservation, edited by C. A. Lawrence and S. S. Block, Lea and Febiger, Philadelphia, 1968.

4. Dugan, V. L., A Mathematical Model for the Thermoradiation Inactivation of Dry Bacillus Subtilis Var. Niger Spores, SC-RR-70-203, Sandia Laboratories, Albuquerque, New Mexico, April 1970.

5. Powers, E. M., Microbial Profile of Laminar Flow Clean Rooms, X-600-65-308, Goddard Space Flight Center, Greenbelt, Maryland, September 1965.

6. Michaelson, G. S., Ruschmeyers, O. R., and Vesley, D., The Bacteriology of Clean Rooms, NASA-CR-890, National Aeronautics and Space Administration, October 1967.

7. Corn, W. B. and Kethley, T. W., "Dispersion of Airborne Bacteria in Clean Rooms," presented at the Fifth Annual Technical Meeting and Exhibit of the Americal Association for Contamination Control, Houston, Texas, March 29-April 1, 1966.

8. McDade, J. J., Determination of the Microbiological Profile of Clean Rooms, SM-49062, Douglas Aircraft, October 1965.

9. Viking Project: Planetary Quarantine Provisions, M73-109-0, National Aeronautics and Space Administration, February 20, 1969.

10. Angelotti, R., et al, "Review of the JPL-Martin Report on a Microbial Burden Prediction Model," Report to the Planetary Quarantine Advisory Committee, April 1969.

11. McDade, J. J., et al, "Environmental Microbiology and the Control of Microbial Contamination," Spacecraft Sterilization Technology, SP-108, National Aeronautics and Space Administration, 1966.

12. McDade, J. J., et al, "Techniques for the Limitation of Biological Loading of Spacecraft Before Sterilization," Sterilization Techniques for Instruments and Materials as Applied to Space Research, edited by P. H. A. Sneath, COSPAR Technique Manual Series, Manual No. 4, 1968.

13. Corn, M., "The Adhesion of Solid Particles to Solid Surfaces, I. A. Review," J. Air. Poll. Control Assoc., 11 (1961), pp. 523-529.

14. Corn, M., "Adhesion of Particles," Aerosol Sciences, edited by C. N. Davies, Academic Press, New York, 1966.

15. Jordan, D. W., "The Adhesion of Dust Particles," Brit. J. Appl. Phys. Suppl. 3 (1954), pp. 5194-5197.

16. Corn, M. and Stein, F., "Mechanisms of Dust Redispersion," Surface Contamination, edited by B. R. Fish, Pergaman Press, London, 1967.

17. Letter from Normal J. Petersen to C. A. Trauth, Jr., dated September 13, 1968.

18. Fredrickson, A. G., "Stochastic Models for Sterilization," Biotechnol. Biolog. 8 (1966), pp. 167-182.

19. NASA Standard Procedures for the Microbiological Examination of Space Hardware, NHB-5340.1, National Aeronautics and Space Administration, August 1967.

20. Fooks, J. H., "An Approach to the Estimation of Microbial Contamination on Spacecraft," Proceedings of the Sixth Annual Meeting of the American Association for Contamination Control, Washington, D.C., May 1967.

21. Feller, W., "On a General Class of 'Contagious' Distributions," Annals. of Math. Stat. 14 (1943), pp. 389-400.

22. Haight, F. A., Handbook of the Poisson Distribution, John Wiley, New York.

23. Roark, A. L., The Determination of Quantitative Microbial Sampling Requirements for Apollo Modules, SC-RR-69-23, Sandia Laboratories, Albuquerque, New Mexico, 1969.

24. Planetary Quarantine Department 1740, Sandia Laboratories Quarterly Report - Planetary Quarantine Program, QR14, September 1969.

25. Green, H. L. and Lorne, W. R., Particulate Clouds: Dust, Smokes and Mists, D. Van Nostrand Company, Inc., Princeton, New Jersey, 1964.

26. Planetary Quarantine Department 1740, Sandia Laboratories Quarterly Report - Planetary Quarantine Program, QR17, July 1970.

27. Parzen, E., Stochastic Processes, Holden-Day, San Francisco, 1962.

28. Feller, W., "On the Integro-Differential Equations of Purley Discontinuous Markoff Processes," Vol. 48 (1940), Trans. Amer. Math. Soc., pp. 488-515.

29. Feller, W., "On Boundaries and Lateral Conditions for the Kolmogorov Differential Equations," Vol. 65 (1957), Ann. Math., pp. 527-570.

30. Coddington, E. A. and Levinson, N., Theory of Ordinary Differential Equations, McGraw-Hill, New York, 1955.

31. Feller, W., An Introduction to Probability Theory and Its Applications, Volume 1, John Wiley, New York, 1968.

32. Tierney, M. S., The Chances of Retrieval of Visible Microorganisms Deposited on the Moon by Unmanned Lunar Probes, SC-M-68-539, Sandia Laboratories Report, August 1968.

33. Breiman, L., Probability and Stochastic Processes with a View Toward Applications, Houghton Mifflin Co., Boston, Mass., 1969.

34. Personal Communication, C. A. Trauth, Jr., 1970.

35. Hemmerde, W. J., Statistical Computation on a Digital Computer, Blaisdell Publishing Company, Waltham, Mass., 1967.

DISTRIBUTION:

NASA, Code SC
Grants and Contracts
400 Maryland Avenue, SW
Washington, D. C. 20546 (25)

L. B. Hall, NASA
Code SB
400 Maryland Avenue, SW
Washington, D. C. 20546 (25)

B. W. Colston
Director, Space & Special Programs
Division
Office of Operations
U. S. Atomic Energy Commission
Albuquerque, New Mexico 87115

L. P. Daspit, Jr.
Viking Project Quarantine Officer
Viking Project Office, NASA
Langley Research Center
Hampton, Virginia 23365

University of California, LRL
P. O. Box 808
Livermore, California 94551
Attn: Tech. Info. Div.
For: Report Librarian

Los Alamos Scientific Laboratory
P. O. Box 1663
Los Alamos, New Mexico
Attn: Report Librarian

Richard G. Bond
School of Public Health
College of Medical Science
University of Minnesota
Minneapolis, Minnesota 55455

John H. Brewer
Star Route 2
Brownwood, Texas 76801

Harold Walker
Director of Research Services
Graduate College
University of New Mexico
Albuquerque, New Mexico

Frank B. Engley, Jr.
Chairman, Department of Microbiology
School of Medicine
University of Missouri
Columbia, Missouri

Gilbert V. Levin
Biospherics, Inc.
4928 Wyaconda Rd.
Rockville, Maryland 20853

Irving J. Pflug
Professor of Environmental Health
University of Minnesota
College of Medical Sciences
Minneapolis, Minnesota 59455

Gerald Silverman
Department of Nutrition and Food Science
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

John A. Ulrich
School of Medicine
University of New Mexico
Albuquerque, New Mexico

Samual Schalkowsky
Exotech Incorporated
525 School Street, SW
Washington, D. C. 20024

Boris Mandrovsky
Aerospace Technology Division
Library of Congress
Washington, D. C.

Mark A. Chatigny
Research Engineer
Naval Biological Laboratory
Naval Supply Center
University of California, Berkeley
Oakland, California 94625

Richard G. Cornell
Associate Professor of Statistics
Department of Statistics
Florida State University
Tallahassee, Florida

DISTRIBUTION (cont.):

Martin S. Favero
Department of Health, Education
and Welfare
CDC-Phoenix Field Station
4402 North 7th Street
Phoenix, Arizona 85014

F. N. LeDoux
Head, Structural & Mechanical
Applications Section
Goddard Space Flight Center
Greenbelt, Maryland

Q. Ussery
Code NC3, Quality Assurance Branch
Manned Spacecraft Center, NASA
Houston, Texas

F. J. Beyerle
George C. Marshall Space Flight Center
Manufacturing Engineering Laboratory
Code R-ME-MMC
Huntsville, Alabama 35812

J. Gayle
Code SO-PLN, Rm 2123, HQS. Bldg.
Kennedy Space Center, NASA
Cape Canaveral, Florida

Murray Schulman
Division of Biology and Medicine
Headquarters, AEC
Washington, D. C. 20545

N. H. MacLeod
Space Biology Branch
Code 625, Bldg. 21, Rm 161
Goddard Space Flight Center
Greenbelt, Maryland 20771

J. E. Campbell
U. S. Public Health Service
222 E. Central Parkway
Cincinnati, Ohio 45202

G. Rotariu
Process Radiation Staff
Division of Isotopes Development
Headquarters, AEC
Washington, D. C. 20545

Martin G. Koesterer, Microbiologist
Bioscience Operation
General Electric
P. O. Box 8555
Philadelphia, Pennsylvania 19101

Carl Bruch
Chief, Bacteriology Branch
Division of Microbiology
Food and Drug Administration
3rd and C., SW, Rm 3876
Washington, D. C. 20204

John W. Beakley
Department of Biology
University of New Mexico
Albuquerque, New Mexico

Loren D. Potter, Chairman
Department of Biology
University of New Mexico
Albuquerque, New Mexico

Loris W. Hughes
Department of Biology
New Mexico State University
University Park, New Mexico

Richard W. Porter
Corporate Engineering Staff
General Electric Company
570 Lexington Avenue
New York, New York

Fred L. Whipple
Smithsonian Astrophysical Observatory
Cambridge, Mass. 02138

J. J. McDade
Biohazards Group
Pitman-Moore Company
Dow Chemical Company
P. O. Box 10
Zionsville, Indiana 46077

Otto Hamberg
Aerospace Corporation
Building A2, Rm 2019
2350 East El Segundo Blvd.
El Segundo, California

DISTRIBUTION (cont.):

Lawrence P. Chambers
NASA Headquarters
Office of Manned Space Flight
Code MLR
Washington, D. C. 20546

Arthur H. Neill
Code SB
400 Maryland Avenue, SW
Washington, D. C. 20546

Richard Green
Jet Propulsion Laboratory
4800 Oak Grove Dr.
Pasadena, California 91103

Rudy Puleo
Public Health Service
Spacecraft Bioassay Laboratory
Drawer Y
Cape Canaveral, Florida 32900

USAEC, Division of Technical
Information
P. O. Box 62
Oak Ridge, Tennessee 37830
Attn: Reference Branch
P. E. Postell

Carl Sagan
Cornell University
Center for Radiophysics and Space
Research
Space Science Building
Ithaca, New York 14850

Document Library
Lovelace Foundation for Medical
Education and Research
5200 Gibson Blvd., SE
Albuquerque, New Mexico 87108

Martin S. Tierney
Group J-10
Los Alamos Scientific Laboratory
Los Alamos, New Mexico

E. C. Pollard
Professor of Biophysics
Pennsylvania State University
618 Life Sciences Building
University Park, Pennsylvania 16802

Robert Angelotti
Deputy Director
Division of Microbiology
Food and Drug Administration
Health, Education and Welfare
200 C. Street, SW
Washington, D. C. 20546

Vance I. Oyama, Chief
Life Detection Systems Branch
NASA, Ames Research Center
Moffett Field, California 94035

Byron W. Brown, Jr.
Department of Community and
Preventive Medicine
Stanford University School of Medicine
Stanford University Medical Center
Stanford, California 94305

Don G. Fox
Sterility Control Officer
NASA Headquarters, Code SB
400 Maryland Avenue, SW
Washington, D. C. 20546

A. A. Rothstein
Manager, Planetary Quarantine
Martin Marietta Corporation
Mail No. 8401
Denver, Colorado 80201

Hellel S. Levinson
U. S. Army Natick Laboratory
Natick, Massachusetts

A. Anellis
U. S. Army Natick Laboratory
Natick, Massachusetts

B. S. Schweigert, Chairman
Department of Food Science
College of Agriculture
Michigan State University
East Lansing, Michigan 48823

H. O. Halvorson
Biochemistry Department
St. Paul Campus
University of Minnesota
St. Paul, Minnesota

DISTRIBUTION (cont.):

Jack Kaye
11607 Georgetowne Court
Patomic, Maryland 20854

H. W. Johnson, LTC
U.S. Army Medical Research and
Development Command
Washington, D.C. 20314

Donald A. Kautter
Dept. of HEW
Food and Drug Administration
Div. of Microbiology
BF-135
200 C Street S.W.
Washington, D.C. 20204

Mr. James Martin
Viking Project Engineer
Langley Research Center, NASA
Langley Station
Hampton, Virginia 23365

J. A. Hornbeck, 1
Staff, 100
W. J. Howard, 1000
D. B. Shuster, 1200
W. A. Gardner, 1500
H. E. Lenander, 1600
T. M. Burford, 1700
C. Winter, 1710
D. R. Morrison, 1720
J. W. Worrell, Jr., 1721
D. P. Peterson, 1724
R. G. Clem, 1730
H. D. Sivinski, 1740 (25)
R. W. Henderson, 2000
C. B. McCampbell, 2310
B. H. VanDomelen, 2345
S. J. Buchsbaum, 5000
L. C. Hebel, 5200
A. W. Snyder, 5220
R. M. Jefferson, 5221
J. E. McDonald, 5300
L. M. Berry, 5500
D. W. Ballard, 7361
G. A. Fowler, 9000
J. H. Scott, 9200
A. Y. Pope, 9300
L. E. Hopkins, Jr., 9500
R. S. Gillespie, 3411
J. L. Gardner, 3422
L. S. Ostrander, 8232
W. K. Cox, 3422-1 (15)