A Proposed Method to Determine Spatio-Temporal Dynamics of Bacterial Growth and Relative Efficacy of Treatments

Terrence P McGarty
White Paper No 139
September, 2016

We examine a set of metrics for ascertaining if two growth media have differing rates of bacterial growth. We propose a modified non-parametric approach which adds significant additional information while being non-destructive.

Notice

This document represents the personal opinion of the author and is not meant to be in any way the offering of medical advice or otherwise. It represents solely an analysis by the author of certain data which is generally available. The author furthermore makes no representations that the data available in the referenced papers is free from error. The Author also does not represent in any manner or fashion that the documents and information contained herein can be used other than for expressing the opinions of the Author. Any use made and actions resulting directly or otherwise from any of the documents, information, analyses, or data or otherwise is the sole responsibility of the user and The Author expressly takes no liability for any direct or indirect losses, harm, damage or otherwise resulting from the use or reliance upon any of the Author's opinions as herein expressed. There is no representation by The Author, express or otherwise, that the materials contained herein are investment advice, business advice, legal advice, medical advice or in any way should be relied upon by anyone for any purpose. The Author does not provide any financial, investment, medical, legal or similar advice in this document or in its publications on any related Internet sites.

Contents

1 INTRODUCTION

The use of a bactericidal solution along with a bacteriostatic nano coating results in first the drastic reduction of bacterial growth and then in the inhibition of regrowth. We consider the following experiment.

Using a specific pathogen, apply it to a growth medium and then swipe the medium with a bactericidal medium and then a bactericidal and NSe medium. Then watch regrowth over time. Try to ascertain any added efficacy by the addition of the NSe.

1.1 OBJECTIVE

The objective here is to take the steps in the above general description and then develop a metric to measure efficacy. Specifically:

- The problem is to determine the efficacy in terms of reducing bacterial regrowth
- The test is to use coated petri dishes and then swipe with a swipe with and w/o NSe.
- The tests would then consider several NSe concentrations, 0 to some max, and measure regrowth from 0 to say 72 hours
- The technique uses the regrowth profile and uses Fourier Transform analysis and power spectra measurements vs time.

Why use a Fourier Transform, FT, approach? The simple reason is that the regrowth is stochastic in space. The FT is reflective of some of those features. We will explain this in detail later.

However, we also present a detailed methodology using least square nonlinear estimators yielding LMSE of the rate as well as the variance of that estimate. However, it depends on a complex execution of the Kushner-Stratonovich equation which I have linearized. The problem may be:

"If all you have is a hammer then everything looks like a nail!"

Specifically, the problem here is to ascertain if regrowth with NSe is less than that without any NSe. One step is to estimate the respective growth rates and then compare them. The problem than is; how good is the estimate and if a hypothesis test of NSe superiority is made based upon that data, how efficient is that test? We examine several options herein. We show how complex this may be. We also propose a non-parametric tests that examines spatial as well as temporal characteristic.

What do we mean by spatio-temporal analysis? Simply we mean that we believe that a better discriminant should include not just the temporal factors but the spatial as well. Thus the metric we propose is one which incorporates both.

We start with an examination of the proposed metric and then return to more classic approaches.

We can provide a brief summary. Let us assume that we have a bacteria process with growth that has a growth rate of say λ_1 and if treated with some process it is λ_2.

1.2 MEASUREMENTS

We will not consider more complex measurements for a variety of reasons. For classic growth rates one can use spectrophotometer methods applying Beer's law for liquid samples over time. This is a destructive form and does not allow measurements of NSe. Also flow cytometry methods are similarly ineffective. Tagging RNA or other markers on bacteria and then counting via some equivalent of a spectrum counting mechanism is costly and time consuming. We thus propose a simpler approach albeit we argue equally valid.

1.3 HYPOTHESIS TESTS

Now we want to determine if the growth rate is different, or even if it the same. We then need a simple hypothesis test to determine of H_1 is true or not. Namely:

H_1 means $\lambda_1 > \lambda_2$ and H_0 otherwise. Then we would also like to determine how confident we are in that proposition.

There is a multiplicity of ways of determining this. For example:

1. We can determine the probability density of the counts depending on the growth rates and use a Neyman Pearson test. That is fairly robust and efficient. We can determine these probabilities as we describe herein but they are complex and do not seem to yield useful results. However, they do yield some insight to the growth complexity.

2. Assume that the growth rates are random processes as are the growth themselves and further assume that the measurement is noisy. We can then apply linearized Kushner-Stratonovich equations. The result is the least squares estimate of the birth rate and a calculation of the variance and covariances. Then we could assume a Gaussian profile and use that to determine the performance of the hypothesis test. This is very robust but requires great data gathering and calculations.

3. We can calculate a monotonic measure that is such that the measure M has the following characteristic:

$$M(\lambda_1) > M(\lambda_2) \forall \lambda_1 > \lambda_2$$

Then if we choose M properly we can say that the hypothesis is met. Yet the performance issue still remains. We demonstrate herein how to deal with that as well.

2 EXPERIMENTAL PROCEDURE

The overall objective is to determine the bacteriostatic capability of the inclusion of nano Selenium on surfaces after having been treated by a wide variety of bactericidal treatments. The reinfection is something which is a result of the existing in situ environment and does not require a new treatment of a pathogen.

The experimental procedures are as follows:

Take a petri dish with a growth medium.

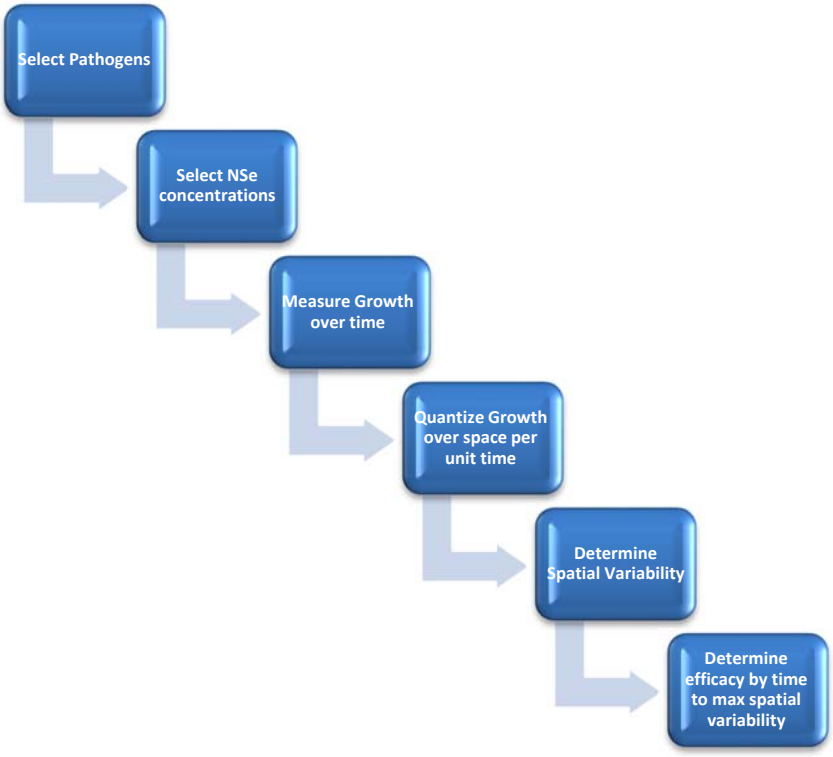
Coat the dish with a selected pathogen, S. aureus as an example, then allow some growth.

Then wipe a section of the disk with the solution and the solution plus a bacteriostatic nano surfacing adjunct. Do so with an X like swab. The center of the X like swab should be at least 40 mm.

Then examine the center of the X swab for the bactericidal effects. Now over the following periods, 4, 8, 12, 24, 36, 48, 72 hours, examine the X swab by photographing along with it with a ruler measuring distances in mm. The width of the center should be at least 40 mm in maximum width cleared by the bactericidal solution.

Collect photographs of the regrowth during these periods. They MUST be done in a timely manner and with highest possible digital resolution. Also the tests should be done for a single pathogen but for varying concentrations of nano inhibitors. This is shown below:

Overall the process appears as shown below:



The nano inhibitors should be such that the coverage is estimated at 50, 100 and 150 ng/cm^2.

Now classic means to measure bacterial growth often a destructive of surface coatings, namely analysis using spectrophometers at 600 nm[1]. Other means use tagged RNA which again is not useful, and complex counting systems. Our proposed system of measurement of regrowth is simple:

1. Leave the surface unaffected. It is essential to determine efficacy that the nano treated surface be left unchanged.

2. Then measure regrowth in situ in as simple a manner as possible.

3. Select a measurement technique which reflects both the temporal and spatial regrowth. Generally, most regrowth reflects only temporal by measuring gross reinfection.

2.1 SURFACE COATINGS

As we have noted elsewhere:

In the above we note that the regrowth of bacteria is dramatically reduced by the application of a Se surface at concentrations of 20-40 micrograms/milliliter. It has also been observed that

[1] See Prescott et al pp 116-125

surface coatings of a density of 100 ng/sq. cm of 20-70 nm diameter nano Se on the surface are also adequate.

Thus we seek roughly 100 ng/sq. cm surface coatings.

Surface concentration Required (ng/sq. cm)	100	50	150
Density of Se			
gray: g/cm3	4.81	4.81	4.81
alpha: g/cm3	4.39	4.39	4.39
vitreous: g/cm3	4.28	4.28	4.28
Density Average in ng/cm3	4,493,333,333	4,493,333,333	4,493,333,333
Density Average in ng/nm3	0.0000000000045	0.0000000000045	0.0000000000045
Size of Nano Particle nm diameter	30	30	30
Volume nano Se nm3	14,137	14,137	14,137
Area nano Se nm2	707	707	707
Weight per nano Se particle ng	0.00000006	0.00000006	0.00000006
Number Se nano particles per sq. cm	1,571,896,997	785,948,498	2,357,845,495

Thus we would compute the preparation as follows:

1. If we want 100 ng/cm^2 then we can determine how many gm we need in a solution which saturates a wipe.

2. Take a dry wipe and weigh it.

3. Take the solution and saturate the wipe. Weigh the total and determine the weight of the solution. Knowing the density of the solution we can determine the volume used to saturate.

4. Estimate the transfer percent from the wipe to the surface. Generally, we have seen that it can be 50%. Thus the solution is half transferred to the surface.

5. Determine the surface area for this determined transfer percentage. Namely we can get a transfer effectiveness of X% per cm^2 and this then allows us to ascertain the coating amount and/or the dosing amount.

6. If then we want 100 ng/cm^2 and we have a 50% transfer rate in one sq. m of surface then we know that we want 2X100X100X100 ng of NSe or 2 mg of NSe dissolved in the solution.

7. This principle can thus be carried forth for any other similar range of coverage.

2.2 SPECTRAL CHARACTERISTICS

The regrowth of a pathogen is common. It is a spatially stochastic process. Namely centers of regrowth appear and they in turn expand. Regrowth is spotty as shown below:

Note the spotty regrowth in the area where the bacteria were originally killed. Now we seek a reliable metric to ascertain the efficacy of the alternative approaches. Time will result in regrowth and it will be a spatially distributed random field in character. Techniques have been developed to analyze such phenomenon and they rely on spectral density calculations. Namely, one can calculate the spatial randomness of the regrowth by determining the Fourier Transform, then its power spectrum, and finally the total power as a function of time.

The FT will provide insight into the spatial characteristics. Specifically:

Assume the data can be quantized in a 0.1 manner, namely no growth or growth. Then let L be the total length of the sample, say 20 mm. Let us assume we than sample the sample each 0.5 mm for a total of N = 40 samples.

Let:

$$s(x_i) = s(i) = \begin{cases} 0 \text{ if no growth} \\ 1 \text{ if growth} \end{cases}$$

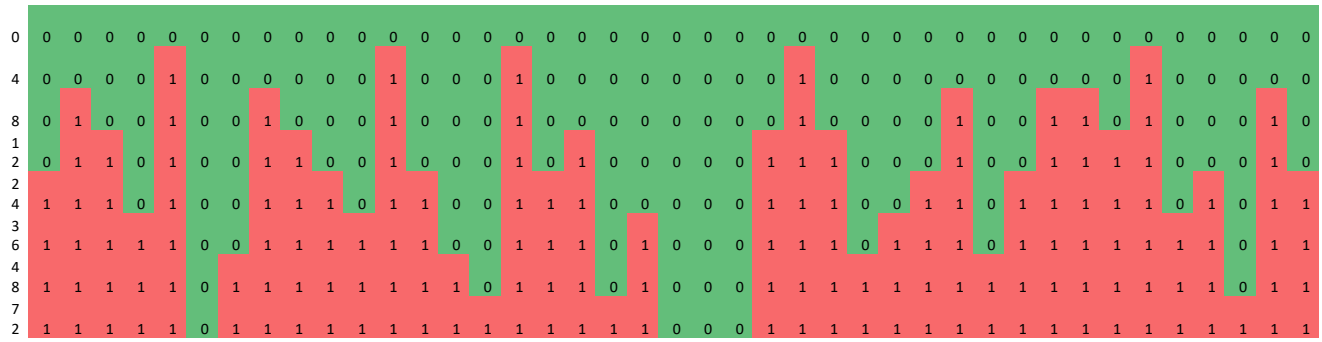
$$i = 1, N+1$$

$$N = \frac{L}{d}$$

$L =$ sample width in total

d= length of individual samples

We show an example below. Here we show the 41 cells horizontally and red is growth and green is no growth. The cells going down vertically are at various time intervals as noted. Our objective is to obtain the Fourier Power spectrum of these cells as a function of time.



The spectrum for a specific wave number is defined as:

$$S(f) = \int_0^L s(x)\exp(-j2\pi fx)dx$$

$$= \sum_{n=1}^{N+1} s(n)\exp(-j2\pi fn\delta) = \sum_{n=1}^{N+1} s(n)\cos(2\pi fn\delta) + j\sum_{n=1}^{N+1} s(n)\sin(2\pi fn\delta)$$

and the power spectrum given by

$$G(f) = S(f)S^*(f)$$

or

$$G(f) = \left[\sum_{n=1}^{N+1} s(n)\cos(2\pi fn\delta)\right]^2 + \left[\sum_{n=1}^{N+1} s(n)\sin(2\pi fn\delta)\right]^2$$

Thus for each f, where there are 2 time the f values as there are x values, we must calculate these products. It then yields a spectrum as a function of the f wave number. Then we can integrate this to yields a power value for each time unit. Namely we have:

$$H(t : sample\ \mathrm{K}) = \int_0^{f_{\max}} G(f, sample\ \mathrm{K})df$$

Now one more simple point. For each of the samples we must subtract out the average so that the samples we use have zero average. That is when we perform the above we define:

$$s(t) = s(n) = s_{measured}(t) - s_{avg}$$

where

$$s_{avg} = \sum_{n=1}^{N+1} s_{measured}(n)$$

This normalization allows for the elimination of spectral bias.

Now what we can then do is that for each sample K, where that means a selenium concentration level, we plot H(t) versus t.

What does this plot mean? Simply:

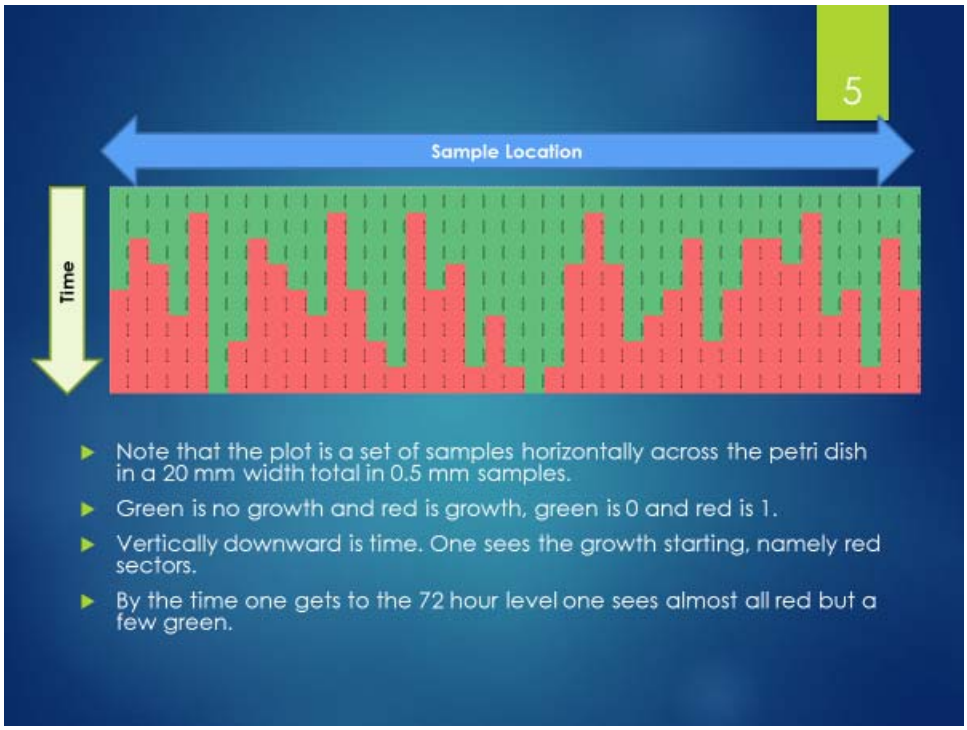
1. For a t where all samples are zeros, namely the same across all spatial cells, we have zero H(t). That is total continuity.

2. For a t where there is very spotty growth, we have "high" spatial frequency components and when we calculate H(t) for that t value we get a large number reflecting the multiplicity of growth.

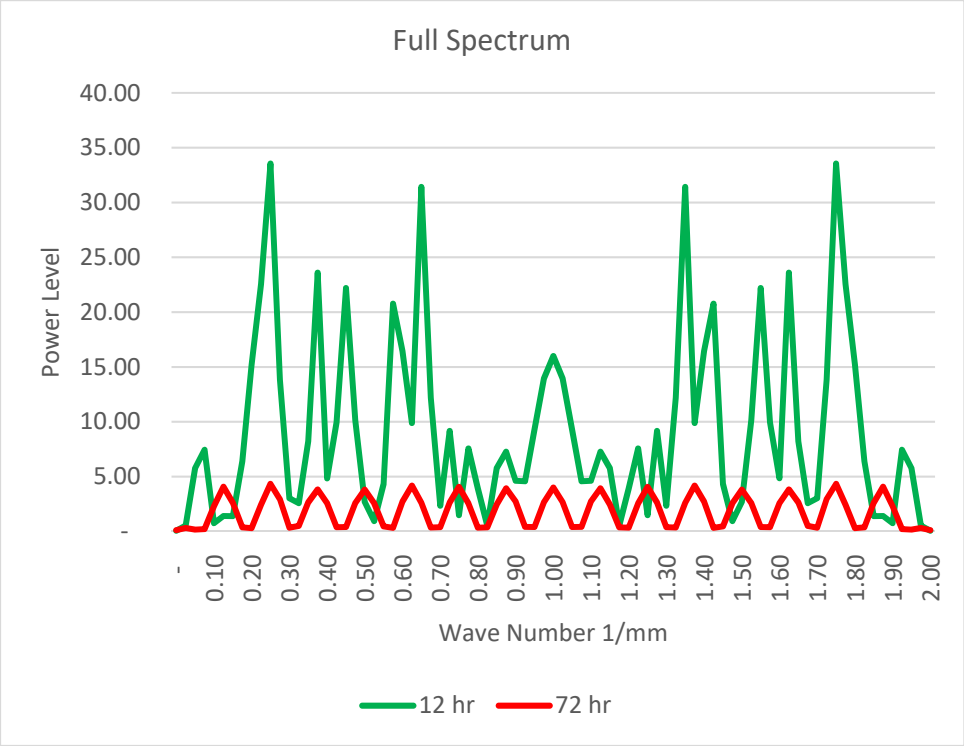
3. Thus we would assume that at t=0 we have an H(t) of 0, at t=a large value we have all growth and thus H(t) also of zero. The maximum H(t) is the time where there is a multiplicity of lots of little disconnected growth. We shall show that with the above diagram.

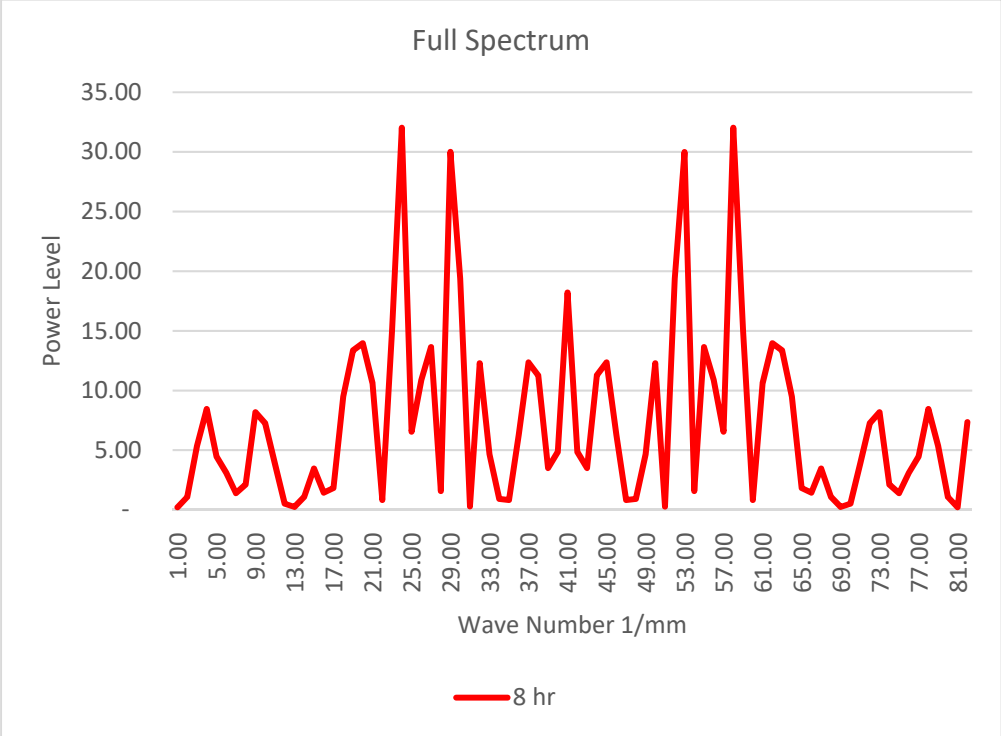
2.3 ANALYSIS

We examine the details on the growth pattern below:



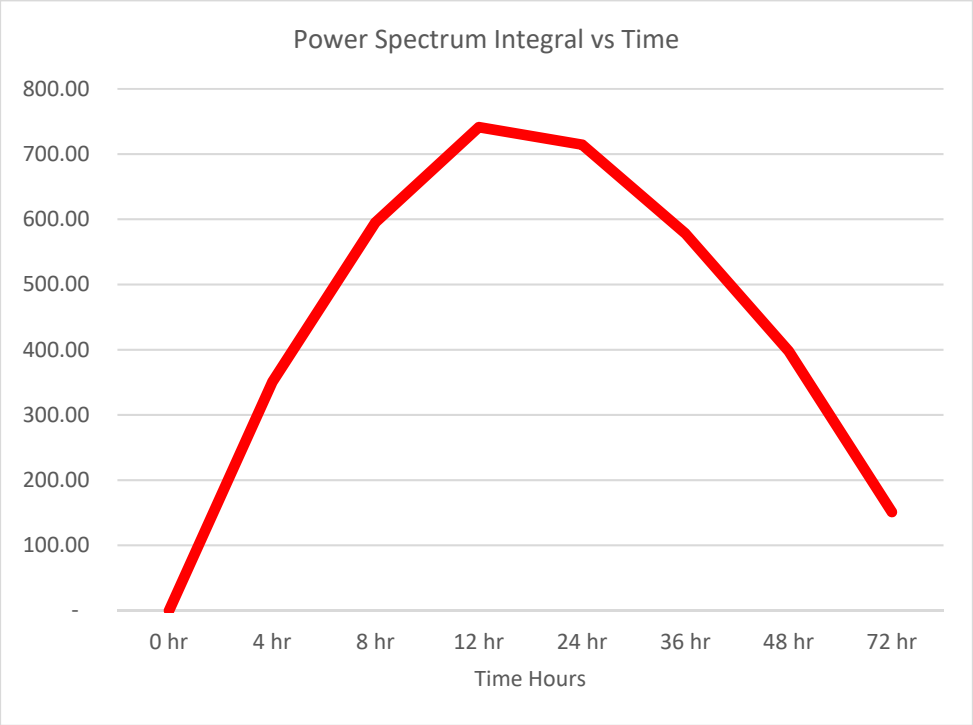
First we present the FT at various times:





Then we show below the Total Power versus time of these spectra. The following slide is the total power spectrum versus time for the sample initially presented above. Note that initially

there is zero power, namely zero fluctuations. Then there is a peak where the growth is maximum. Then is drops as the growth saturates. Time of maximum growth should be maximized for maximum efficacy.



What does this plot mean? Simply:

1. For a t where all samples are zeros, namely the same across all spatial cells, we have zero H(t). That is total continuity.

2. For a t where there is very spotty growth, we have "high" spatial frequency components and when we calculate H(t) for that t value we get a large number reflecting the multiplicity of growth.

3. Thus we would assume that at t=0 we have an H(t) of 0, at t=a large value we have all growth and thus H(t) also of zero. The maximum H(t) is the time where there is a multiplicity of lots of little disconnected growth. We shall show that with the above diagram.

Note on the integrated spectrum we have a peak at 12 hours. That means that the specimen had reached a maximum of regeneration, lots of growth and still lots of no growth.

Examine the actual spectra for 12 and 72 hours. It is symmetric about 1.0 and there are peaks representing the semi periodic peaks representing the spotty regrowth.

Note if we compare the 8 to the 12 hour charts we see first a pair of higher frequency peaks, namely smaller clusters, and second the area under is smaller.

Note in the above the following:

1. There is zero power at the start. Namely there is no growth and the signal is flat and thus has a zero spectrum and in turn a zero integral.

2. Note that as time increases and growth increases the total power increases as multiple cores of growth occur. This technique is useful in this type of analysis because as growth occurs the maximum total power is reached at the point at which the maximum number of disconnected islands of regrowth appear. This we call the maximum island regrowth time, T_{max}.

3. Finally the total power decreases and when it is down to say 10% 0f the peak this is the time to saturation of the sample, T_{sat}. It is a measure of how long it takes to saturate the sample with regrowth.

Thus we can use these two metrics as measures of efficacy.

3 MATHEMATICAL OPTIONS

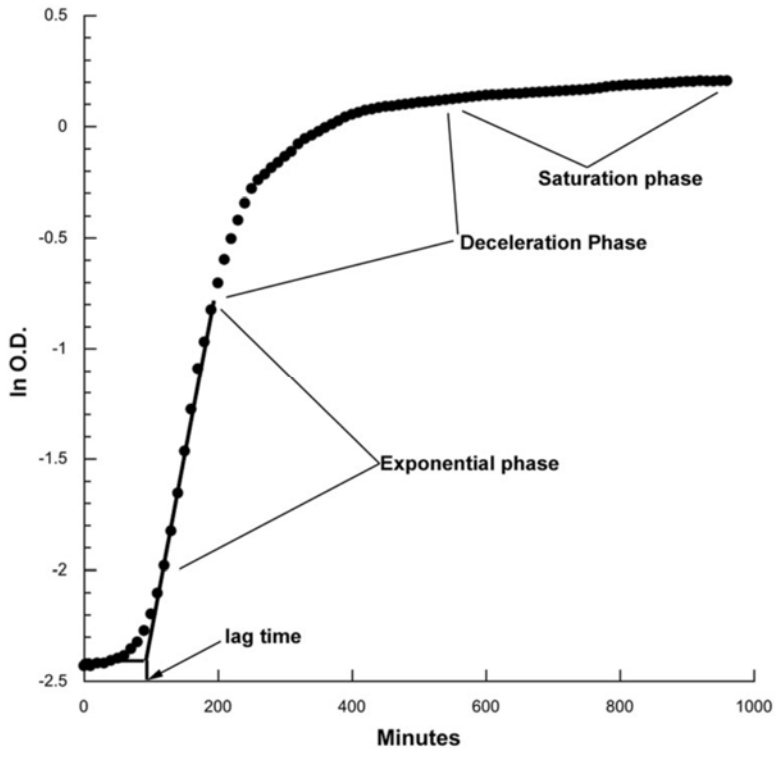
The previous analysis was an ad hoc proposal to examine bacterial growth in a petri dish coated and uncoated. The approach was to generate a simple data driven metric which is reflective of bacterial growth rates and can allows for a ready comparison which is quantitative and reflective of growth. In this section we examine several alternative options where we see this growth determined in detail and then go back and compare it to that of the simplified model.

The general question we pose is:

Bacteria growth are defined by a growth rate, say r. If one looks at a time sequence of growth, say d sec, d being quite small, then a bacteria can remain constant or double. The probability of doubling is say rd. Now does anyone know of a result for the probability density for the number of such cells as a function of r and time? Also can one construct a hypothesis test to examine the growth in two different media and ascertain if one has a higher or lower growth rate than the other and is there a way to calculate a p value?

In ascertaining a statistical validation for a hypothesis test one generally must understand the underlying random process which causes the uncertainty. In the area of bacterial growth, we have a growth rate that has been used for ensemble studies. Namely we assume that the ensemble grows in a deterministic manner and that the variability in that manner is deminimis.

The classic growth rate curve from Hall et al is depicted below:



The exponential phase is what we examine as a determinant of bacteriostatic validity. If we are looking for a metric for ascertaining bacteriostatic efficacy we can estimate the growth rate and then compare one to the other.

Namely:

λ_1 = Growth Rate for Environment 1 (Treated)

λ_0 = Growth Rate for Environment 0 (Un-Treated)

Then we assume that we can measure by means of some efficient estimator each of these rates in the two environments; namely in a treated and untreated environment.

Then we select a hypothesis that says say the treated is better than the untreated as a result of the treated metric being greater than the untreated. However, these metrics are measures based upon noisy or random data so there may be a probability that we say something when it is not correct. This is the classic issue in all hypothesis tests.

In our present case we ask the following:

1. What metric do we use to reflect growth rates?

2. How do we compare those metrics?

3. What is the statistical underpinnings that can cause errors in these measurements?

4. What are the detection and false alarm probabilities of this testing scheme?

These are standard questions but the complexity in this case is more difficult than usual.

3.1 A SIMPLE PROPOSAL

A simple proposal would be as follows:

1. Take the two samples to be compared and treat then accordingly.

2. Periodically say once an hour, measure N cells on each sample and per Hall et al count the bacteria in each micro cell.

3. Using the models discussed in Hall et al, and using a linear estimator approach, calculate an estimate of the growth rates for each sample as an average of the estimates across the N cells over M time periods.

4. Then compare growth rates to ascertain a selected hypothesis.

This is highly data intensive. It does yields reasonable estimates of growth rates and the statistics are efficient[2].

However, when trying to ascertain the detection and/or false alarm rates we have to make certain assumptions regarding the underlying random nature of the individual arrival rates. There could be random noise as well as process noise.

3.2 A SLIGHTLY MORE COMPLEX MODEL

We can expand the model one more step. Let us assume the growth rates are themselves a random process as follows:

$$\lambda_0(t) = \lambda_0 + w_0(t)$$
$$\lambda_1(t) = \lambda_1 + w_1(t)$$
$$w_1(t), w_0(t) = \text{ Gaussian zero mean Wiener processes variance } \sigma$$

Thus we can then determine estimates of the growth rates as defined above. Likewise, we can show that if the estimators are efficient that we have:

$$\widehat{\lambda_i} = \lambda_i + n_i$$
$$and$$
$$\widehat{\lambda_0} - \widehat{\lambda_1} = \lambda_0 - \lambda_1 + n_k$$
$$n_k = n_0 + n_1 = n_0 - n_1$$

That is the difference in the estimated growth rates are the difference in their mean values plus noise. This then allows us to use a standard Gaussian metric for ascertaining performance is we have the standard deviation of the noise.

That standard deviation we could estimate from the noisiness of the actual estimates that is we calculate the variance of the estimated per se. This then gives us a reasonable means to ascertain performance.

3.3 A NONLINEAR STATE MODEL

We now examine another approach to analyzing the data. We assume that the growth rate is itself a random process, that the growth equation is a random process and that the measurement is a random process. We then use this model to ascertain the least squares estimate and its variance using the classic Kushner-Stratonovich equations[3].

[2] See Fleis et al.

[3] See McGarty, Stochastic. This is the standard approach used in a broad spectrum of areas for nonlinear estimation. It is reasonable here. We have used it in a similar case in the late 1960s for the estimation of earth limb reference points.

Let us begin with a simple example as follows:

$$\lambda(t) = \lambda_0$$

or

$$\frac{d\lambda(t)}{dt} = 0 + w_1(t)$$

and

$$\frac{dN(t)}{dt} = \lambda(t)N(t) + w_2(t)$$

and

$$z(t) = N(t) + n(t)$$

Here the w (t) processes are derivatives of Wiener processes and thus are white noise, n (t) is also a Wiener process, and z (t) is the simple measurement of the growth. We can put this in a state model as follows:

$$\frac{dx_1(t)}{dt} = 0 + w_1(t)$$

and

$$\frac{dx_2(t)}{dt} = x_1(t)x_2(t) + w_2(t)$$

or

$$\frac{dx(t)}{dt} = f(x(t), t) + w(t)$$

where

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

and

$$z(t) = x_2(t) + n(t) = Cx(t) + n(t)$$

The above is written in classic state equations. Now the w and n are derivatives of a Wiener process so that they, the derivatives are white noise. Specifically:

$$w_i(t) = \frac{dW_i(t)}{dt}$$

$$E\left[W_i(t)W_j(s)\right] = Q_{ij}\min(t,s)$$

and

$$n_i(t) = \frac{dN_i(t)}{dt}$$

$$E\left[N_i(t)N_j(s)\right] = R_{ij}\min(t,s)$$

Now we seek to find a least square estimate of x, the vector, based upon measurements over some time period and also determine the variance of that estimate. This we directly obtain from the Kushner Stratonovich equations. First we linearize the model (McGarty, Stochastic p 282):

$$f(x) = f(\hat{x}) + A(\hat{x})(x-\hat{x}) + \frac{1}{2}\sum_{i=1}^{n}\gamma_i(x-\hat{x})^T B_i(\hat{x})(x-\hat{x})$$

where

$$A = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} \cdots \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_n}{\partial x_1} \cdots \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$

and

$$B_i = \begin{bmatrix} \frac{\partial^2 f_1}{\partial^2 x_{1n}} \cdots \frac{\partial^2 f_1}{\partial x_n \partial x_1} \\ \frac{\partial^2 f_n}{\partial x_1 \partial x_n} \cdots \frac{\partial^2 f_n}{\partial^2 x_n} \end{bmatrix}$$

We will deal with the first term only.

$\hat{x}(k+1) = [I - K(k+1)C(k+1)]\Phi(k+1,k)\hat{x}(k) + K(k+1)z(k+1)$

and

$P(k+1) = [I - K(k+1)C(k+1)]\Phi(k+1,k)P(k)\Phi^T(k+1,k) + [I - K(k+1)C(k+1)^T + [I - K(k+1)C(k+1)]Q(k)$

$[I - K(k+1)C(k+1)]^T + K(k+1)R(k+1)K^T(k+1)$

and

$$K(k+1) = \Phi(k+1,k)P(k)\Phi^T(k+1,k)C^T(k+1)$$

$$[C(k+1)\Phi(k+1,k)P(k)\Phi^T(k+1,k)C^T(k+1) + C(k+1)Q(k)C^T(k+1) + R(k+1)]^{-1}$$

where

$$x(k+1) = \Phi(k+1,k)x(k) + w(k)$$

$$z(k+1) = C(k+1)x(k+1) + n(k+1)$$

This is the discrete linearized model. Note that it assumes knowledge of the statistics of the three separate noise processes. But we can have an estimate which we know if MMSE, optimum, unbiased and contains a variance.

We may be uncertain about the accuracy of the initial covariance, as stated such inaccuracies may influence P (k), but if P (k) is stable, lack of P (0) certainty may disappear as a serious defect.

Q (k) and R (k) inaccuracies may be quite common. The system noises may be self-induced by the designer to compensate for inaccuracies in his model, for linearization approximation, or for actual phenomenological noises. In general, this noise is the most difficult to model since most scientific experiments are of an exploratory nature to begin with and the phenomena are being investigated, leaving perturbing effects until last. The measurement noise is, in general, the easiest to model and monitor, Examples of phenomenon noise and measurement noise are given by a seismic example where the phenomenon noise may result from measuring equipment or human error.

The above is the optimal design. One can then ascertain in the large that the estimate has a mean and a variance. We can then compare one estimate to another using that data. However, this is a data intensive approach. Thus we examine another approach.

3.4 MONOTONICITY

Now if we cannot estimate the growth rate and instead have a simpler measure then how can it be used? Thus assume:

$T_1 = \hat{\lambda}_1 - \hat{\lambda}_0$

and

$T_2 = f(\lambda_1) - f(\lambda_0)$

and

$T_1 = g(T_2)$

or

$T_2 = h(T_1)$

generally

Let us further assume the functions are all monotonic, continuous, and differentiable.

The using a simple Taylor series we can expand these. This can be seen as follows:

$$h(\lambda + n) = h(0) + \frac{\partial h}{\partial \lambda}\lambda + \frac{\partial h}{\partial n}n + O(\lambda, n)$$

or

$$h(\lambda + n) = a\lambda + bn + O(\lambda, n)$$

Namely for some reasonably small segment we could linearize the metric. The conclusion is that:

1. Estimating growth rates is difficult

2. Using a parametric, albeit discretized but monotonic measure, can in the small provide a reasonable metric

3. Applying a hypothesis test on two estimated growth rates can be done and likewise with the monotonic function as a metric the linearized analogy indicates that the same holds true there.

5. Performance of the test can be ascertained in the standard manner but now using estimated noise values from the test itself.

6. Since an analytical solution to the statistics of the metric are often formidable a simple method of using existing data is acceptable.

3.5 THE METRIC

In the analysis of efficacy of a bacteriostatic technique one uses a metric to compare one method to another. The metric must reflect the rate of growth as well as the time to regrowth. It must also reflect a large enough portion of the sample to make it representative while at the same time no require an unmanageable amount of data analysis. Finally, the metric must be a simple metric, a measure which is repeatable and comparable.

However, whenever we have a metric on data the calculation of that metric may have a statistical variation. Thus we may state a simple hypothesis test as follows for some metric, sat time to saturation:

$$H_1 : T_{Sat,M1} \geq T_{Sat,M0}$$
$$or$$
$$H_0 : T_{Sat,M1} < T_{Sat,M0}$$

We have elsewhere shown how this saturation constant may be obtained.

Let us consider a simpler example. Namely let us assume we can calculate this metric by just counting cells, say 40 cells, in a 20 mm width at 0.5 mm each. Then we define:

T_{Sat} = time at which M cells of N are noted as filled
$N = 20$
$M = 18$

This is a simple case.

Now lest us examine two samples, one with no NSe and one with NSe. We make the assumption that they have different growth rates, as in the above hypothesis test. We want to do that test and see with what accuracy we have a valid result.

Let us thus assume that the bacteria have the following growth rates:

$$\lambda_{No\_Se}$$
$$\lambda_{Se}$$
$$and$$
$$\lambda_{No\_Se} > \lambda_{Se}$$

We then make the measurements:

$$T_{Sat\_No}$$
$$T_{Sat\_Se}$$

From the simple algorithm above.

If each cell is independent, then the probability that any cell is contaminated is independent and we define contamination as some minimal level of growth so that it is observed. Namely a cell is declared as contaminated and thus counted if and only if there are in excess of some specific cell

count. We do not have to know what that is but we know it applies equally to any samples tested uniformly.

Thus:

$$P\left[Cell\_Counted, t\right] = P\left[n(t) > n_0\right]$$

where

$$P\left[n(T) > n_0 \middle| Cell_k\right] =$$

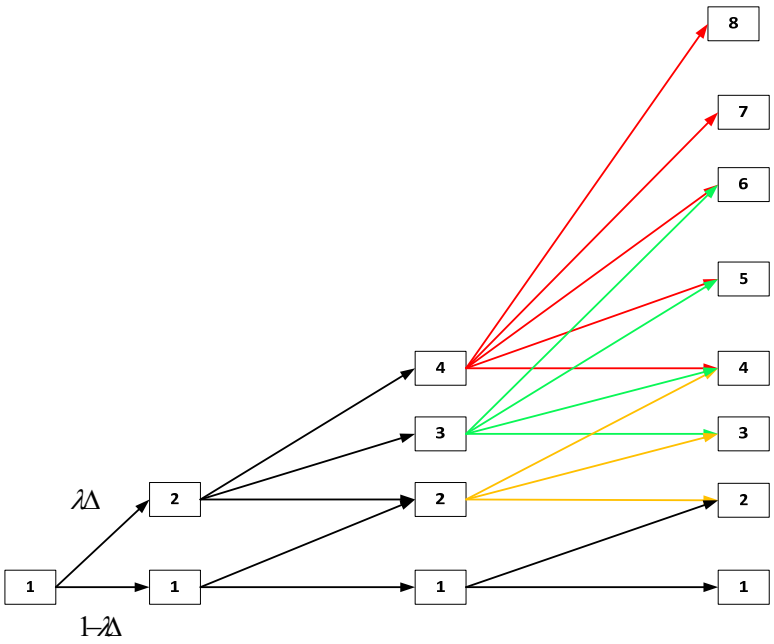
3.6 A SIMPLIFIED ANALYSIS

The model below provides a simple graphic. Note that the probability that a cell divides in some time is:

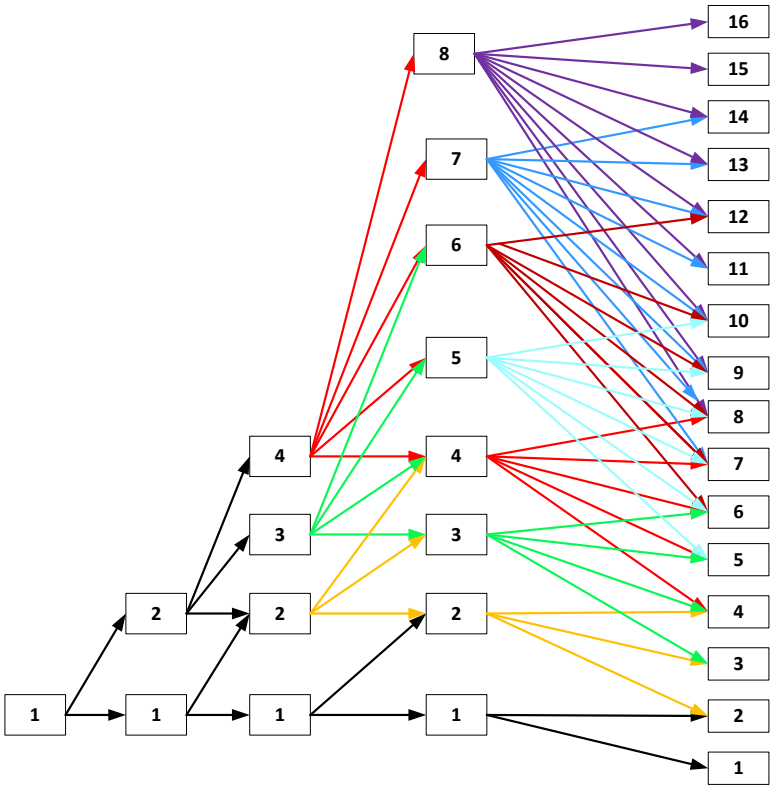
$$P\left[\text{Cell divides in time interval } \Delta\right] = \lambda\Delta$$

$$P\left[\text{Cell does not divide in time interval } \Delta\right] = 1 - \lambda\Delta$$

This is what the graphic depicts. Note however that one starter cell can or cannot divide creating just one or two. Two cells may now act independently. Cell one may divide or not and cell 2 may divide or not so we may have as a result 2, 3 or 4 cells. Likewise, as we progress we may have 2^N cells at time N or any combination less than that. This means that we have a probability distribution of there being any number of cells from the max number down to 1 in this model. If we in addition have two growth rates, then we have two separate probability densities.



and



The two probability densities are:

$p(x|\lambda_1)$
$p(x|\lambda_2)$

Each has a separate form but they differ in mean and standard deviation. The calculation of these densities is quite complex and there does not appear to be any closed form solution. Thus using standard decision techniques even for this approach is problematic at best.

Consider the case of having 8 states and then moving to 16. This is a somewhat complicated step but it does present a basis for a general algorithm. First we show graphically below what we can expect for all possible state changes.

INSERT

Now if we have a typical state change we can define the probability of the next state as a function of the prior probability. For example, take going to state 10 in this phase. How can we get there? Simply:

1. Go from 8 with 2 doublings
2. Go from 7 with 3 doublings
3. Go from 6 with 4 doublings

4. Go from 5 with 5 doublings

Now we define P [k, m] as the probability of being in state k at total state count m. Thus in our case P [5, 8] is the probability of being in state 5 when the total number is 8. We want to calculate P [10, 16], namely the probability of being in state 10 when there are 16 possible states. Let us further assume that p is the probability of a state change and q is 1-p. Then we have for P [10, 16] the following:

$$P[10,16] = C(8,2)P[8,8]p^2q^6 + C(7,3)P[7,8]p^3q^5 + C(6,4)P[6,8]p^4q^2 + C(5,5)P[5,8]p^5q^0$$

and

$$C(n,k) = \frac{n(n-1)...(n-1+k)}{k(k-1)...1}$$

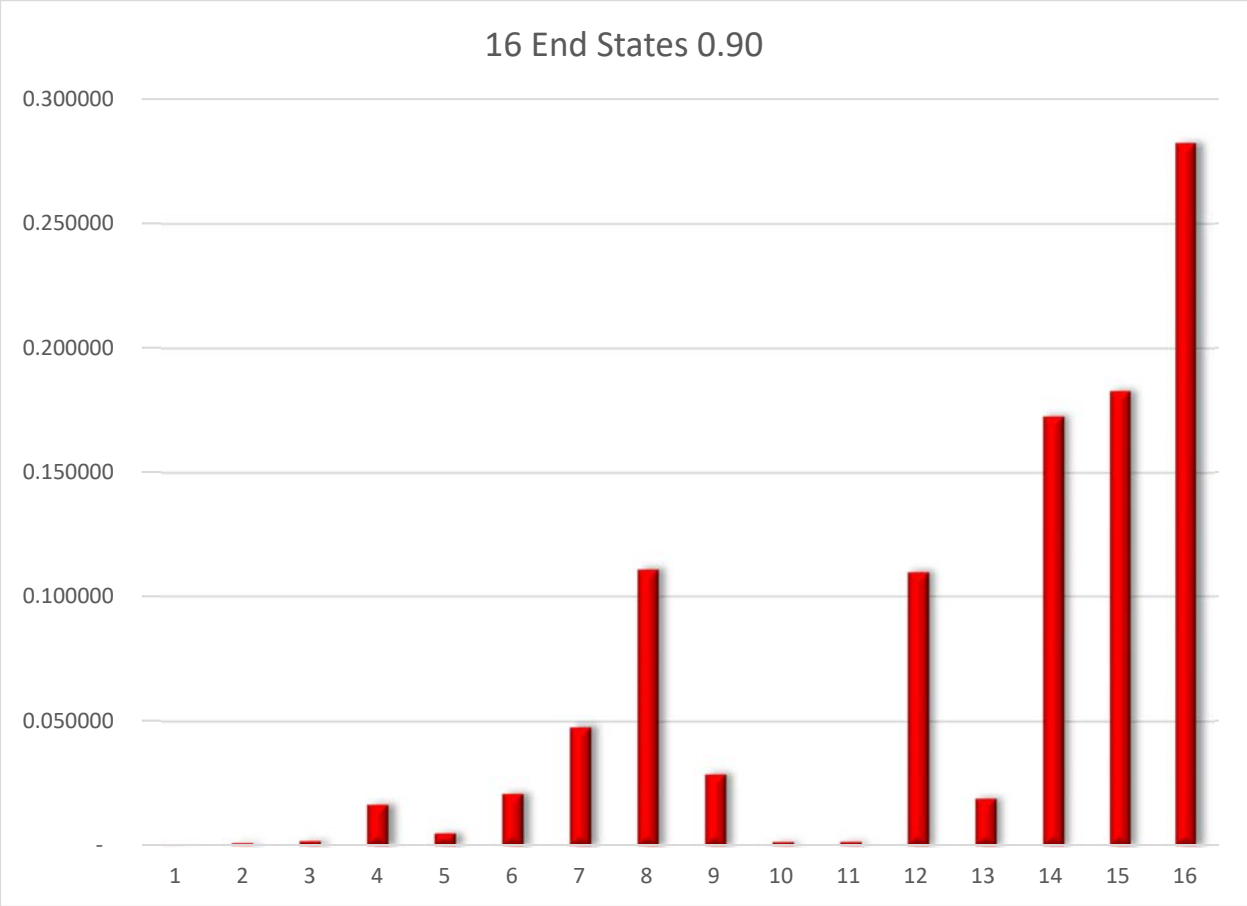
This should be fairly obvious. It lists all the states that can lead to a 10 from the original set of 8. The full set of combinatorics for 8 to 16 are shown below.

		Number Changes								
End State		0	1	2	3	4	5	6	7	8
	1	C(1,0)								
	2	C(2,0)	C(1,1)							
	3	C(3,0)	C(2,1)							
	4	C(4,0)	C(3,1)	C(2,2)						
	5	C(5,0)	C(4,1)	C(3,2)						
	6	C(6,0)	C(5,1)	C(4,2)	C(3,3)					
	7	C(7,0)	C(6,1)	C(5,2)	C(4,3)					
	8	C(8,0)	C(7,1)	C(6,2)	C(5,3)	C(4,4)				
	9		C(8,1)	C(7,2)	C(6,3)	C(5,4)				
	10			C(8,2)	C(7,3)	C(6,4)	C(5,5)			
	11				C(8,3)	C(7,4)	C(6,5)			
	12					C(8,4)	C(7,5)	C(6,6)		
	13						C(8,5)	C(7,6)		
	14							C(8,6)	C(7,7)	
	15								C(8,7)	
	16									C(8,8)

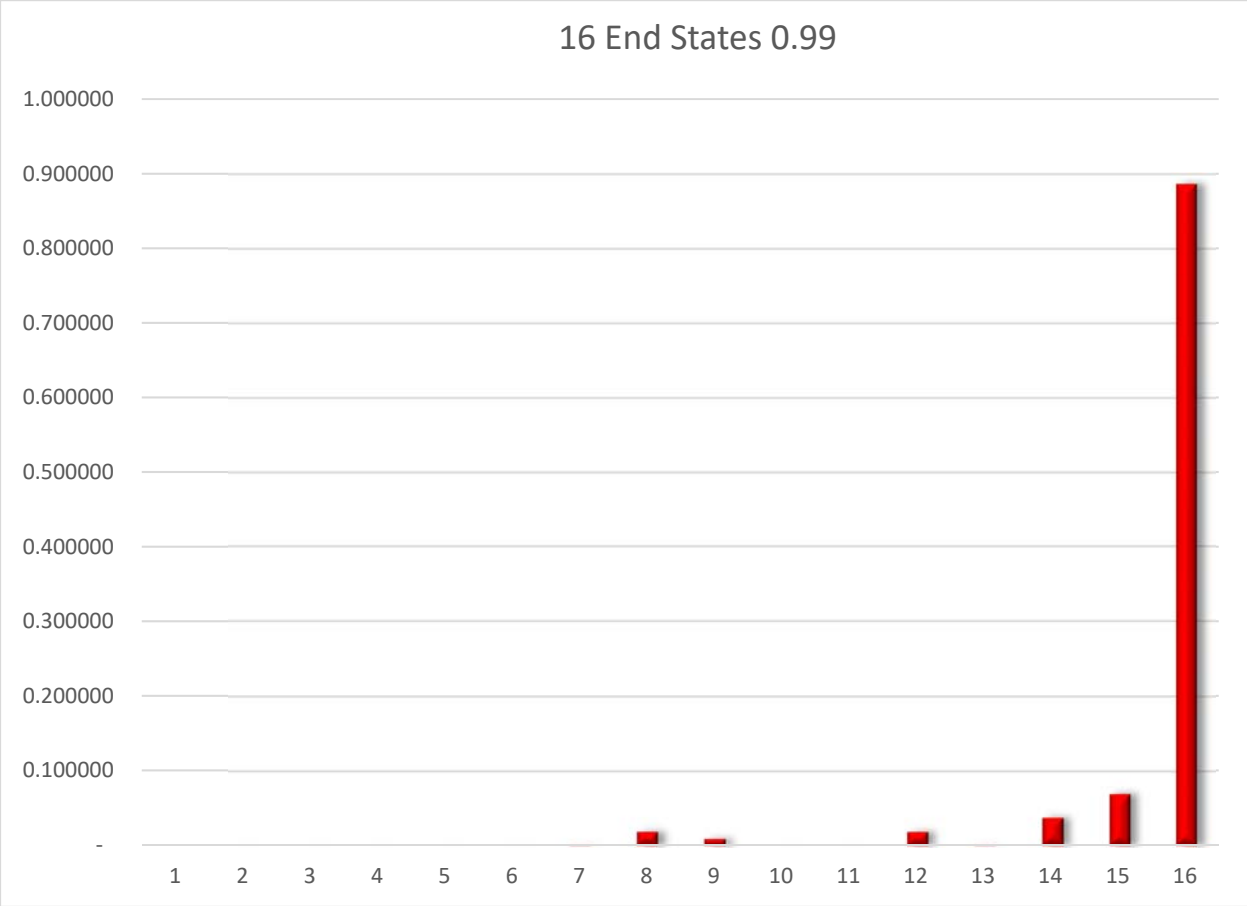
Using this algorithm and for a p=0.9 and q=0.1 we obtain the probability as shown below. This density function is the function for:

$$P\left[N(t) = n \middle| \lambda = 0.9, N_{count} = 16\right]$$

The chart below is for p=0.90.

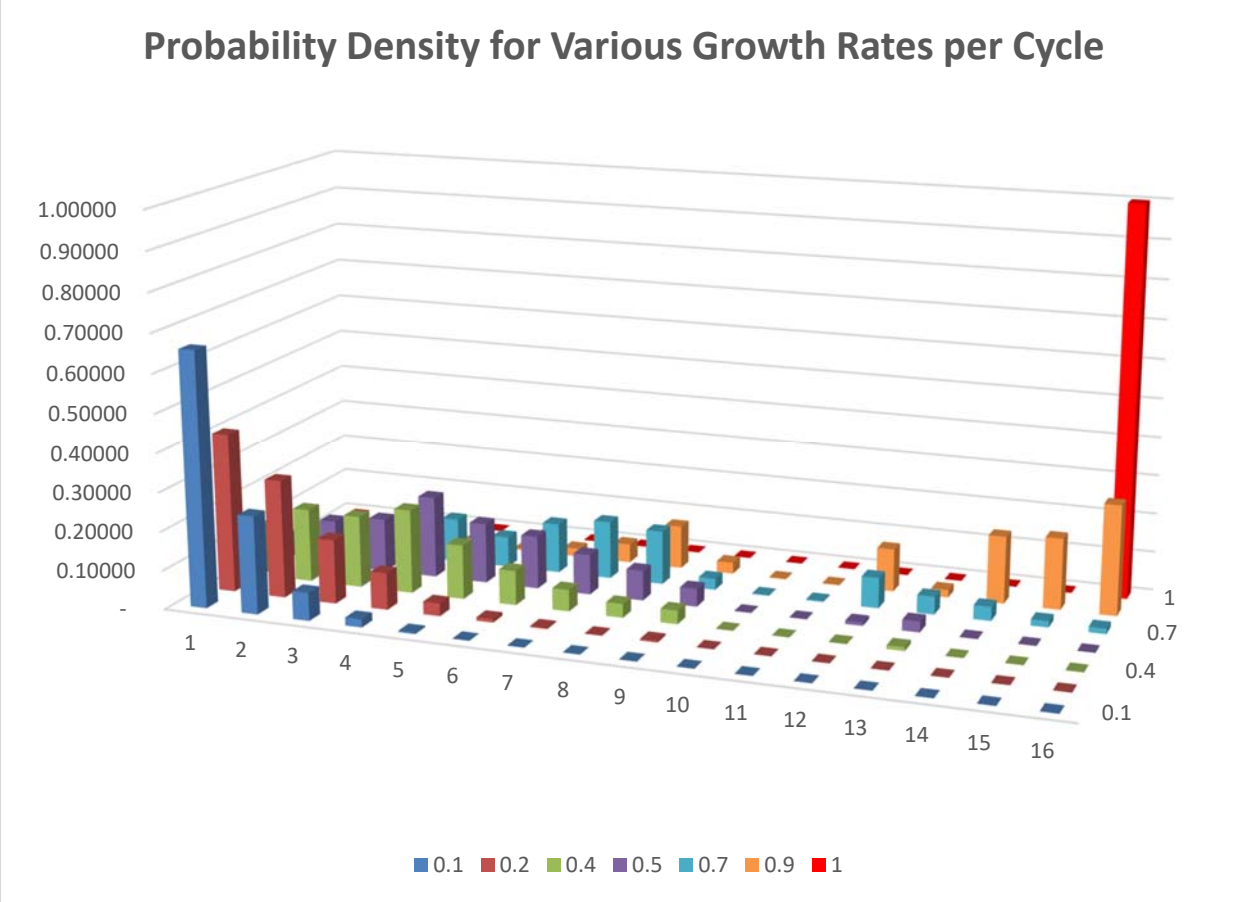


The chart below is for 0.99



Note the significant difference. A lower transition rate, slower growth if you will, is spread across a wide base of possible states. That is not the case for a higher growth rate. A likelihood ration is thus the ratio of the two. This is complex.

The best way to visualize this is to examine the density for various growth rates per cycle, from 0.1 to 1.0. We show this below.



As the growth rate approaches zero the density peaks there and as it approaches 1 it peaks at the max. However, for intermediate growth rates it is everywhere. The discriminant is highly multimodal and unwieldy in this form.

Now the algorithm for calculating is straightforward but complex in detail. Specifically:

1. For any N set of states we list out the set of 2N follow on states. As we had graphically done before.
2. Then we use the tabular form above to determine the C (n, k) values for each P [k, N] where we need to complete the table of N columns and 2N rows. The completion need only be conceptual in the algorithm
3. Transitions are allowed as depicted above. Namely, if we have an end state of say n, and we have k transitions, that is k doublings, we must have k doublings and j remain constant, so that the original state is k+j and further the end state is n so that n=k+j. We can algorithmatize this process simply by moving through the possibilities and allowing only those that meet the criteria.
4. Then we create the sums which contain the products of the C, P, p, and q that meet the specific integer states.
5. Finally add then up appropriately and display.

There does not appear to be any closed form solution to this nor does there appear to be an extreme bound that is useful. The above graph points out several factors. First there are multiple peaks and second the end point peak is Large but not overwhelming so.

However, another simple approach is to use what we have in bacterial growth. The problem however is this model is deterministic and not random. Namely, we have for the number of bacteria the following:

$$\frac{dN(t)}{dt} = \alpha N(t)$$

or

$$\frac{dN(t)}{N(t)} = \alpha t$$

Either can be used in exponential growth.

Thus based upon the prior analysis, limited as it may be, we suggest that the following can be done:

1. Use the quantized metric to ascertain comparable metrics for hypothesis tests. Namely to determine the following:

$$H_1 : \lambda_1 \geq \lambda_2$$
$$H_0 : otherwise$$

We can use the test

$$H_1 : T_{Sat\_No} \geq T_{Sat\_Se}$$
$$H_0 : otherwise$$

2. Furthermore we can ascertain the p value of the test by assuming the noise is Gaussian with zero mean and a variance determined by the variance of the detected data. This assumes an unbiased estimator.

4 OBSERVATIONS

The metric used in the proposed efficacy test is clearly a compromise and used for the purpose of ascertaining an estimate of relative efficacy with a simple test and measurement. Our objective is to provide some justification of the test. Attempting to provide a detailed description appears quite complex and we provide what appears as a reasonable analysis of the usefulness of the test *as well as a means to provide a measure of its robustness.

Clearly additional analysis will be required but this analysis provides some reasonable foundation for the application of the test as proposed.

We can make the following observations:

1. Note on the integrated spectrum we have a peak at 12 hours. That means that the specimen had reached a maximum of regeneration, lots of growth and still lots of no growth

2. Examine the actual spectra for 12 and 72 hours. It is symmetric about 1.0 and there are peaks representing the semi periodic peaks representing the spotty regrowth.

3. Note if we compare the 8 to the 12 hour charts we see first a pair of higher frequency peaks, namely smaller clusters, and second the area under is smaller.

5 REFERENCES

1. Christian, R., et al, Comparison of Methods for Measurement of Bacterial Growth Rates in Mixed Batch Cultures, Applied and Environmental Microbiology, May 1982, p. 1160-1165.

2. Dengremont, E., and J. M. Membre, Statistical Approach For Comparison Of The Growth Rates Of Five Strains Of Staphylococcus Aureus, Applied And Environmental Microbiology, Dec. 1995, P. 4389–4395

3. Fleis. J., et al, Statistical Methods for Rates and Proportions, Wiley (NY) 2003.

4. Hall, B. et al, Growth Rates Made Easy, Mol. Biol. Evol, October 28, 2013.

5. McGarty, T. P., Stochastic Systems and State Estimation, Wiley (New York) 1973.

6. Murray, J., Mathematical Biology, Springer (New York) 1989.

7. Prescott L., et al, Microbiology, 5th Ed McGraw Hill (New York) 2002.

8. Rohatgi, V., Statistical Inference, Dover (Mineola) 2003

9. Snyder, D., Random Point Processes, Wiley (New York) 1975

10.

6 APPENDIX

This Appendix contains some constants, basic physical measures as well as a suggested protocol for measuring.

6.1 SOME VALUES

The following are a set of some values of size, mass, etc. that may be of use in this analysis.

6.1.1 Cell sizes:

Bacteria (E. coli): ~1mm diameter, 2mm length (2,000,000 μm), ~1mm3 in volume; 109cells/ml in an overnight culture (OD600~1), 4 sq. mm surface

Yeast (Saccharomyces cerevisiae): ~5 mm diameter, ~50 mm3 in volume

Mammalian (HELA) cell - ~2,000 mm3 in volume, adherent cell on a slide ~20 mm diameter ~100,000 cells in a confluent well of a 96 multiwall plate

6.1.2 Organelles and cell constituents:

Mammalian cell nucleus ~10-micron diameter
Mitochondria ~1-2-micron length, ~0.2-0.7-micron diameter
Chloroplast ~4-micron length, ~1-micron diameter
Cell membrane ~5-10 nm,
"Average" protein ~2nm,
water molecule ~0.2nm

6.2 TEST PROCEDURE

This is a draft description of the tests which will be performed using various antibacterials along with Vexti bacteriostatic inhibitors using nano selenium.

6.2.1 Pathogens Proposed

The following pathogens shall be tested as available at Vexti or Northeastern facilities.

6.2.2 Bacteria

1. Enterobacter aerogenes
2. Pseudomonas aeuginosa
3. Staphylococcus aureus
4. Enterococcus faecails VRE

5. Staphylococcus aureus MRSA
6. Pseudomonas aerogenes MBL

6.2.3 Virus

1. Norovirus (feline Calicivirus)
2. Influenza A, H3N2 avian Virus

6.2.4 Bactericidal Products

The products used in the testing should be selected so as to have industry standard bactericidal efficacy.

6.3 TEST PROTOCOLS

Tests shall be performed in accordance with EPA Guidelines:

1. Product Performance Test Guidelines
OCSPP 810.2200:
Disinfectants for Use on Hard Surfaces—Efficacy Data Recommendations

2. Product Properties Test Guidelines
OPPTS 830.6317
Storage Stability

Other techniques shall be performed under generally accepted industry standards.

6.3.1 Spatial and Temporal Analysis

1. Select a pathogen
2. Select NSe concentrations, say 50, 100, 150 ng/sqcm
3. Coat 4 petri dishes, wipe with cloth and with no NSe and the 3 NSe concentrations
4. Measure photographically the results at 0, 4, 8, 12, 24, 36, 48, 72 hours. That is 8X4 or 32 photographs
5. Using a Photoshop metric to enter data into the analysis and plot power versus time for each application
6. Determine efficacy of each NSe coating and benchmark with none.

6.3.2 Results to be Presented

The results will be of the following form. It shall be an N X M table for selected times demonstrating the efficacy utilizing the above referenced test procedures across a selection of wipes (N) and a selection of pathogens (M).

The data will be from the Total Power measures and will specifically the peak time and the maximum time. Efficacy will be related to how large these values are and the statistical significance of the spread between these values.

The results will generally appear as shown below (or as modified by Vexti dependent on laboratory results and a separate Table will be prepared for T_{Max} and T_{Sat} each:

Wipe/Pathogen	Bleach Germicidal Disposable Wipe	HB Germicidal Disposable Wipe	AF3 Germicidal Disposable Wipe
NSe Concentration (ng/sqcm)	0,50,100, or 150	0,50,100, or 150	0,50,100, or 150
Enterobacter aerogenes	T_{Max} and T_{Sat} each	T_{Max} and T_{Sat} each	T_{Max} and T_{Sat} each
Pseudomonas aeuginosa	T_{Max} and T_{Sat} each	T_{Max} and T_{Sat} each	T_{Max} and T_{Sat} each
Staphylococcus aureus MRSA	T_{Max} and T_{Sat} each	T_{Max} and T_{Sat} each	T_{Max} and T_{Sat} each

The table above will be prepared and completed using time intervals to be appropriately selected per protocols, generally at 0, 4, 8, 12, 24, 36, 48, 72 hours. The tables will be prepared without and with NSe treatments.

Namely the Table will be prepared for NSe concentrations of 0, 50, 100, 150 ng/sqcm.

6.3.3 Statistical Analyses

It is recognized that the metrics used have statistical variation and thus statements such as the following:

$$T_{Sat,None} < T_{Sat,NSe}$$

Require some statistical measure. This will be examined separately but given the large set of samples used in the determination of these numbers a statistical significance seems high if the numbers are materially different.