Journal of Microbiological Methods 107 (2014) 214–221

Contents lists available at ScienceDirect

Journal of Microbiological Methods

journal homepage: www.elsevier.com/locate/jmicmeth

Estimation method for serial dilution experiments

Avishai Ben-David [a,*], Charles E. Davidson [b]

[a] RDECOM, Edgewood Chemical Biological Center, Aberdeen Proving Ground, MD 21010, USA
[b] Science and Technology Corporation, Belcamp, MD 21017, USA

ARTICLE INFO

Article history:
Received 9 May 2014
Received in revised form 29 August 2014
Accepted 30 August 2014
Available online 7 September 2014

Keywords:
Titration
Serial dilution
Viable bacterial counts
Agar plates
Too numerous to count (TNTC)
Density (concentration) of microorganisms

ABSTRACT

Titration of microorganisms in infectious or environmental samples is a corner stone of quantitative microbiology. A simple method is presented to estimate the microbial counts obtained with the serial dilution technique for microorganisms that can grow on bacteriological media and develop into a colony. The number (concentration) of viable microbial organisms is estimated from a single dilution plate (assay) without a need for replicate plates. Our method selects the best agar plate with which to estimate the microbial counts, and takes into account the colony size and plate area that both contribute to the likelihood of miscounting the number of colonies on a plate. The estimate of the optimal count given by our method can be used to narrow the search for the best (optimal) dilution plate and saves time. The required inputs are the plate size, the microbial colony size, and the serial dilution factors. The proposed approach shows relative accuracy well within ± 0.1 $\log_{10}$ from data produced by computer simulations. The method maintains this accuracy even in the presence of dilution errors of up to 10% (for both the aliquot and diluent volumes), microbial counts between 10^4 and 10^{12} colony-forming units, dilution ratios from 2 to 100, and plate size to colony size ratios between 6.25 to 200.

1. Introduction

Quantitative estimation of the number of viable microorganisms in bacteriological samples has been a mainstay of the microbiological laboratory for more than one-hundred years, since Koch first described the technique (Koch, 1883). Serial dilution techniques are routinely used in hospitals, public health, virology, immunology, microbiology, pharmaceutical industry, and food protection (American Public Health, 2005; Hollinger, 1993; Taswell, 1984; Lin and Stephenson, 1998) for microorganisms that can grow on bacteriological media and develop into colonies. A list of bacteria that are viable but nonculturable (VBNC), the detection of such microorganisms, and the process of resuscitation of cells from VBNC state are addressed by Oliver (2005, 2010). In the work presented here it is assumed that the microorganisms are culturable.

The objective of the serial dilution method is to estimate the concentration (number of colonies, organisms, bacteria, or viruses) of an unknown sample by counting the number of colonies cultured from serial dilutions of the sample, and then back track the measured counts to the unknown concentration.

Given an unknown sample which contains n_0 colony forming units (CFUs), a series of J dilutions are made sequentially each with a dilution factor α. From each of the J dilutions a fraction α_p^{-1} is taken and spread (plated) on an agar plate (assay) where colonies are counted. Thus, in general there are two dilution factors: α and α_p. For example, $\alpha = 10$ indicates a 10-fold dilution, e.g., by diluting successively 0.1 ml of sample into 0.9 ml of media; and $\alpha_p = 1$ means that the entire volume is spread (plated) on the agar plate. For an experiment with a larger dilution factor α_p, multiple plates may be spread at the same dilution stage. For example, $\alpha_p = 20$ represent a 5% plating of the dilution, and thus up to 20 replicates could be created. At each dilution the true number of colonies is $n_j = n_0\alpha^{-j}\alpha_p^{-1}$ and the estimated number is $\hat{n}_j$. The estimated quantities are denoted with a "hat" (estimated quantities can be measured quantities, or quantities that are derived from measured or sampled quantities); symbols without a "hat" denote true quantities (also known as population values in statistics) that do not contain any sampling or measurement error. In this work both n_j and n_0 are "counts", i.e., number of colonies. Knowing the aliquot volume, one can easily convert counts to concentration (for example CFU/ml).

The importance of serial dilution and colony counting is reflected by the number of standard operating procedures and regulatory guidelines describing this methodology. In all of these guidelines the optimal number ($\hat{n}_j$) of colonies to be counted has been reported (Park and Williams, 1905; Wilson, 1922; Jennison and Wadsworth, 1940; Tomasiewicz et al., 1980; FDA, 2001; Goldman and Green, 2008) as 40–400, 200–400, 100–400, 25–250, 30–300. It is interesting to note that these

Abbreviations: TE, total error; RE, relative error; MPN, most probable number; TNTC, too numerous to count; VBNC, viable but nonculturable; CFU, colony forming unit.

* Corresponding author. Tel.: +1 410 436 6631; fax: +1 410 436 1120.
E-mail addresses: avishai.bendavid@us.army.mil (A. Ben-David), charles.e.davidson2.ctr@us.army.mil (C.E. Davidson).

http://dx.doi.org/10.1016/j.mimet.2014.08.023
0167-7012/

references do not specify the area in which the colonies grow, nor the diameter of the particular organism assayed. The result is that titration and counting colonies is done within a range that may be inadequate, and may introduce considerable error. In our work these parameters are addressed.

The main challenge in serial dilution experiments is the estimation of the undiluted microorganisms counts n_0 from the measured $\hat{n}_j$. There are two competing processes (Tomasiewicz et al., 1980) that affect the accuracy of the estimation: sampling errors and counting errors. Sampling errors are caused by the statistical fluctuations of the population. For example, when sampling an average of 100 colonies, the fluctuations in the number of the population are expected to be $\pm\sqrt{100}$ when the sampling process is governed by a Poisson probability (Poisson and Binomial distributions are often used in statistical analysis to describe the dilution process (Hedges, 2002; Myers et al., 1994)) where the standard deviation equals square-root of the mean; the relative error (ratio of the standard deviation to the mean) is $\sqrt{100}/100 = 0.1$. Thus, the larger the sample size is, the smaller the relative sampling error; hence, one would like to use a dilution plate with the largest number $\hat{n}_j$ (i.e., the least diluted sample, $j \to 1$). However, as the number of colonies increases, counting error is introduced due to the high probability of two (or more) colonies to merge (due to overcrowding) and become indistinguishable, and be erroneously counted as one colony. An optimum (a "sweet spot") between these two processes (sampling and counting error) needs to be found for using the optimal dilution $\hat{n}_j$ (i.e., the optimal jth plate) with which to estimate n_0. Cells can grow into colonies in various ways. Wilson (1922) states that when two cells are placed very close together only one cell will develop, and when two cells are situated at a distance from each other both cells may grow and then fuse into one colony. Either way, the end result is the appearance of one colony which causes counting error.

Estimation of bacterial densities from the most probable number (MPN) method (Cochran, 1950) requires multiple replicates of the jth dilution plate, and analyzes the frequency of plates with zero colonies instead of using counts directly. MPN is often used to measure microbes in milk, water and food (Blodgett, 2010). The MPN method (Cochran, 1950) "is of low precision, as is to be expected from a method that does not use direct counts. Large number of samples [replicate agar plates] must be taken at each dilution if a really precise result is wanted". In our work we seek a method where the counts from a single plate are used to estimate bacterial concentrations.

A simple method to estimate the number of colonies n_0 in the unknown sample from the counted number of colonies $\hat{n}_j$ at the jth assay is presented. Our method is easy to implement. The method selects the optimal count (i.e., a best single plate) with which n_0 is estimated. There are only a few inputs needed: the incubation plate size, the microbial colony size, and the dilution factors (α and α_p). The dilution error (although present in the serial dilution experiment) is not an input. The relative accuracy of our method is well within $\pm$ 0.1 $\log_{10}$ (i.e., within 100% error) which is much better than the common requirement of $\pm$ 0.5 $\log_{10}$ (i.e., within 500% error) that is often regarded as accepted accuracy in many biological experiments.

2. Material and methods

The measured (counted) number of colonies $\hat{n}_j$ is related to the true number of colonies n_j by $n_j = \hat{n}_j + \delta_j$ where δ_j is a bias that accounts for uncounted colonies due to the merging (overcrowding) of nearby colonies, and thus, $n_j \geq \hat{n}_j$. The challenge is to obtain an estimate of δ_j from a single measurement $\hat{n}_j$ of the jth Petri dish. The challenge is met in an ad-hoc manner. The following assumptions are made: (i) The true n_j (when no colonies are miscounted) is described by a Poisson probability for which the variance equals the mean (Forbes et al., 2011), (ii) The probability density function for $\hat{n}_j$ (with the effect of merging of nearby colonies) is a displaced version (by δ_j) of the Poisson distribution for n_j, see Fig. 1. Hence the variance of the probability density function

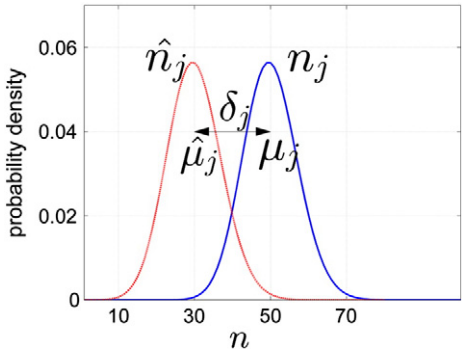

Fig. 1. The concept of δ_j is given with Poisson and a shifted-Poisson probability density functions for the true n_j, and for the counted (measured) $\hat{n}_j$, respectively. $\hat{n}_j$ is the observed (counted) number of colonies on plate j of the serial dilution process, for which the true number of colonies is n_j with a mean μ_j. Due to uncounted colonies in the measuring process (i.e., merging of colonies that are counted as one colony due to overcrowding), $\hat{n}_j < n_j$.

for n_j is the same as that for $\hat{n}_j$, and therefore the variance $Var(\hat{n}_j) = Var(n_j) = E\left[\left(n_j - \mu_j\right)^2\right]$. $E(\cdot)$ is the expectation operator (i.e., an average process), and μ_j is the mean of n_j. There is not enough data to compute a variance with the expectation operator (because only one plate with one value of $\hat{n}_j$ for a dilution $\alpha^{-j}\alpha_p^{-1}$ is available). Therefore, we define a measure of "spread" given by $V_n = (n_j - \mu_j)^2$, that is computed from a single value of n_j. The "spread" can be solved for μ_j by $\mu_j = n_j - \sqrt{V_j}$. The distance $\delta_j < \mu_j$ (see Fig. 1), and thus there must be a constant $c < 1$ such that $\delta_j = c\mu_j = cn_j - c\sqrt{V_n}$. Because we do not want to modify the individual n_j values (with c), we instead construct δ_j with a constant $k > 1$ where $\delta_j = n_j - k\sqrt{V_n}$, and k only modifies the spread $\sqrt{V_n}$. Our notion of "spread" is a substitute for the notion of variance (that we cannot compute) and is weak; hence, no harm is done in adjusting it with a fudge factor k. The factor k is a function of the geometry of the serial dilution experiment. At this point it suffices to state that k exists (the numerical value of k is addressed later). With the Poisson and displaced Poisson assumptions for n_j and $\hat{n}_j$, we set $V_n = V_{\hat{n}} = \hat{n}_j$, leading to $\delta_j = n_j - k\sqrt{\hat{n}_j}$. We don't have access to n_j, and therefore we replace the unknown population value n_j with the measured $\hat{n}_j$. This is inspired by the principle that is often used in signal processing of replacing unknown population parameters with maximum-likelihood estimates (as is done, for example, in the generalized likelihood ratio test (Scharf and Friedlander, 1994)).

With an undetermined k (for the moment) a model-estimate of the true δ by $\hat{\delta}$ is given by

$$\left\{\begin{array}{l} \hat{\delta}_j = \hat{n}_j - k\sqrt{\hat{n}_j} \\ \hat{\delta}_j > 0 \quad for \quad 1 < k < \sqrt{\hat{n}_j} \\ \hat{\delta}_j = 0 \quad for \quad k \geq \sqrt{\hat{n}_j} \end{array}\right\}. \tag{1}$$

The inequalities in Eq. (1) are necessary to ensure $\hat{\delta}$ to be a non-negative quantity less than the value of μ. Eq. (1) implies that the counting error is negligible $\left(\hat{\delta} \to 0\right)$ when $\hat{n}_j < k^2$. Given $\hat{\delta}_j$ in Eq. (1) we proceed to estimate n_0 by

$$\hat{n}_0 = \left(\hat{n}_j + \hat{\delta}_j\right)\alpha^j\alpha_p. \tag{2}$$

Eq. (2) produces an estimate of n_0 by correcting the count for the expected number of missed colonies $\hat{\delta}_j$, and by multiplying by the total dilution factors $\alpha^j\alpha_p$ (to reverse the effect of the serial dilution α^{-j} and the

plating dilution α_p^{-1} in $n_0 \to n_j$). The question remains "which jth plate to use for $\hat{n}_j$?" The rule of thumb that is commonly used (Tomasiewicz et al., 1980) is to choose a plate with $30 \le \hat{n}_j \le 300$ or $25 \le \hat{n}_j \le 250$. With our method $\hat{n}_j$ is chosen as follows.

The standard definition of total error (also known as mean squared error, see p. 330 in Casella and Berger, 2002) for an estimated quantity $\hat{n}_0$ is given by $\sqrt{E\left[(\hat{n}_0-n_0)^2\right]} = \sqrt{\mathrm{var}(\hat{n}_0) + [E(\hat{n}_0)-n_0]^2} = \alpha^j\alpha_p\sqrt{(\hat{n}_j+\delta_j) + \left(\delta_j-\hat{\delta}_j\right)^2}$ where $[E(\hat{n}_0)-n_0]^2$ is the error due to the bias of the estimates. The total error combines the effects of precision (via the variance of the estimated $\hat{n}_0$) and accuracy (via the bias in estimating $\hat{n}_0$). In computing total error, one performs many Monte Carlo simulations (where n_0, δ_j, and all other parameters are known) for the dilution process, constructs estimates for n_0, and computes the variance and the mean from all the estimates of n_0. This is what we do later in Section 3 (Results) when we validate our estimation method. However, the true δ_j is not known in a real scenario (only $\hat{n}_j$, α^j, and α_p are known), and thus we resort to the estimated $\hat{\delta}_j$ with our model as is given in Eq. (1) to construct an estimator of total error. Our total error is given as a sum of the sampling error and the counting error. The sampling error which is due to the variance of the sampled n_j on a plate is given by $\mathrm{var}(n_j) = \mathrm{var}\left(\hat{n}_j+\hat{\delta}_j\right) = \hat{n}_j+\hat{\delta}_j$ (using the Poisson distribution for $n_j = \hat{n}_j+\delta_j$ whose Poisson intensity parameter λ is estimated as $\lambda = E(n_j) = \hat{n}_j+\hat{\delta}_j$), and the counting error on a plate that is due to the residual bias given by $\left(\delta_j-\hat{\delta}_j\right)^2$ which we approximate as $\hat{\delta}_j^2$ (because the true δ_j is not known). Our total error (TE) model is given by

$$\left\{\begin{aligned} TE_j &= \sqrt{(\text{sampling error})^2 + (\text{counting error})^2} \\ &= \alpha^j\alpha_p\sqrt{Var\left(n_j\right)+\hat{\delta}_j^2} \\ &= \alpha^j\alpha_p\sqrt{\left(\hat{n}_j+\hat{\delta}_j\right)+\hat{\delta}_j^2} \end{aligned}\right\}. \tag{3}$$

The total error in estimating n_0 originates from sampling error (the variance of the true n_j that is transformed by $\alpha^j\alpha_p$ to the space of n_0) and the counting error due to $\hat{\delta}_j$ (also transformed to the n_0 space). We construct a table for TE_j values for each serial dilution plate, and choose the $\hat{n}_j$ value that produces the smallest TE_j (example is shown later in Table 1). With this optimal $\hat{n}_j$, n_0 is estimated with Eq. (2). The predicted total error in Eq. (3) underestimates the truth (to be shown in the results section) because usually $\left(\delta_j-\hat{\delta}_j\right)^2 > \hat{\delta}_j^2$. Note that $\hat{\delta}_j$ in Eq. (1) cannot exceed the value $\hat{n}_j-\sqrt{\hat{n}_j}$ (which occurs when $k=1$), whereas $\delta_j-\hat{\delta}_j$ is unbounded, because the true $\delta_j \le n_j$ is only bounded by n_j, where $n_j \ge \hat{n}_j$ (Fig. 1). Furthermore, $\hat{\delta}_j$ is set to zero in Eq. (1) for $k \ge \sqrt{\hat{n}_j}$ whereas the true δ_j only approaches zero when the colony size approaches zero (shown later in Eq. (8)). Hence, usually $\delta_j > \hat{\delta}_j$. Note that TE becomes very large as the counts $\hat{n}_j \to 0$ because α^j increases rapidly ($\alpha^j \propto 1/\hat{n}_j$, small counts occur at large j values with huge dilution).

The estimated uncertainty (error) for the solution $\hat{n}_0$ (Eq. (2) and its relative error (RE), are given by

$$\left\{\begin{aligned} \hat{n}_0 &= \left(\hat{n}_j+\hat{\delta}_j\right)\alpha^j\alpha_p \pm TE_j \\ RE &= \frac{\Delta\hat{n}_0}{\hat{n}_0} = \frac{TE_j}{\hat{n}_0} = \frac{\sqrt{\left(\hat{n}_j+\hat{\delta}_j\right)+\hat{\delta}_j^2}}{\left(\hat{n}_j+\hat{\delta}_j\right)} \end{aligned}\right\} \tag{4}$$

where TE is given in Eq. (3), $\hat{\delta}_j$ is given Eq. (1), and the value of k (with which $\hat{\delta}_j$ is computed) is given later in Eq. (5). The estimated uncertainty with Eq. (4) underestimates the total error (discussed in Section 3).

Relative error as a function of $\hat{n}$ is shown in Fig. 2 for a nominal 10-fold dilution ratio ($\alpha = 10$) and for 100% plating ($\alpha_p = 1$) as a function of colony size (2 to 10 mm diameter) on a 10 cm diameter agar plate. The decreasing part of the curves (e.g., $\hat{n}<100$ for $d_{\mathrm{colony}} = 3$mm) in the figure represents the effect of the decrease of the sampling error as a function of increased $\hat{n}$, whereas the increasing portion of the curves (e.g., for $\hat{n}>100$, $d_{\mathrm{colony}} = 3$mm) represents the effect of counting error that increases with $\hat{n}$. The location of the minimum of the curves depends on the trade-off between sampling and counting errors. The shape of the error curve shows that as the colony size increases it is advantageous to select a plate with smaller number of colonies and that the optimal counts (for a minimum error) is contained in a narrower range of counts. Observing the behavior of the analytical TE_j curve as a function of k and comparing to the expected behavior of the minimum location (as a function of sampling error and colony size) allows us to determine a value for k.

If the total dilution factor $\alpha\alpha_p$ is decreased, then more dilutions steps will be necessary to obtain a countable plate, and the sampling error at a given count $\hat{n}$ will increase (due to compound error of the multiple steps), thereby pushing the location of the minimum of the total error curve to larger values (to mitigate the compound error). Thus, it is advantageous to sacrifice an increase of counting error to mitigate the fact that sampling error increased by seeking a serial dilution plate with more colonies. For example, consider a scenario with $n_0 = 10^6$ and assume that $n_j = 100$ is the optimal count in a serial dilution experiment with $\alpha = 10$ and $\alpha_p = 1$. Four sequential dilutions ($j = 4$) are needed in order to obtain $n_j \cong 100$. For $\alpha = 2$, thirteen dilutions ($j = 13$) are needed to obtain $n_j \cong 100$ and the compounded sampling error is larger. Thus, the optimal count for $\alpha = 2$ will occur at a value larger than 100.

If the colony size decreases relative to the plate size, it is less likely for two colonies to merge, and counting error at a given count $\hat{n}$ will have decreased, again pushing the location of the minimum of the total error curve to larger values (i.e., we are willing to sacrifice some of the benefit of the reduced counting error in order to reduce sampling error by seeking a serial dilution plate with more colonies).

Since the location of the minimum of TE_j increases with k (see Eq. (6), below), k should be inversely proportional to relative colony size, and inversely proportional to the combined dilution factor $\alpha\alpha_p$.

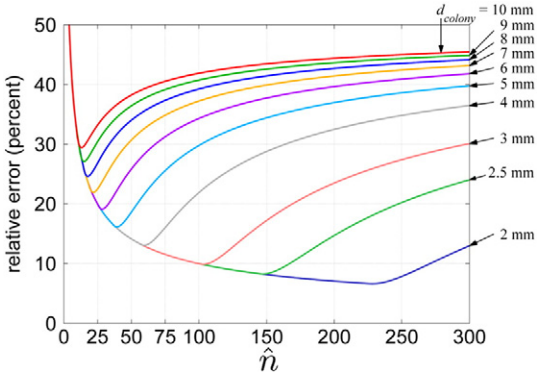

Fig. 2. Relative error curve in percent (Eq. (4)) as a function of measured number of colonies $\hat{n}$ on a plate for different values of colony size (diameter). The experimental condition for the TE curve is $d_{plate} = 10$ cm (diameter), dilution ratios $\alpha = 10$, and $\alpha_p = 1$. The optimal measured $\hat{n}$ that minimizes the total error is the one that is closest to the minimum of the curve (e.g., around $\hat{n} \cong 100$ for $d_{colony} = 3$mm). Eq. (6) predicts the value of $\hat{n}$ at the minimum of the total error curve (e.g., $\hat{n} = 101$ for $d_{colony} = 3$mm). Microbial count in the range $25 < \hat{n} < 170$ is expected to produce an estimate of n_0 within 20% error for $d_{colony} = 3$mm. The typically cited optimal range of 25–250 counts is predicted by our model for a colony of $d_{colony} = 2.5$mm (also within 20% error).

Therefore, the previously undetermined value of k in (Eq. (1)) is set to

$$\left\{\begin{array}{l} 1<k=3\times k_{\text{experiment}} \\ \frac{1}{3}<k_{\text{experiment}}=\frac{1}{\alpha\alpha_p}\times\sqrt{\frac{area_{\text{plate}}}{area_{\text{colony}}}}\cong\frac{1}{\alpha\alpha_p}\times\frac{d_{\text{plate}}}{d_{\text{colony}}} \end{array}\right\} \quad (5)$$

where the proportionally constant is set to be 3 (the value of 3 gave us good empirical results in our extensive testing, especially for $\alpha\alpha_p = 10$), and the inequality $\frac{1}{3} < k_{\text{experiment}}$ is to ensure that $k > 1$ so that $\delta_j < \mu_j$ (see Fig. 1). In most experiments the colony sizes are mono-dispersed and have a circular shape, hence a ratio of diameters (plate to colony) can be used for k.

We seek the location $\hat{n}$ of the minimum of the total error curve for $TE(\hat{n},k)$ in Eq. (3), where $\hat{\delta}_j$ and k are given by Eqs. (1) and (5). Although this is a constrained minimization problem subject to $k \le \sqrt{\hat{n}}$, the constraint does not need to be enforced, and it is sufficient to minimize Eq. (3) ignoring the constraint.

From Eq. (2) we use $\alpha^j\alpha_p = \frac{\hat{n}_0}{\hat{n}_j+\hat{\delta}_j} \cong \frac{\hat{n}_0}{\hat{n}_j}$ (i.e., we neglect δ_j in the denominator) in Eq. (3) to obtain a continuous function of the measured $\hat{n}_j$. We then, by taking a derivative with respect to $\hat{n}_j$ and solving for the derivative to be zero, obtain the location $\hat{n}$ of the minimum of the total error curve. This approximation gives us a nice compact solution with less than 13 counts error (when $\hat{\delta}_j$ is not neglected) in the location of optimal $\hat{n}$ for all possible $k > 1$ (less than 4 counts error for $k > 2$, and less than 10% error for $k > 4$). Our solution for the optimal $\hat{n}$ that produce a minimum total error is given by,

$$\left\{\begin{array}{l} \hat{n}=\frac{4+k^2+k^4+\sqrt{16+8k^2+2k^6+k^8}}{2k^2} \\ \hat{n}\cong k^2 \quad for \quad k>4.5 \\ 1<k=3k_{\text{experiment}} \end{array}\right\}. \quad (6)$$

The solution is *only* a function of k. Eq. (6) shows that the optimal $\hat{n}$ increases with k (as was discussed earlier). For $k > 4.5$, $\hat{n} \cong k^2$ is within 5% of the analytic solution given by Eq. (6), and within 15% of the more complicated solution that can be derived using the correct expression for $\alpha^j\alpha_p$. The estimate of the optimal count $\hat{n}$ can be used to narrow the search for the best dilution plate $\hat{n}_j$ and save time.

In principle $TE(\hat{n},k)$ (Eq. (3)) can be solved for a given value of $TE = const$ to produce a range $\hat{n}_{\min} < \hat{n} < \hat{n}_{\max}$ for which the total error is bounded (i.e., $\hat{n}_{\min}$ and $\hat{n}_{\max}$ are given by the intersection between the curves given in Fig. 2 and a horizontal line at a constant relative error level). This range can be obtained pictorially from Fig. 2 for the specific serial dilution experiment (i.e., $\alpha = 10$, $\alpha_p = 1$, $d_{\text{plate}} = 10$cm), for other serial dilution parameters the figure $RE(\hat{n})$ must be redrawn (since k is a function of α, α_p, d_{colony}, d_{plate} — see Eq. (5) — and $\hat{\delta}(\hat{n})$ is a function of k through Eq. (1)). For example, in Fig. 2, counts in the range $25 < \hat{n} < 170$ is expected to produce an estimate of n_0 within 20% error for $d_{\text{colony}} = 3$mm. The typically cited optimal range of 25–250 counts on a plate that is used to estimate n_0 is predicted by our model (for $d_{\text{plate}} = 10$cm and the commonly used 10-fold dilution with $\alpha_p = 1$) for a colony size of $d_{\text{colony}} = 2.5$mm (i.e., $\frac{d_{\text{plate}}}{d_{\text{colony}}} = 40$, $k = 12$) to within error bound of 20% in estimating $\hat{n}_0$. In Fig. 3 we show the best $\hat{n}$ on a plate (that minimize the total error in Eq. (3)) as a function of the experimental scenario for a given ratio of plate size to colony size ($d_{\text{plate}}/d_{\text{colony}}$) as is predicted by Eq. (6) for 10-fold dilution ($\alpha = 10$) and plating dilution $\alpha_p = 1$ (i.e., plating the entire dilution volume). The figure shows that the optimal count $\hat{n}$ increases with $d_{\text{plate}}/d_{\text{colony}}$, where the likelihood of miscounting colonies due to overlapping and merging together increases as $d_{\text{plate}}/d_{\text{colony}}$ decreases. Fig. 3 can be used as a quick guideline for selecting the best available agar plate from which to estimate n_0.

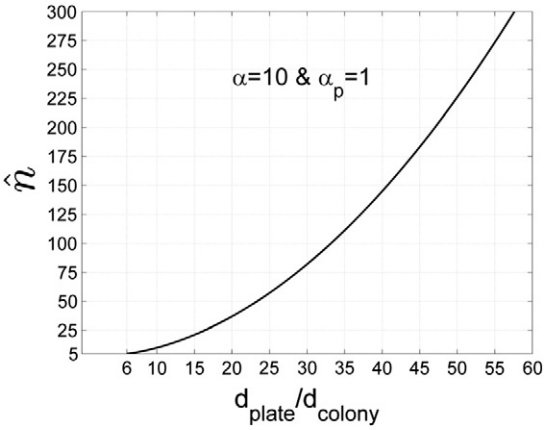

Fig. 3. Optimal counts $\hat{n}$ on a plate (that minimizes the total error in Eq. (3)) as a function of the ratio $d_{\text{plate}}/d_{\text{colony}}$ in a serial dilution process with 10-fold dilution ($\alpha = 10$) and 100% plating ($\alpha_p = 1$). The likelihood of miscounting colonies due to overcrowding increases as $d_{\text{plate}}/d_{\text{colony}}$ decreases.

The inverse of Eq. (6) gives an estimate of the ratio $\frac{d_{\text{plate}}}{d_{\text{colony}}}$ as a function of optimal $\hat{n}$ (i.e., the counts that minimizes TE) on an agar plate, as

$$\frac{d_{\text{plate}}}{d_{\text{colony}}}=\frac{\alpha\alpha_p}{6\sqrt{2}}\sqrt{4\hat{n}-4+2\sqrt{4\hat{n}^2-20\hat{n}+9}+\frac{9+3\sqrt{4\hat{n}^2-20\hat{n}+9}}{\hat{n}}}. \quad (7)$$

The size of a colony (d_{colony}) increases as a function of time (through growth) on the agar plate. Let us assume that one wants to count a given number of colonies $\hat{n}$ on a plate (say, $\hat{n} = 50$). The minimum total error for counting $\hat{n} = 50$ on the jth plate will be achieved when d_{colony} satisfies the ratio $\frac{d_{\text{plate}}}{d_{\text{colony}}}$ that is predicted with Eq. (7). Thus, if the rate of growth of a particular organism is known (say, τ mm/day), and a particular $\hat{n}$ is desired a priori, the time for incubation in the serial dilution experiment can be predicted by d_{colony}/τ where d_{colony} is computed with Eq. (7) for a given desired value of $\hat{n}$. Thus, if a plate with $\hat{n} = 50$ is observed in the serial dilution experiment it is assured to produce minimum total error.

The estimation algorithm for $\hat{n}_0$ from measured $\hat{n}$ on a single agar plate is summarized by the following six steps (knowledge of the error in sampling the aliquot and diluent volumes is not required):

Step 1: Compute $k_{\text{experiment}} = \frac{1}{\alpha\alpha_p}\times\sqrt{\frac{area_{\text{plate}}}{area_{\text{colony}}}} \cong \frac{1}{\alpha\alpha_p}\times\frac{d_{\text{plate}}}{d_{\text{colony}}}$ (see Eq. (5)) for the serial dilution experiment, where α is the dilution factor (e.g., for 10-fold sequential dilution, $\alpha = 10$), α_p is the plating dilution ratio when a fraction of the dilution volume is placed on an agar plate (e.g., when 5% of the dilution volume is plated, $\alpha_p = 20$), and the ratio of plate size to colony size $\sqrt{\frac{area_{\text{plate}}}{area_{\text{colony}}}} \cong \frac{d_{\text{plate}}}{d_{\text{colony}}}$ is estimated from the agar plate. If $k_{\text{experiment}} > \frac{1}{3}$, proceed to step 2. Our algorithm is only valid for $k = 3k_{\text{experiment}} > 1$.

Step 2: Count the number of colonies $\hat{n}_j$ on the J serial dilution plates. The relevant plates are those with $\hat{n}_j$ that are close to $\hat{n}$ (given with Eq. (6), and in Fig. 3 for the standard dilution $\alpha = 10$ and $\alpha_p = 1$). Estimating the best $\hat{n}$ with Eq. (6) may eliminate the need to count irrelevant serial dilution plates that are clearly far from the optimal $\hat{n}$ and save time.

Step 3: Compute $\hat{\delta}_j$ with Eq. (1). Note that k in Eq. (1) is $k = 3k_{\text{experiment}}$. Thus, for $\hat{n}_j < 9k^2_{\text{experiment}}$ the model predicts negligible counting error and thus $\hat{\delta}_j \to 0$.

Step 4: Compute total error TE_j for each plate with Eq. (3). Select the jth plate for which the error is the smallest. For this optimal value of j we now have $\hat{n}_j$ and $\hat{\delta}_j$. This agar plate is defined as the

best (optimal) plate from which to estimate $\hat{n}_0$.

Step 5: Estimate the sought after number of CFUs $\hat{n}_0$ with Eq. (2).

Step 6: Estimate the uncertainty (error) in the estimated solution $\hat{n}_0$ with Eq. (4). The estimated uncertainty underestimates the true uncertainty.

One can compute TE_j, RE_j, and candidate solutions $\hat{n}_0$ with Eqs. (3) and (4) for any (or all) of the J serial dilutions plates and create a table with $\{\hat{n}_j, \hat{\delta}_j, TE_j, \hat{n}_0, RE_j\}$ for each plate, and scrutinize the different candidate solutions for which the optimal solution with the smallest error. Instead of forming a table and selecting the $\hat{n}_j$ with the smallest TE, one could simply select the value of $\hat{n}_j$ closest to the optimal value given by $\hat{n}$ in Eq. (6), but this could lead to a suboptimal choice because of the asymmetry of the total error shown in Fig. 2 (e.g., where for $d_{\text{colony}} = 3$mm the error increases faster for $\hat{n} < 100$ than for $\hat{n} > 100$).

3. Results

A series of simulations to demonstrate our method for estimating n_0 from serial dilution experiments is presented. In the simulation, 10,000 realizations were constructed for each agar plate of the dilution experiment. In all the simulations $\alpha_p = 1$ is assumed (the entire volume of each dilution is plated on an agar plate). We assume that only $\hat{n}_j$ plates with less than 500 colonies are to be counted (a reasonable upper bound in many practical scenarios). In the simulations, volumetric errors in aliquot volume x and diluent volume y (e.g., inserting $x = 0.1$ml into $y = 0.9$ml for 10-fold dilution, $\alpha = 10$) are modeled with normal statistics with standard deviation given in percent error (denoted as "dilution error" in subsequent figures). The number of bacteria (n_j) transferred at each dilution stage is sampled with binomial statistics where $\alpha^{-1} = x/(x+y)$ is the likelihood for a bacterium to be transferred; the Poisson distribution is a special case of the binomial distribution (Forbes et al., 2011).

The condition for resolving two colonies on a plate is selected to be the Rayleigh criterion of resolution in optics (Born and Wolf, 1980) for which in the context of the serial dilution problem, two overlapping colonies each with a diameter d_{colony} are determined to be "just resolved" when the distance between the centers of the two colonies is larger than $0.5d_{\text{colony}}$. The number of miscounted colonies δ_j is simulated with Matern-II hard-core Poisson point-process (Chiu et al., 2013, Ch. 5.4), where the hard-core radius (r_{hardcore}) is set to be the Rayleigh criterion of resolution $0.5d_{\text{colony}}$, is given by

$$\left\{\begin{array}{l}\delta_j = n_j(1-p_{\text{retain}}) \\ p_{\text{retain}} = \dfrac{d_{\text{plate}}^2}{4n_j r_{\text{hardcore}}^2}\left[1-\exp\left(-\dfrac{4n_j r_{\text{hardcore}}^2}{d_{\text{plate}}^2}\right)\right] \\ r_{\text{hardcore}} = \dfrac{d_{\text{colony}}}{2}\end{array}\right\} \tag{8}$$

Table 1
Example of the 6-step estimation algorithm for an unknown sample with $n_0 = 10^6$counts to be estimated by a serial dilution method with plate diameter of $d_{\text{plate}} = 10$cm, colony diameter of $d_{\text{colony}} = 3$mm, and dilution ratios $\alpha = 10$ and $\alpha_p = 1$. $\hat{\delta}_j$ is computed with Eq. (1) and k is computed with Eq. (5); The estimated total error (TE) is computed with Eq. (3), the solution $\hat{n}_0$ and its relative error (RE) are computed with Eq. (4). The 4th dilution plate exhibits the smallest total error (i.e., optimal, best plate) and produces an estimate of $\hat{n}_0 = 1.2 \times 10^6$. The predicted optimal $\hat{n}_j$ with Eq. (6) is 100 and thus also points to the selection of the 4th plate as the optimum.

Dilution stage j	$\hat{n}_j$ (counts)	$\hat{\delta}_j$ (counts)	TE_j (counts)	$\hat{n}_0$ (counts)	RE (%) predicted
3	469	252	0.254×10^6	0.72×10^6	35
4	113	7	0.130×10^6	1.2×10^6	11
5	8	0	0.282×10^6	0.80×10^6	35
6	2	0	0.414×10^6	2.0×10^6	71

where p_{retain} is the probability for a single colony to be observed (i.e., to be counted), and $1 - p_{\text{retain}}$ is the probability that a single colony is miscounted. Two colonies whose centers are located within an area $\pi r_{\text{hardcore}}^2$ fuse together and appear as one colony. In Eq. (8) $\delta_j \to 0$ when $\frac{d_{\text{colony}}}{d_{\text{plate}}} \to 0$ (i.e., the likelihood of overlapping colonies approaches zero). Using δ_j of Eq. (8), the observed (measured) counts are given by $\hat{n}_j = n_j - \delta_j$.

The details of the extensive and complex statistical model for the simulations are deferred to a separate publication. That said, we want to give the reader some insights to the overall governing principles behind the simulations. The simulation of n_j is consistent with a compound binomial–binomial probability density function, where the compound binomial–binomial distribution describes a cascade serial dilution $n_0 \to n_{j=1} \to n_{j=2} \to n_j \to \ldots \to n_j$; $n_0 \to n_{j=1}$ is a simple binomial process, $n_{j=1} \sim \text{Binomial}\left(n_0, \frac{1}{\alpha\alpha_p}\right)$, and each of the intermediates n_j in the J cascade process is a random variable, leading to a compound binomial–binomial distribution that is itself a binomial $n_j \sim \text{Binomial}\left(n_0, \frac{1}{\alpha_p}\prod_{i=1}^{j}\frac{1}{\alpha_i}\right)$, where each $\alpha_i = \frac{x+y}{x}$. The binomial–binomial distribution describes the serial dilution process in the absence of dilution errors in the aliquot x and diluent y volumes. The randomness in α (due to errors in x and y at each dilution stage, hence $\alpha_i = \frac{x_i+y_i}{x_i}$) is combined with the binomial–binomial in the Monte-Carlo simulations to include volumetric errors. It is beneficial to the reader to simply think of $n_j = n_0\alpha^{-j}\alpha_p^{-1}$ (at the expense of accuracy) as 10,000 samples (realizations) drawn from a Poisson probability density function with intensity λ (the Poisson population parameter), $\lambda = n_0\alpha^{-j}\alpha_p^{-1}$.

We vary four factors in the simulations: dilution error of 0 to 10% at each dilution stage (for both the aliquot and diluent volumes), n_0 in the range 10^4 to 10^{12}, dilution ratio α in the range 2 to 100, and $\frac{d_{\text{plate}}}{d_{\text{colony}}}$ (6.25 to 200). Thus, a large portion of a 4-dimensional space relevant to the method is explored. In this work we use the simulation results to verify and test our proposed estimation algorithm

The parameters of merit for testing the estimation algorithm are accuracy, precision, and total error. Statistical properties were computed from 10,000 $\hat{n}_0$ estimations where each of the 10,000 $\hat{n}_j$ measurements was processed as a single plate with our algorithm. Standard definitions for accuracy, precision, and total error are used. Accuracy is the distance from the truth, accuracy $= E(|\hat{n}_0 - n_0|)$. Precision is the standard deviation of the solution $\hat{n}_0$ around its mean value $E(\hat{n}_0)$, precision $= STD(\hat{n}_0)$. We strive to obtain small values for both accuracy and precision. The true total error computed by $TE = \sqrt{\text{accuracy}^2 + \text{precision}^2} = \sqrt{E\left[(\hat{n}_0 - n_0)^2\right]}$ and is a function of n_0. These three figures of merit are presented in the figures below as percentages by normalizing to the true n_0 (which is *only* known in simulations). The predicted relative error RE (Eq. (4)) is also presented. The prediction of total error (and the relative error) is solely dependent on the solution $\hat{n}_0$ without any knowledge of the true value for n_0. The predicted total error may underestimate the truth by as much (in an extreme case) as five-fold (e.g., predicting 10% total error where the true total error is 50%), nevertheless, the prediction is still useful and is within the same order of magnitude as the truth. The median and mean values of the underestimation in our simulations are within factor of two of the true values.

An example of a single experiment for which the 6-step solution process is implemented is given in Table 1. In this example $n_0 = 10^6$ counts, and the parameters of the serial dilution process are plate diameter of $d_{\text{plate}} = 10$cm, colony diameter of $d_{\text{colony}} = 3$mm, dilution ratios $\alpha = 10$, and 2% dilution errors. The table shows that the best plate to use is the 4th dilution assay for which the total error and the predicted relative error (Eqs. (3) and (4)) are minimized. The 4th dilution, using Eq. (2) estimates the unknown sample to be $\hat{n}_0 = 1.2 \times 10^6$ counts.

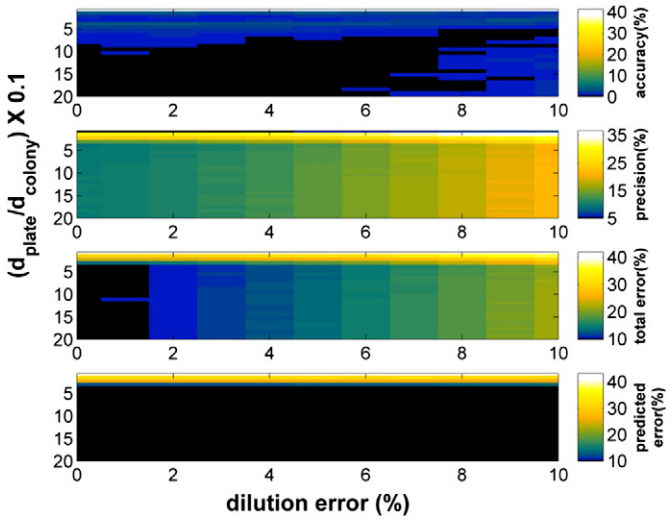

Fig. 4. Performance of estimation method (accuracy, precision, total error, and predicted total error) for a low initial count $n_0 = 10^4$, a nominal 10-fold dilution ratio, $\alpha = 10$, and $\alpha_p = 1$. Performance is a function of the error (0 to 10%) in each stage of the dilution process for both the aliquot and diluent volumes, and the ratio between the plate-diameter and the colony-diameter (6.25 to 200). The figure shows that the estimation method performs with error that is much less than $\pm$ 0.1 $\log_{10}$ units (i.e., 100%). Most accuracy and precision values fall below 30%. The predicted total error for which the true n_0 is not known underestimates the true total error (computed with knowledge of n_0) by a factor of 1.25 (median value). The color-scale ranges from black (lowest value) to white (highest value).

For a given series of dilutions, it is possible that the counts associated with a plate other than the optimum will (by sheer luck) give an estimated $\hat{n}_0$ closer to the truth than the optimal plate. For example, if $\hat{n}_{j=6}$ was equal to 1 instead of 2, the predicted value of $\hat{n}_0$ would be exactly the correct value of 1×10^6. On average, however, the optimal (best) plate ($j = 4$) gives a better estimate due to a much smaller relative error (11% versus 71%).

In this example, using the predicted optimal $\hat{n} = 100$ with Eq. (6), or Fig. 3, could have directly pointed to the 4th plate and eliminated the need to count the other plates. The statistics of the 4th plate over all the 10,000 simulations gives a mean count $E(n_{j=4}) = 96$ with a relative error of 12%. Ninety percent of the time (i.e., 0.05 confidence limit) selecting the 4th plate as the best plate results in true relative errors (for which the true n_0 must be known) between -23% and $+21\%$.

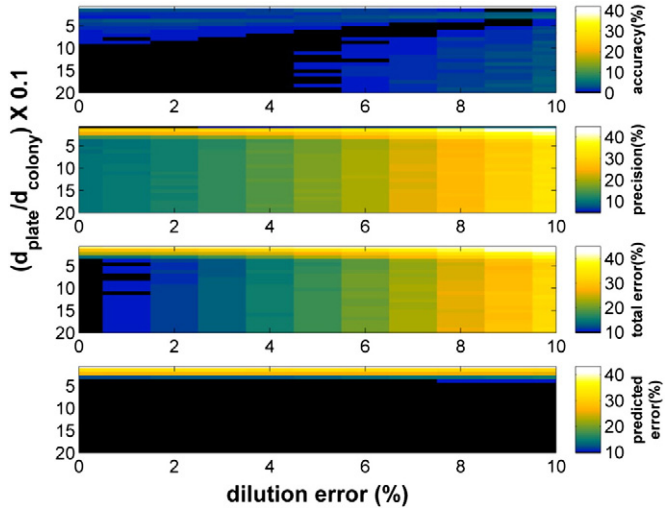

Fig. 5. Same as Fig. 4 but for a medium initial count of $n_0 = 10^7$. The performance is less good than shown in Fig. 4, but still overall is much better than $\pm$ 0.1 $\log_{10}$ units.

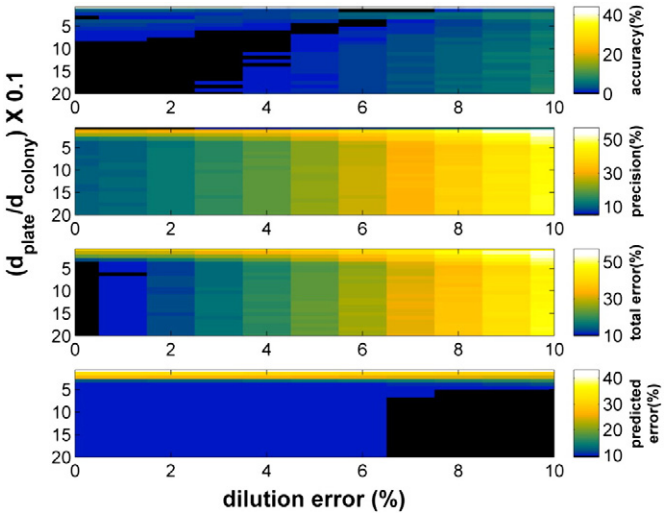

Fig. 6. Same as Fig. 4 but for a high initial count of $n_0 = 10^{12}$. The figure shows that the performance is similar to that shown in Figs. 4 and 5, and is still within $\pm$ 0.1 $\log_{10}$ in spite of the fact that to reach a countable plate (e.g., $\hat{n}_j$<500) from the initial value $n_0 = 10^{12}$ about 10 sequential dilutions are needed, for which the error in the dilution volumes (aliquot and diluent) is compounded.

Our predicted relative error (11%), with Eq. (4) underestimates the true relative errors by two-fold.

In Figs. 4 to 6 the performance of the estimation method is shown for three scenarios: a low initial count ($n_0 = 10^4$ in Fig. 4), medium initial count ($n_0 = 10^7$ in Fig. 5), and high initial count ($n_0 = 10^{12}$ in Fig. 6). In all figures the color-scale ranges from black (lowest value) to white (highest value). For these scenarios the dilution ratio is held at a nominal value $\alpha = 10$. The performance is presented in color images as a function of error in the dilution process (0 to 10%) and the ratio $d_{\text{plate}}/d_{\text{colony}}$ (6.25 to 200). Fig. 4 shows that the accuracy (top panel) in most cases is much smaller than 10% for this wide range of parameters; the precision in most areas of the plot is very good, less than 30%, and the total error is less than about 40%. The predicted total error (lower panel) is smaller (by factors 1.25 and 1.4 for the median and

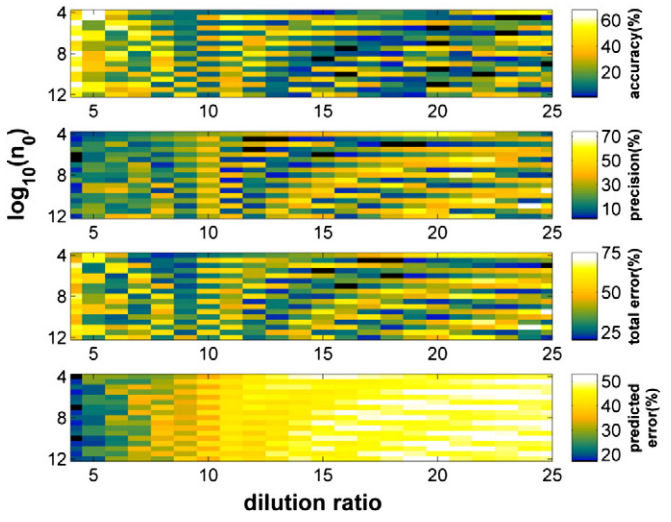

Fig. 7. Performance of estimation method (accuracy, precision, total error, and predicted total error) as a function of dilution ratio α, and n_0, for a small ratio $d_{\text{plate}}/d_{\text{colony}} = 10$ for which the likelihood of counting error due to overlapping of colonies is high. The plating dilution is kept at $\alpha_p = 1$. The error in the dilution process (for both the aliquot and diluent volumes) is held constant at a level of 5%. The maximum value of α that satisfies $1 < k = 3 \times k_{experiment}$ (Eq. (5)) is 30, the upper limit for which our method is valid. The figure shows that the estimation method performs with error less than $\pm$ 0.1 $\log_{10}$ units (i.e., 100%). The color-scale ranges from black (lowest value) to white (highest value).

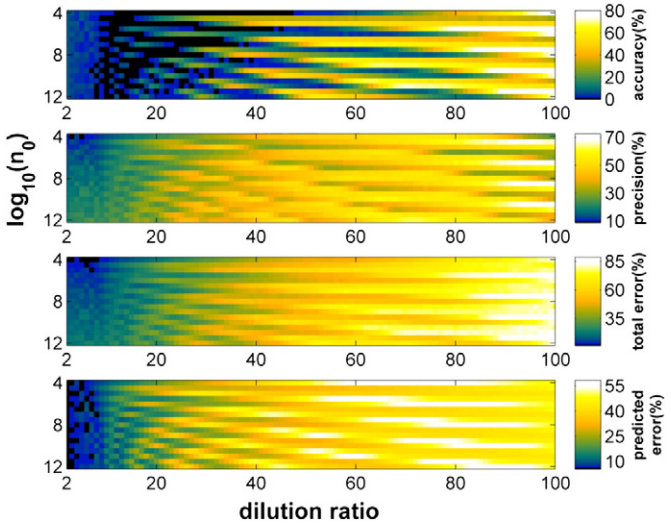

Fig. 8. Same as Fig. 7 but for a larger $d_{plate}/d_{colony} = 50$ for which the likelihood of overlapping colonies is lower. The estimation method performs with error less than $\pm$ 0.1 $\log_{10}$ units.

mean values, respectively) than the true error (2nd panel from the bottom), as was anticipated earlier in the discussion of Eq. (3). In Fig. 5 we show results for a medium initial count $n_0 = 10^7$ and in Fig. 6 we show results for a high initial count $n_0 = 10^{12}$. The performance for these higher counts is less good (slightly) than the one shown in Fig. 4. The results show that our estimation method works well and is within $\pm$ 0.1 $\log_{10}$ units (i.e., within 100%), even for a high count where the number of dilutions needed to bring the measured $\hat{n}_j$ to a reasonable number that can be counted (say, less than 500) is large, and thus the compounded error in the dilution process is large. The predicted total error in Figs. 5 and 6 underestimate the true error by factor of 1.5 and 2 for median values, and factors 1.8 and 2.4 for mean values.

In Figs. 7 to 9 the performance of the estimation method is explored as a function of the dilution ratio α that ranges from 2 to 100. For a given dilution ratio α, one only has a few discrete $\hat{n}_j$ values, of which the closest to the optimal $\hat{n}$ (Eq. (6)) will produce the smallest total error. As α decreases (e.g., from 10-fold dilution to 2-fold dilution) the number of plates increases and the likelihood that a measured $\hat{n}_j$ is closer to the optimal $\hat{n}$ increases. For simplicity, the error in the dilution process (for each dilution stage) is kept at a constant level of 5%. We

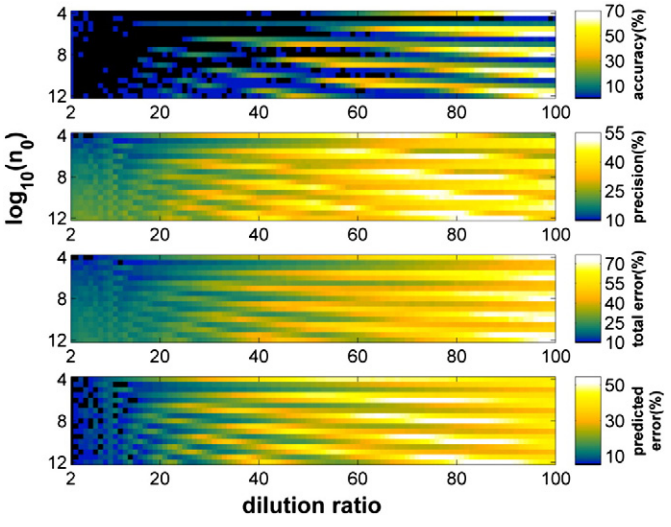

Fig. 9. Same as Fig. 7 but for a higher ratio $d_{plate}/d_{colony} = 100$ that further reduces the likelihood of overlapping colonies. The figure shows that the estimation method performs with error less than $\pm$ 0.1 $\log_{10}$ units.

show results for three d_{plate}/d_{colony} ratios (low, medium, and high) that affect the likelihood of counting error. A low ratio of d_{plate}/d_{colony} implies high likelihood of counting error (large colonies that can fuse and overlap). For a given n_0 and a dilution stage j, a plate with high value of d_{plate}/d_{colony} is more sparsely covered than a plate with a low value d_{plate}/d_{colony}. As d_{plate}/d_{colony} increases (for a fixed n_0) each colony becomes a separate small identifiable dot on the plate, and the counting error is very small. In Fig. 7 we show the method's performance for a small ratio $d_{plate}/d_{colony} = 10$ (e.g., a 10 cm plate and a very large 1 cm colony, hence, a very high likelihood for counting error). Due to the limitation of our estimation method (Eq. (1)) that $k = 3 \times k_{experiment}$ must be greater than one, $k_{experiment} = 10/\alpha$ is limited by $\alpha < 30$ and only results for $\alpha < 25$ are shown. The figure shows that for a very large range $10^4 \le n_0 \le 10^{12}$ the performance is within $\pm$ 0.1 $\log_{10}$ units. In Figs. 8 and 9 we show the performance for d_{plate}/d_{colony} of 50 and 100 respectively, for which the likelihood of counting error is reduced. The figures show that the estimation method performs with error much less than $\pm$ 0.1 $\log_{10}$ units. For these larger ratios we can study the performance for a larger α (since the constraint that $k > 1$ is less restrictive); we show results for $2 < \alpha < 100$. In Fig. 7 (for which the likelihood of counting error is large due to the small value of d_{plate}/d_{colony}) the predicted total error is, on average, the same as the true total error (median value for ratio of true total error to predicted total error of 0.9, and mean value of 1.1). The predicted total error in Figs. 8 and 9 underestimates the true error by factor 1.6 and 1.5 for median values, and 1.6 and 1.6 for mean values.

4. Conclusions

We introduced a simple method for estimating the undiluted initial counts n_0 from observed counts of a serial dilution experiment which depends on the serial dilution factor α, the plating dilution factor α_p, and the dilution errors. The dilution process is given by $n_j = n_0 \alpha^{-j} \alpha_p^{-1}$ where, for example, $\alpha = 10$ is for 10-fold sequential dilution, and $\alpha_p = 20$ is for a scenario that 5% of the dilution volume is plated. Our method predicts the best (optimal) $\hat{n}_j$ agar plate with which to estimate $\hat{n}_0$ counts with the smallest error. Knowing the aliquot volume, one can easily convert counts to concentration (for example CFU/ml).

The number (density) of viable microbial organisms is estimated from a single dilution plate (assay), hence, no replicates are needed. The procedure is outlined as a 6-step process. In a serial dilution experiment for microorganisms that can grow on bacteriological media and develop into colonies, the main factors that affect the process are area of agar (dilution plate) in which the colonies grow and the size of colonies being studied. In addition the dilution factor α affects the compound volumetric dilution error in the j dilution stages that are required to reduce the unknown sample n_0 (e.g., 10^6 organisms) to a countable number $\hat{n}_j$ (e.g., usually less than a few hundreds). Our method selects the best dilution plate with which to estimate the microbial counts, and takes into account the colony size and plate area that both contribute to the likelihood of miscounting the number of colonies on an agar plate. The only two input parameters in our estimation method are (*i*) the ratio of a plate area to a typical colony area $\sqrt{\frac{area_{plate}}{area_{colony}}} \cong \frac{d_{plate}}{d_{colony}}$ which assumes that the bacterial colonies on each of the j plates are roughly of the same size and are roughly of a circular shape, and (*ii*) the nominal dilution ratios α and α_p for the dilution and plating process. Knowledge of the error in the dilution process (in sampling the aliquot and diluent volumes) is not required. It is interesting to note that the guidelines (e.g., Tomasiewicz et al., 1980; FDA, 2001) for selecting the optimal dilution plate (e.g., 30 to 300 counts or 25 to 250 counts) do not specify the area in which the colonies grow, nor the diameter of the particular organism assayed. In our estimation method these parameters are addressed.

Our estimation method attempts to choose an optimal measured $\hat{n}_j$ by striking a balance between sampling error and counting error, where relative sampling error increases as j increases (fewer colonies on a plate), and counting error decreases as j increases (less likelihood for colonies to merge and overlap, and thus be miscounted). The combined effect of these two processes is described by the total error curve (Eq. (3)). We assumed that the spread (related to variance) of the true n_j and the measured (estimated) $\hat{n}_j$ are roughly the same, and can be described with a Poisson (and shifted Poisson) statistics for which the variance equals the mean. With this assumption we described the behavior of the bias $\hat{\delta}$ (Eq. (1)), due to miscounting of overlapping colonies, that is introduced in predicting the true n_j as $n_j = \hat{n}_j + \hat{\delta}_j$. With the selection of the optimal dilution stage, j, and its associated values of $\hat{n}_j$ and $\hat{\delta}_j$, we estimate n_0 (Eq. (2)) and predict the error of the estimation (Eq. (4)). Our predicted error underestimates the true error (by about a factor two in our simulations). The optimal observed $\hat{n}$ (that produces a minimum total error) increases with the parameter $k_{\text{experiment}} = \frac{1}{\alpha\alpha_p} \times \frac{d_{\text{plate}}}{d_{\text{colony}}}$ which characterizes the serial dilution experiment. The behavior of the optimal $\hat{n}$ with $k_{\text{experiment}}$ is qualitatively correct for a dilution process which suffers from sampling and counting errors, in addition to compounding dilution errors from all the j dilution stages prior to the optimal dilution. An analytical estimate of the optimal $\hat{n}$, given in Eq. (6) as a function of $k_{\text{experiment}}$, can be used to narrow the search for the optimal $\hat{n}_j$ among all the dilution plates and save time. An optimal value for $\hat{n}_j$ for a standard 10-fold dilution as a function of $d_{\text{plate}}/d_{\text{colony}}$ is shown in Fig. 3 and can give guidance for the best (optimal) agar plate to be used in estimating n_0 in our 6-step estimation process. The shape of the error curve (Fig. 2) and the optimal counts (Fig. 3) show that as the colony size increases it is advantageous to select a plate with smaller number of colonies, and that the optimal counts (for a minimum error) is contained in a narrower range of counts. An analytical solution of $TE = const$ (Eq. (3)) can be pursued in order to produce a range $\hat{n}_{\min} < \hat{n} < \hat{n}_{\max}$ for which the total error is bounded. In this work we were only interested in selecting the best single plate with which to estimate n_0. However, the range $\hat{n}_{\min} < \hat{n} < \hat{n}_{\max}$ for a given error can be obtained pictorially from Fig. 2 by the intersection between the curves in Fig. 2 and a horizontal line at $RE = const$. For example, for 10-fold dilution, 100% of the dilution volume is plated, colony diameter of 3 mm, and 10 cm (diameter) agar plate, the microbial count on an agar plate in the range $25 < \hat{n} < 170$ is expected to produce an estimate of n_0 within 20% error. The typically cited optimal range of 25–250 counts on a plate that is used to estimate n_0 is predicted by our model for $\alpha = 10$, $\alpha_p = 1$ and $d_{\text{plate}} =$ 10cm diameter, for a colony diameter of $d_{\text{colony}} = 2.5$mm to within error bound of 20% in estimating $\hat{n}_0$ (Fig. 2). For other serial dilution parameters a new figure can be easily created.

We extensively tested our estimation method through numerical simulations as a function of four parameters: dilution errors in the range of 0 to 10% (for both the aliquot and diluent volumes), n_0 (10^4 counts to 10^{12} counts), dilution ratio α (2 to 100), and $\frac{d_{\text{plate}}}{d_{\text{colony}}}$ (6.25 to 200). Our method produced accuracy, precision, and total error well within $\pm$ 0.1 $\log_{10}$ units (i.e., within 100% error).

Acknowledgments

We are indebted to Dr. Jose-Luis Sagripanti for his guidance and encouragement to pursue this work. His wealth of knowledge and patience is much appreciated.

References

American Public Health Association, 2005. Standard methods for the examination of water and wastewater. American Public Health Association, Washington DC (Available at http://www.standardmethods.org/).
Blodgett, R., 2010. Most probable number from serial dilution. Appendix 2 in U.S. Food and Drug Administration, Bacteriological Analytical Manual (BAM). (Available at http://www.fda.gov/Food/FoodScienceResearch/LaboratoryMethods/ucm109656.htm).
Born, M., Wolf, E., 1980. Principles of Optics, 6th edition. Pergamon, New-York, NY.
Casella, G., Berger, R.L., 2002. Statistical Inference, 2nd edition. Duxbury, Pasific Grove, California.
Chiu, N.K., Stoyan, D., Kendall, W.S., Mecke, J., 2013. Stochastic Geometry and Its Applications, 3rd edition. Wiley, West Sussex, United Kingdom.
Cochran, W.G., 1950. Estimation of bacterial densities by mean of the "Most Probable Number". Biometrics 6, 105–116.
FDA, 2001. Bacteriological analytical manual. Available at http://www.fda.gov/Food/FoodScienceResearch/LaboratoryMethods/ucm063346.htm.
Forbes, C., Evans, M., Hastings, N., Peacock, B., 2011. Statistical Distributions, 4th edition. Wiley, New-York, NY.
Goldman, E., Green, L.H., 2008. Practical Handbook of Microbiology, 2nd edition. CRC Press, Boca Baton, Florida.
Hedges, A.J., 2002. Estimating the precision of serial dilutions and viable bacterial counts. Int. J. Food Microbiol. 76, 207–214.
Hollinger, F.B. (Ed.), 1993. AIDS Clinical Trials Group Virology Manual for HIV Laboratories. NIH (NIAID Division of AIDS), Rockville MD, pp. MIC-1–MIC-5.
Jennison, M.W., Wadsworth, G.P., 1940. Evaluation of the errors involved in estimating bacterial numbers by the plating method. J. Bacteriol. 39, 389–397.
Koch, R., 1883. Über die neuen Untersuchungsmethoden zum Nachweis der Mikrokosmen in Boden, Luft und Wasser. Vortrag auf dem XI. Deutschen Ärztetag in Berlin. Vereinsblatt für Deutschland, Komnüssions-Verlag von F. C. W. Vogel, Leipzig, pp. 137,274–137,284 (Available at http://edoc.rki.de/documents/rk/508-274-284/PDF/274-284.pdf).
Lin, L.I., Stephenson, W.R., 1998. Validating an assay of viral contamination. In: Peck, R., Haugh, L.D., Goodman, A. (Eds.), Statistical Case Studies: A Collaboration Between Academe and Industry. SIAM, Philadelphia, PA.
Myers, L.E., McQuay, L.J., Hollinger, F.B., 1994. Dilution assay statistics. J. Clin. Microbiol. 32, 732–739.
Oliver, J.D., 2005. The viable but nonculturable state in bacteria. J. Microbiol. 43, 93–100.
Oliver, J.D., 2010. Recent findings on the viable but nonculturable state in pathogenic bacteria. FEMS Microbiol. Rev. 34, 415–425.
Park, W.H., Williams, A.W., 1905. Pathogenic microorganisms. Available at Nabu Press https://archive.org/stream/pathogenicmicroo00parkrich#page/n3/mode/2up.
Scharf, S.L., Friedlander, B., 1994. Matched subspace detectors. IEEE Trans. Signal Process. 42 (8), 2146–2157.
Taswell, C., 1984. Limiting dilution assays for the determination of immunocompetent cell frequencies.III. Validity tests for the single-hit Poisson model. J. Immunol. Methods 72, 29–40.
Tomasiewicz, D.M., Hotchkiss, D.K., Reinbold, G.W., Read, R.B., Hartman, P.A., 1980. The suitable number of colonies on plates for counting. J. Food Prot. 43, 282–286.
Wilson, G.S., 1922. The proportion of viable bacteria in young cultures with special reference to the technique employed in counting. J. Bacteriol. 7, 405–446.