

Research Article

Mathematical Modeling of Biofilm Structures Using COMSTAT Data

Davide Verotta,¹ Janus Haagenen,² Alfred M. Spormann,³ and Katherine Yang¹

¹*Department of Clinical Pharmacy, School of Pharmacy, University of California San Francisco, San Francisco, CA, USA*

²*Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kogle Alle 6, 2970 Hørsholm, Denmark*

³*Department of Civil and Environmental Engineering, James H. Clark Center, Stanford University, Rm E250, 318 Campus Drive, Stanford, CA 94305, USA*

Correspondence should be addressed to Davide Verotta; davide.verotta@ucsf.edu

Received 10 August 2017; Revised 14 November 2017; Accepted 26 November 2017; Published 20 December 2017

Academic Editor: Michele Migliore

Copyright © 2017 Davide Verotta et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Mathematical modeling holds great potential for quantitatively describing biofilm growth in presence or absence of chemical agents used to limit or promote biofilm growth. In this paper, we describe a general mathematical/statistical framework that allows for the characterization of complex data in terms of few parameters and the capability to (i) compare different experiments and exposures to different agents, (ii) test different hypotheses regarding biofilm growth and interaction with different agents, and (iii) simulate arbitrary administrations of agents. The mathematical framework is divided to submodels characterizing biofilm, including new models characterizing live biofilm growth and dead cell accumulation; the interaction with agents inhibiting or stimulating growth; the kinetics of the agents. The statistical framework can take into account measurement and interexperiment variation. We demonstrate the application of (some of) the models using confocal microscopy data obtained using the computer program COMSTAT.

1. Introduction

Biofilms are structured communities of bacteria enclosed in an extracellular matrix composed of polysaccharides, proteins, and extracellular DNA adherent to a surface [1]. Unlike planktonic bacteria, biofilms exhibit differences in metabolism, antibiotic tolerance, and ability to evade the immune system, making infections due to biofilms difficult to treat [2]. Biofilms are a main cause of acute and chronic infections, including foreign-body infections, otitis media, and urinary tract infections.

When a population of microorganisms organized in a biofilm grows, it is likely to pass over different phases. The growth may follow a period of dormancy, if the environmental conditions before the beginning of the growth are not optimal. Eventually, cells start to divide and structure, and the biofilm grows into a period in which the overall rate of cell division prevails over that of their death. Under favorable conditions, the growth may be considered to be unlimited (hence exponential) for some time, but eventually

[3] physiological and physical limits such as (i) exhaustion of available nutrients, (ii) accumulation of inhibitory metabolites or end products, and (iii) exhaustion of space intervene. As a result, the growth rate decreases and the colony reaches its maximum size. Repeated or continual exposure to these environmental or physiological stressors can then result in a decline in the biofilm size [4].

In vitro biofilms are often used in studies concerning therapy, dealing with the reactions of bacterial populations to various agents: drugs, for example, antibiotics, or mutagens [5–8]. Those are applied externally to the biofilm structures and change their environment or directly eliminate (kill) the bacteria or decrease their reproductive capacity.

A variety of mathematical models have been used to describe bacterial growth in investigation of dynamics in environment depending only on the activity of the bacteria [9, 10]. Similarly, a number of models have been proposed to describe the action of different agents, in particular drugs and their interaction [11–13]. The objective of this paper

is to obtain an integrated framework that provides models describing the different stages of biofilm growth, the action of different agents, and the simultaneous modeling of resulting kinetics for live and dead biofilm. The statistical issues associated with the use of these data, in particular the treatment of different sources of variability, are also addressed.

We demonstrate the use of the resulting general models using data obtained by means of confocal microscopy and COMSTAT [14]. COMSTAT takes the image stacks created by the confocal microscope as source data and produces up to ten image analysis features for quantification of biofilm structures which are output as one or more text files. The models we describe in this paper apply to univariate measurements: total biomass, area in a specific layer, average thickness, and volumes of microcolonies identified at the substratum (COMSTAT also obtains multivariate data, such as thickness distribution, which can be used to quantify the three-dimensional structures in the biofilm. The modeling of such data is the subject of current modeling investigation and will be reported in future communications.).

2. Methods

2.1. Mathematical Modeling. Modeling of biofilm growth requires the specification of three components. The first, a function $g(\cdot)$, describes the growth of the biofilm in absence of agents limiting or promoting growth. The second, a nonnegative function $h(\cdot)$, describes the interaction with the agents. The third, $C(\cdot)$, describes the temporal variation (kinetics) of the agents acting on the biofilm. The general model we consider expresses the rate of change of the biofilm as follows:

$$\begin{aligned} \frac{dB(t)}{dt} &= g(t, B(t)) \pm h(C(t), B(t)) \\ B(0) &= B_0, \end{aligned} \quad (1)$$

where t is time, $B(t)$ is biofilm amount present in the system at time t , B_0 is the initial conditions or biofilm amount at $t = 0$, and $C(t)$ is the concentration of agent at time t . The \pm sign indicates that the model can describe either inhibition or stimulation.

2.1.1. Biofilm Growth. A large number of models have been proposed to describe bacterial growth, for example, [15–17]. For the purpose of this paper, we only describe the simplest model for unlimited growth (exponential) and three semiempirical models for limited growth. Exponential biofilm growth assumes that the rate of growth is proportional to the amount of cells present in the system, and that there is no limitation to growth; thus

$$\frac{dB(t)}{dt} = g(t, B(t)) = k_b B(t), \quad (2)$$

where k_b is the growth rate of the biofilm. The analytic solution of (2) is an exponential growth:

$$B(t) = B_0 e^{k_b t} \quad (3)$$

with doubling time, t_d , being equal to $\ln(2)/k_b$.

In general, exponential growth can only describe the early stages of biofilm growth. A number of semiempirical models can be used to take into account the decrease in proliferation that results from limitations of nutrient supply or mechanical constraints or metabolites accumulation. We consider the Logistic, Gompertz, and Bertalanffy models. The Logistic [18] takes the following form:

$$g(t, B(t)) = k_b B(t) \left(1 - \frac{B(t)}{B_{\max}}\right), \quad (4)$$

where B_{\max} is the maximum biofilm level that can be reached. The corresponding analytic solution is as follows:

$$B(t) = \frac{B_{\max} B_0}{B_0 + (B_{\max} - B_0) e^{-k_b t}}. \quad (5)$$

The Gompertz model [19–21] takes the following form:

$$g(t, B(t)) = k_b B(t) \ln\left(\frac{B_{\max}}{B(t)}\right) \quad (6)$$

with analytic solution:

$$B(t) = B_{\max} e^{\ln(B_0/B_{\max}) e^{-k_b t}}. \quad (7)$$

Finally the Bertalanffy model [22] assumes that growth occurs proportionally to biofilm surface area, while biofilm loss is proportional to biofilm:

$$g(t, B(t)) = k_b B(t)^{2/3} - k_d B(t), \quad (8)$$

where k_d is the death rate for biofilm. The solution of (8) is also sigmoidal in shape and tends to an asymptote as time increases, where the birth and death term balance each other out. A general version of the Bertalanffy model takes the following form [22]:

$$g(t, B(t)) = k_b B(t)^\lambda - k_d B(t)^\delta, \quad (9)$$

where $0 < \lambda < \delta$. The Logistic equation is a special case of this model with $\lambda = 1$, $\delta = 2$.

2.1.2. Agents Interaction with Biofilm. The simplest model describing the interaction of bacterial biofilm with an agent assumes that the action of the agent is proportional to the product agent and biofilm:

$$h(C(t), B(t)) = \theta_1 C(t) B(t), \quad (10)$$

where θ_1 now quantifies the interaction. For an inhibitory agent combining (10) with exponential growth (2) obtains

$$\frac{dB(t)}{dt} = k_b B(t) - \theta_1 C(t) B(t). \quad (11)$$

According to this model if the agent concentration is kept at constant level k_b/θ_1 , the growth rate is zero. That concentration is the minimal concentration of agent killing bacteria or BIC. The BIC is frequently used to characterize bacterial growth in planktonic, in vitro experiments, and

clinical settings (e.g., [23]). Note that if the growth rate is not exponential, the BIC is not a constant but depends on the current amount of biofilm. For example, for the Logistic model, incorporating an inhibitory agent action takes the following form:

$$\frac{dB(t)}{dt} = k_b B(t) \left(1 - \frac{B(t)}{B_{\max}}\right) - \theta_1 C(t) B(t) \quad (12)$$

and the BIC is given by the following equation:

$$\text{BIC}(B(t)) = \frac{k_b}{\theta_1} \left(1 - \frac{B(t)}{B_{\max}}\right) \quad (13)$$

that shows how the BIC decreases as a function of $B(t)$. Similarly, for the Gompertz model,

$$\text{BIC}(B(t)) = \frac{k_b}{\theta_1} \ln\left(\frac{B_{\max}}{B(t)}\right). \quad (14)$$

More complex models for the interaction agent/biofilm can be used, for example, threshold models, according to which the agent is effective if its concentration reaches a threshold,

$$h(C(t), B(t)) = \begin{cases} \theta_1 C(t) B(t) & \text{if } C(t) \geq \theta_2 \\ 0 & \text{otherwise,} \end{cases} \quad (15)$$

where the parameter θ_2 is the threshold; or models following saturation kinetics similar to what is used in, for example, pharmacodynamics [13],

$$h(C(t), B(t)) = \frac{\theta_1 C(t)}{\theta_2 + C(t)} B(t), \quad (16)$$

where now, for large concentrations of agent, the killing rate asymptotes to $\theta_1 B(t)$. In presence of two agents, $C_1(t)$ and $C_2(t)$, models (10) and (16) increase in complexity, as the agents can affect the killing rate in different ways. For example, the linear model (10) can be generalized as follows:

$$h(C_1(t), C_2(t), B(t)) = (\theta_1 C_1(t) + \theta_2 C_2(t)) B(t) \quad (17)$$

or

$$\begin{aligned} h(C_1(t), C_2(t), B(t)) \\ = (\theta_1 C_1(t) + \theta_2 C_2(t) + \theta_3 C_1(t) C_2(t)) B(t), \end{aligned} \quad (18)$$

where θ_3 is a killing rate for the agents' interaction. When saturation kinetics are present, as is the case for the single agent model (16), a number of possibilities arise, generating additive, synergistic, and antagonistic models often applied in pharmacodynamic, enzymology, or binding experiments [12]. For example, incorporating a model for competitive interaction between two agents yields

$$\begin{aligned} h(C_1(t), C_2(t), B(t)) \\ = \left(\frac{C_1(t)/C_{1,50} + \alpha (C_2(t)/C_{2,50})}{1 + C_1(t)/C_{1,50} + C_2(t)/C_{2,50}} \right) B(t), \end{aligned} \quad (19)$$

where if $\alpha > 0$ the model shows additivity, and when $\alpha = 0$ it reduces to the competitive antagonism model and yields

$$\begin{aligned} h(C_1(t), C_2(t), B(t)) \\ = \frac{C_1(t)}{C_1(t) + C_{1,50} (1 + C_2(t)/C_{2,50})} B(t). \end{aligned} \quad (20)$$

2.1.3. Agent Kinetics. A mathematical representation of the agents' kinetics is needed to incorporate their action on the biofilm. This in general does not present particular difficulties. Analytical solutions or sets of differential equations can be used to do so [24, 25]. As an alternative one can use "model independent" representations of the data, such as smoothing splines [26], or, when the measurement error in the kinetics data is low an even simpler linear interpolant as we do in the examples reported below.

2.1.4. Modeling Dead Biofilm. Confocal microscopy data allow the measurement of both live and dead cells (see Data section). Indicating by $D(t)$ the dead biofilm present in the system at time t , its rate of generation can be directly obtained from the growth equations reported above by identifying the loss or gain in live cells due to the presence of the agent. For the Logistic growth rate equation (4), this yields

$$\frac{dD(t)}{dt} = k_b \frac{B(t)^2}{B_{\max}} \pm h(C(t), B(t)), \quad (21)$$

and, for the generalized Bertalanffy model equation (8), an inhibitory agent obtains

$$\frac{dD(t)}{dt} = k_d B(t)^\delta + h(C(t), B(t)). \quad (22)$$

For the Gompertz model, the growth equation does not lead directly to express a rate of cell loss since it only expresses a decrease of the growth rate k_b inversely proportional to biofilm growth. A possibility is to express dead biofilm as follows:

$$D(t) = B_0 e^{k_b t} - B(t), \quad (23)$$

the difference between what would result from exponential growth and the actual biofilm level.

2.1.5. Modeling Post-Plateau Biofilm Decrease. To account for effects leading to a post-plateau decline in the biofilm size [4], one can introduce a hypothetical endogenous variable, X , that is reduced proportionally to the amount of biofilm present [27, 28].

$$\frac{dX(t)}{dt} = -k_i B(t), \quad X(0) = 1, \quad (24)$$

where k_i is the rate of reduction. $X(t)$ is arbitrarily set to 1 at time zero; it is positive as $B(t)$ increases, but eventually it becomes negative (as it can be seen by the relationship $X(t) = 1 - k_i \int_0^t B(t) dt$). The growth rate of biofilm takes the following form:

$$g(t, B(t), X(t)) = g(B(t)) X(t), \quad (25)$$

where $g(B(t))$ is given by (4) or (6). For example, for Logistic growth,

$$\frac{dB(t)}{dt} = \left[k_b B(t) \left(1 - \frac{B(t)}{B_{\max}} \right) \right] X(t). \quad (26)$$

The right side of equations (25) and (26) becomes negative as time progresses and leads $B(t)$ to asymptote to zero ($X(t)$ asymptotes to a negative value). The main limitation of the model is that for any value of k_l the biofilm always asymptotes to 0. To avoid this the model can be modified as follows:

$$\frac{dX(t)}{dt} = k_p - k_q X(t) - k_l B(t), \quad X(0) = \frac{k_p}{k_q}, \quad (27)$$

where k_p and k_q are production and elimination rates of the endogenous substance. Equations for dead cells are obtained as before.

We remark that especially in in vivo systems one can observe pre- or post-plateau appearance of cyclical growth, associated with seeding and dispersal [29, 30]. To model these situations one could use a time-varying B_{\max} that changes value after a pure time-delay or depending on a threshold biofilm value.

2.1.6. Modeling Dormancy. Biofilm growth can follow an initial period of dormancy that can be expressed easily using a (pure) lag time. The growth rate becomes

$$g(t, X(t)) = \begin{cases} 0 & t < t_{\text{lag}} \\ g(B(t - t_{\text{lag}})) & \text{otherwise,} \end{cases} \quad (28)$$

where t_{lag} is the lag time.

2.2. Statistical Modeling and Model Selection. Biofilm data usually consist of several measurements made on a number of growing ‘‘cells’’ on different occasions. The measurements present different levels of random variation: among measurements within a given cell (intracell variation), among cells (intercell variation), and interoccasion. A hierarchical nonlinear mixed effect model [31] can be used to represent such data. According to this model the i th observation from the j th cell and k th experiment, y_{ijk} , takes the following form:

$$y_{ijk} = V(t_{ijk}, h(\delta, \eta_j, \nu_k)) + \varepsilon_{ijk}, \quad (29)$$

where t_{ijk} is the corresponding sampling time, V is a nonlinear function describing the relationship between time, parameters, and predicted response, and $h(\delta, \eta_j, \nu_k)$ describes the parameters values for each cell, which depend on δ , a fixed effect mean parameter vector, a vector of intercell random effect parameters, η_j , and a vector of intercell random effect parameters, ν_k , with mean zero and a certain, typically multivariate normal, distribution with variance-covariance matrix Ω . The random effect ε_{ijk} indicates intracell variability (e.g., measurement error) typically assumed (multivariate) normally distributed with mean zero and covariance matrix Σ , which could depend on δ, η_j (we omit details for simplicity)

and an unknown parameter vector σ . Popular methods used to fit mixed effect models to data are implemented in the computer programs such as NONMEM [32] and NLME [33], which allow the computation of corresponding second-order statistics for the estimates, in particular the large sample variance-covariance matrix of the parameters δ , individual parameters as empirical-Bayes estimates, and corresponding predictions, conditional on the estimates for δ, Ω, Σ .

To select between different models one can use statistical model selection criteria together with the usual graphical displays based on predictions, observations, and residuals. Some statistical model selection criteria are the Hannan-Quinn (HQ) [34] and Akaike (AKA) [35]; these criteria penalize the likelihood of the model proportionally to the number of parameters in the model, thus penalizing the selection of models with ‘too many’ parameters. The HQ is the most conservative; it uses twice the log of the number of observations times the number of parameters as penalty, while the AKA uses twice the number of parameters.

2.3. Data. *Pseudomonas aeruginosa* PAO1 tagged with green fluorescent protein (GFP) were studied in flow-chamber experiments described in [36]. In brief, biofilms were grown at 30°C in flow chambers. Each flow chamber was inoculated with 250 μL of an overnight culture of PAO1 diluted to an OD_{600} of 0.05 and left without flow for one hour. After one hour, flow was started with minimal media at a flow rate of 20 ml/h using a peristaltic pump (Watson Marlow 205 S). After cultivation for 24 or 72 hours, flow was stopped and minimal media were replaced with an antibiotic flask containing the desired concentration of either Meropenem (MEM) or Tobramycin (TOB) where administered as an intermittent bolus. This antibiotic flask is connected to bubble traps and the flow chambers containing the cultivated PAO1 biofilms. Flow was restarted and minimal media were pumped from the dilution flask to the antibiotic flask to the flow chambers at a constant rate calculated to mimic the elimination rate constant of the antibiotic. MEM and TOB were obtained from the pharmacy of the University of California San Francisco Medical Center [37]. Concentration-time profiles were based on previously described PK parameters of MEM and TOB from healthy volunteers and patients with cystic fibrosis [38, 39]. The target MEM peak concentration based on human population values was computed to be 107.53 mg/L with a $t_{1/2} = 0.893$ h. The target TOB peak concentration, based on a dose of 10 mg/kg in a 70 kg adult, was 32.79 mg/L with an associated $t_{1/2} = 2.75$ h. Samples were analyzed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) [40].

Microscopic observations of the flow cells were completed using a Leica TCS SP2 confocal laser scanning microscopy (CLSM) with an argon/krypton laser and detectors and filter sets for simultaneous monitoring of GFP (excitation, 488 nm; emission, 517 nm) for live cell staining and propidium iodide (excitation, 543 nm; emission, 565 nm) for dead cell staining. Images were acquired at approximately 1 μm intervals in the z direction down through the biofilm. Each channel of the flow chamber was randomly imaged at

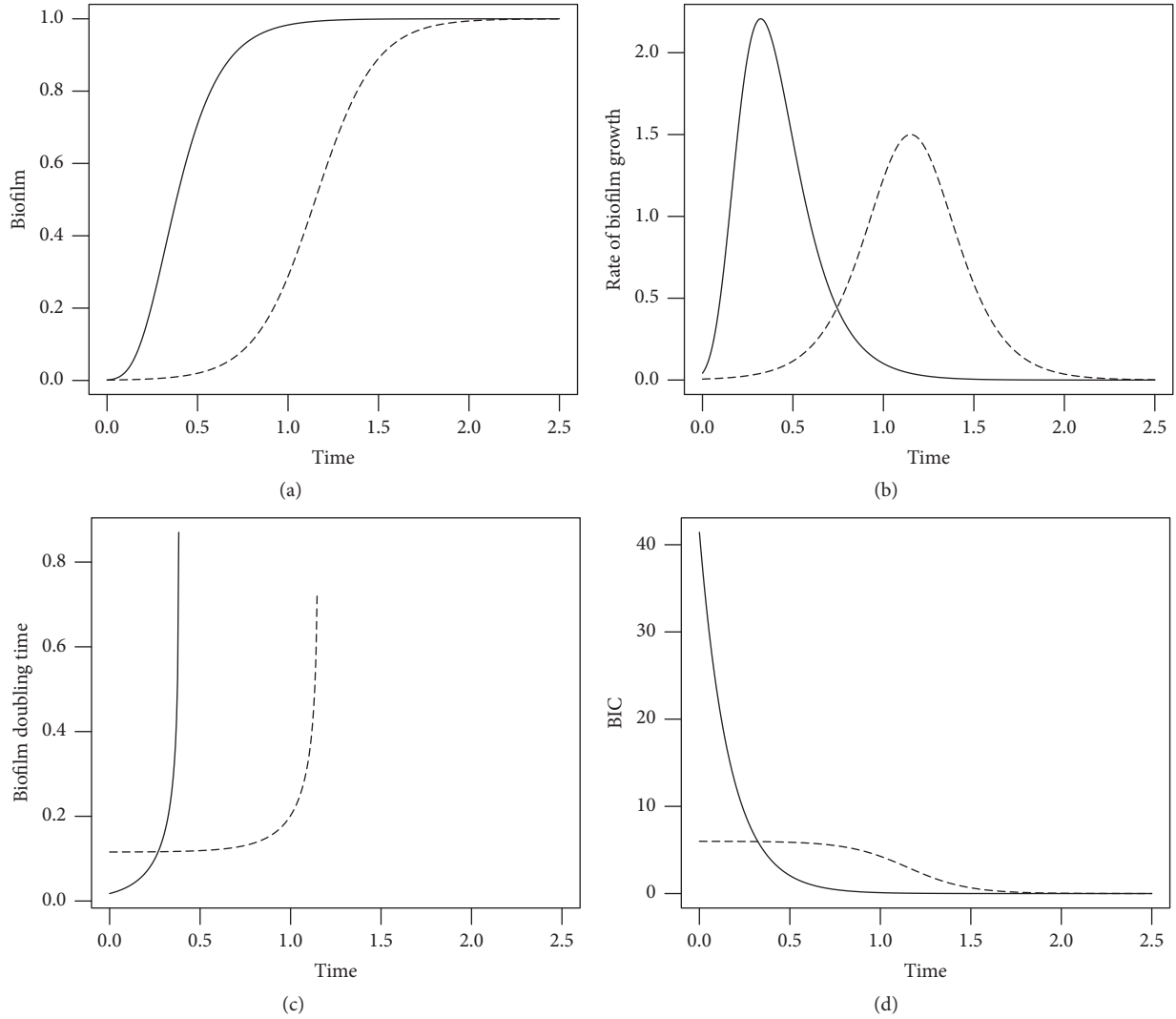


FIGURE 1: Biofilm growth models: simulation. (a) Biofilm growth according to the Gompertz (solid line) and Logistic (dashed line) models. (b) Corresponding time derivatives. (c) Corresponding biofilm doubling time. (d) Corresponding BIC. Parameter values: $B_0 = 0.001$, $k_b = 6$, $B_{\max} = 1$, $\theta_1 = 1$. (Biofilm and time are in arbitrary units.)

two separate locations per time point. CLSM images were analyzed using COMSTAT [14].

3. Results

Biofilm Growth Models: Simulation. Figure 1(a) shows a plot of the Gompertz and Logistic models (see (7) and (5)) with parameters values $B_0 = 0.001$, $k_b = 6$, $B_{\max} = 1$, and time in arbitrary units. Figure 1(b) shows their corresponding time derivative (i.e., the biofilm overall rate of change), and Figure 1(c) the biofilm doubling time. Doubling time is not a constant, as for the exponential model equation (3), but is instead expressed by the following relationships:

$$t_d = \frac{\ln(\ln(2)/e^{-k_b t} + 1)}{k_b}, \quad (30)$$

for the Gompertz, and

$$t_d = -\frac{1}{k_b} \ln\left(\frac{1}{2} - \frac{B_0}{2(B_{\max} - B_0)} e^{k_b t}\right) \quad (31)$$

for the Logistic model. Note how the doubling time is only defined up to the time for which $B(t) = B_{\max}/2$, where it tends to infinity. Figure 1(d) shows the BIC for the two models: note how the BIC is a decreasing function of $B(t)$, because intrinsic growth rate decreases as $B(t)$ reaches the asymptote B_{\max} .

Simulated Data Example 1. Different growth models are often difficult to discriminate when fitted to data. Figure 2 shows data simulated using a Logistic (Figure 2(a)) or Gompertz (Figure 2(b)) model. The data are generated by adding random normally distributed noise with mean 0 and standard deviation 0.1 to the predictions; parameter values are

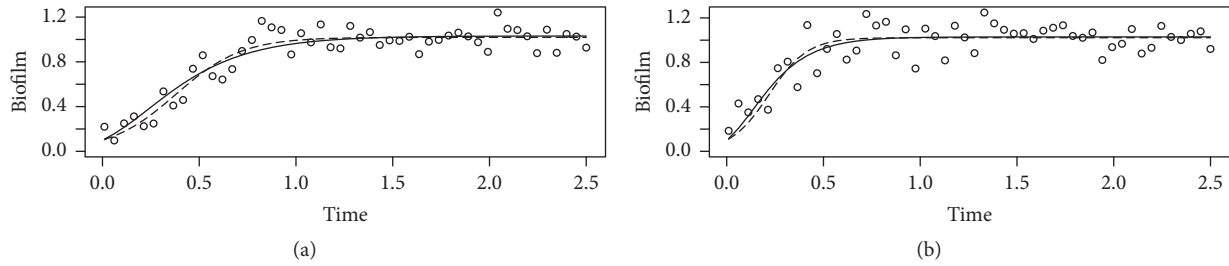


FIGURE 2: Simulated Data Example 1. Gompertz and Logistic models data fit. Circles, data simulated using a Logistic (a) or Gompertz (b) model with added random normally distributed noise; parameter values: $B_0 = 0.001$, $k_b = 6$, $B_{\max} = 1$. Solid lines and dashed lines: Gompertz and Logistic model fit to the data, respectively. See also legend to Figure 1.

$B_0 = 0.001$, $k_b = 6$, $B_{\max} = 1$. Visually, the fits are difficult to distinguish, and they are undistinguishable according to the Akaike criterion. There is difference in Akaike value between the two models of 0.8 for Figure 2(a) and 1.2 for Figure 2(b).

The Bertalanffy model equation (8) can describe rather different kinetics than the Logistic (and Gompertz) depending on the values of its parameters. In general one cannot reproduce data generated by generalized Bertalanffy process using either a Gompertz or Logistic equation.

Simulated Data Example 2. Figure 3 shows an example where data are simulated using the Bertalanffy model with $k_b = 4$, $k_d = 2$, $\lambda = 0.25$, $\delta = 1$ (Figure 3(a)) or $k_b = 4$, $k_d = 2$, $\lambda = 0.1$, $\delta = 1$ (Figure 3(b)), adding normally distributed error with mean 0.0 and standard deviation 0.1. Note the misfit of the Gompertz (dashed line) and especially Logistic model (thin dashed line) when compared to the Bertalanffy fit (solid line). The differences in the objective functions corresponding to the fits are in this case statistically significant according to the Akaike criterion.

Real Data Example 1. Figure 4 shows the result of the fits of the Logistic and Gompertz models to average thickness control data. Note how the average thickness increases and then reaches a plateau. The Logistic (dashed line) and Gompertz (solid line) fits estimate a similar value for B_{\max} but have slightly different shapes. The objective function values for the Logistic and Gompertz models were 373.50 versus 375.35, indicating a statistically insignificant, slightly better fit for the Logistic model. According to the Logistic model the maximum biofilm growth, (B_{\max}), was 52.1 (μm) with a birth rate (k_b) of 0.051 (1/h). According to the Gompertz model the maximum biofilm growth, (B_{\max}), was 53.2 (μm) with a birth rate (k_b) of 0.049 (1/h).

Real Data Example 2. Figure 5 shows the result of the fits to biomass data collected in the same cells. The Logistic (dashed line) and Gompertz (solid line) fits estimate a similar value for B_{\max} , the maximum growth level, and have slightly different shapes before asymptote toward B_{\max} . The objective function values for the Logistic and Gompertz models were 585.86 versus 585.93, indicating that the models cannot be discriminated based on the fit. According to the Gompertz model, the maximum biofilm growth, (B_{\max}), was 42.5 $\mu\text{m}^3/\mu\text{m}^2$ with a birth rate (k_b) of 0.051.3 1/h.

Real Data Example 3. This demonstrates the use of a nonlinear mixed effect model on a complex data set that includes controls, single-dose administration of MEM at TOB at 24 hours (monoexponential decay), triple dose (monoexponential decay) of both drugs at 74 hours, and a combined single-dose infusion of MEM and single dose (monoexponential) or TOB at 72 hours. Figure 6 shows the plot of the weighted residuals versus time for the fit to data; the solid lines show the R [41] function supsmu fit to the residuals. Note the overall lack of trends (a trend in the residuals would indicate a misfit of the model to the data), although there is a minor trend at the lowest and highest prediction values.

Figure 7 shows plots of the observations (open circles) and empirical-Bayes predictions by the model (solid line) for a selection of the data sets. The data are selected based on the objective function value of the individual fit to each data set. Top to bottom, right to left they correspond to the 10% (best), 25%, 40%, 60%, 75%, and 90% (worst) fit of the model. Table 1 reports the parameter estimates, showing the partition of the variability of the parameter estimates. Figure 8 shows a simulation based on those parameter estimates: drug kinetics in Figure 8(a) and the corresponding biomass versus time plot in Figure 8(b).

Real Data Example 4. Finally, as an example of a combined fit to live and dead cells Figure 9(a) shows the fit of the Logistic growth model incorporating drug effect according to (15) to the live biofilm average thickness data pooled from twelve experiments; Figure 9(b) shows the fit of to the dead cells; Figure 9(c) shows the pharmacokinetic profile imposed during in the experiments. Note how the model follows quite well the decrease in average thickness following the drug administration and tracks the corresponding increase of dead cells despite the large variability for both live and dead data measurements. The estimate of the killing rate (θ_1) in this example is 0.00243 (1/h/ μm).

4. Discussion

In this paper, we describe general mathematical models that can be used to quantify biofilm growth and interaction with different agents inhibiting or stimulating growth. The models are semiempirical in nature, in the sense that they do not include a detailed physiological description but allow the

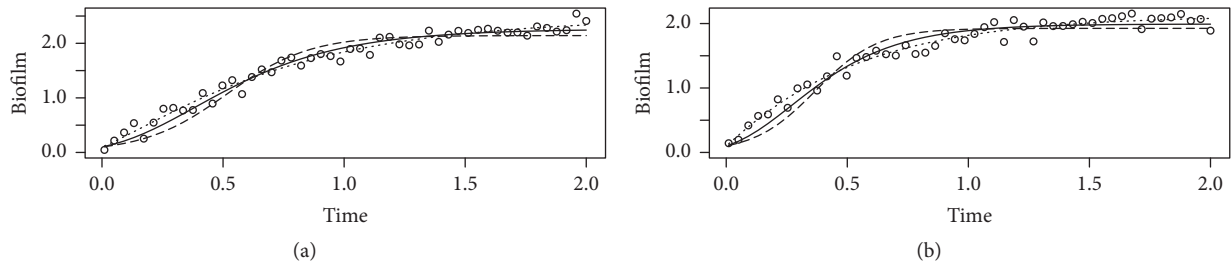


FIGURE 3: Simulated Data Example 2. Bertalanffy versus Gompertz and Logistic models data fit. Circles, data simulated using a Bertalanffy model with added random normally distributed noise. Parameter values: $k_b = 4, k_d = 2, \lambda = 0.25, \delta = 1$; $k_b = 4, k_d = 2, \lambda = 0.1, \delta = 1$ (b). Solid lines, dashed lines, and thin dashed lines: Bertalanffy, Gompertz, and Logistic model fit to the data, respectively. See also legend to Figure 1.

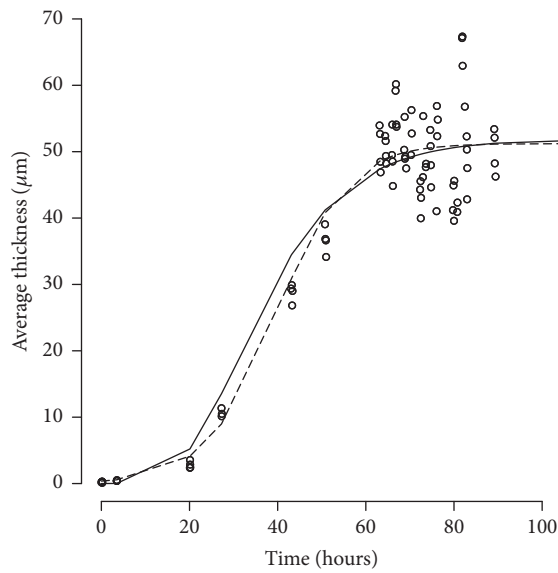


FIGURE 4: Real Data Example 1. Circles, average thickness data with superimposed predictions of the fits of a Logistic (dashed line) and Gompertz (solid line) models.

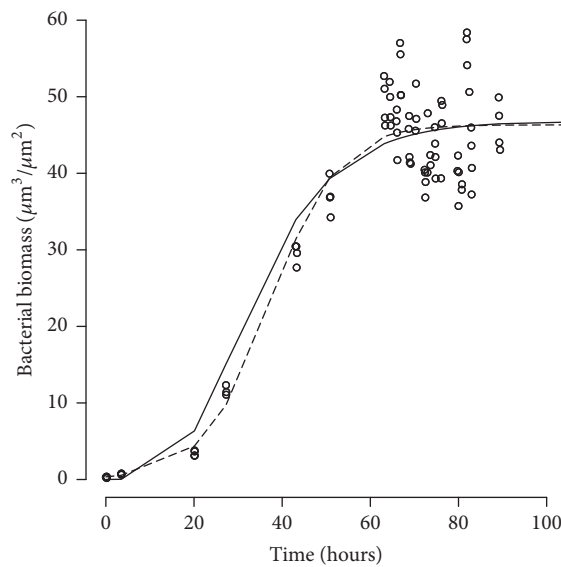


FIGURE 5: Real Data Example 2. Circles, biomass data with superimposed the predictions of the fits of a Logistic (dashed line) and Gompertz (solid line) models.

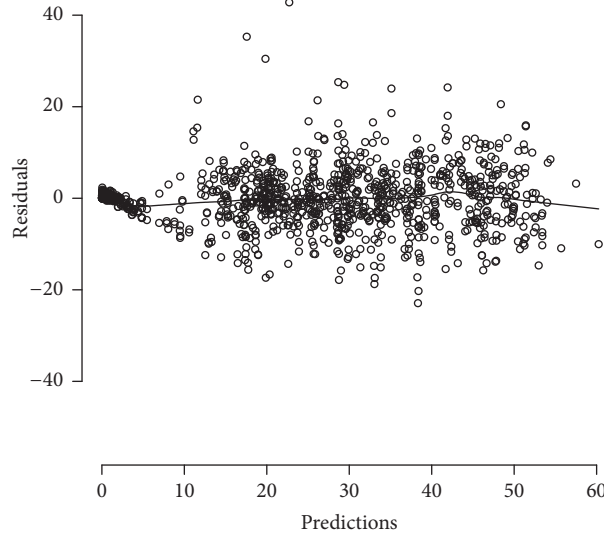


FIGURE 6: Real Data Example 3: model residuals plot. Gompertz model fit to multiple administration Meropenem (MEM) and Tobramycin (TOB) data (see also legend to Figure 8). Open circles, nonlinear mixed effect model weighted residuals versus predictions; solid line, R function upsmsu fit to the weighted residuals.

TABLE 1: Biofilm growth parameters and kills rates associated with MEM and TOB.

Parameter	Parameter estimate	Interchannel variability (%)
B_{\max} ($\mu\text{m}^3/\mu\text{m}^2$)	39.5 (1.34)	21.8
k_b (1/h)	0.0425 (0.0406)	11.1
k_1 (1/h/cu)	0.00301 (0.00011)	14.8
k_2 (1/h/cu)	0.00352 (0.000168)	20.9
k_{12} (1/h/cu ²)	0.000473 (0.000083)	7.1
<i>Residual Variability</i> σ ($\mu\text{m}^3/\mu\text{m}^2$)	8.02	-

B_{\max} , maximum biofilm growth (in absence of drug); k_b , biofilm growth rate (in absence of drug); k_1 , MEM kill rate; k_2 , TOB kill rate; k_{12} , drug interaction kill rate (cu, concentration units, 1000*ng/mL). Standard errors for the estimated parameters are in parenthesis.

quantification of biofilm growth and the interaction with growth-inducing or growth-reducing agents.

With regard to biofilm growth, we show that the commonly used BIC (the minimal drug concentration that inhibits growth) is not a constant when growth is limited, which is almost always the case with biofilms, but depends on the actual amount of biofilm and corresponding growth status.

We show how different growth models are often difficult to distinguish when fitted to data. This is due to the inherent flexibility of the models and also related to the sampling frequency and noise level present in the data. In the examples reported here, the sampling is relatively frequent. But in general, other experimental conditions might allow more frequent sampling and therefore the identification of more complex growth models and/or more complex models for the response to agents affecting biofilm growth.

The main advantage of the use of a mathematical model is that it can provide a characterization of complex data in terms of few parameters. As an example, Table 1 shows how a

model with just five parameters characterizes seven distinct experiments. In addition, a nonlinear mixed effect model allows the identification not only of the parameters characterizing the growth and growth inhibition but also of the sources of variability present in a particular experimental situation. This is important because it allows the comparison of the quality of the data arising from different experimental conditions, and it suggests where experimental variation might need to be reduced. The main advantages of using a nonlinear mixed effect model are its statistical elegance and its efficiency. The main disadvantages are the many distributional assumptions involved in the formulation of the model and the computational burden involved in estimating the likelihood of the observations. A simple alternative, which takes advantage of the relatively abundance of data in in vitro experiments, is to fit data from each cell and experiment independently, obtaining the corresponding parameter estimates, which are then used to characterize the distribution of the parameters in the collection of experiments. This requires no specialized software and might be the most appealing method of choice in many situations.

The most important function of a mathematical model is the characterization of the process underlying data generation (rather than random variation associated with data collection). The models we describe represent a summary of models already present in the literature but also include potentially useful new models describing the kinetics of dead cells, together with models describing agents and agent interaction with growth. Taken together, they can allow the representation of a vast array of biofilm data.

Conflicts of Interest

All authors have no financial and personal relationships with other people or organizations that could inappropriately influence (bias) the work described in this article.

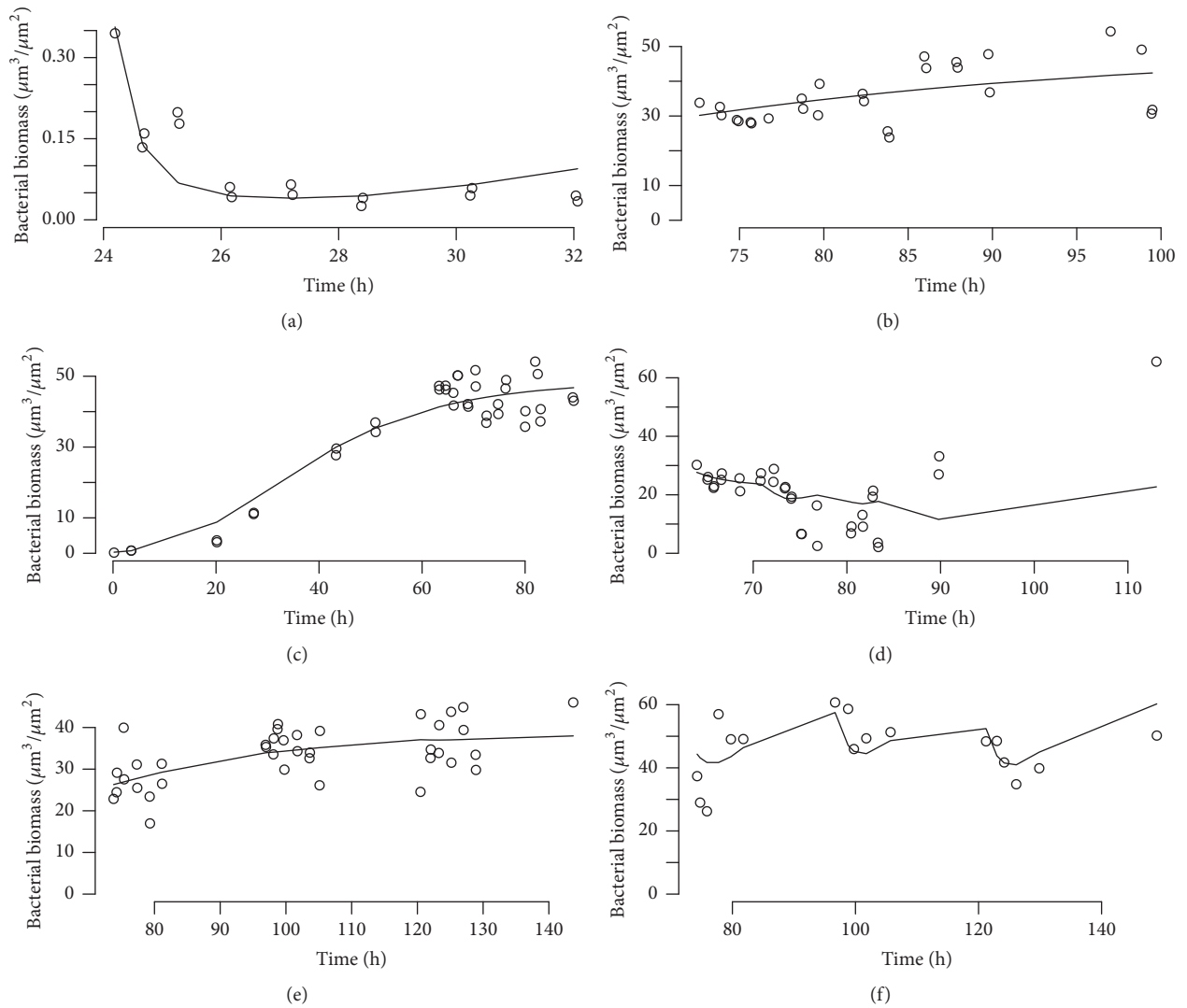


FIGURE 7: Real Data Example 3: observations and model predictions. Gompertz model fit to multiple administration MEM and TOB data (see text). In all panels: open circles, observations; solid lines, model predictions. Panels: top to bottom, right to left: 10% best fit, 25%, 40%, 60%, 75%, and 90% (worst fit).

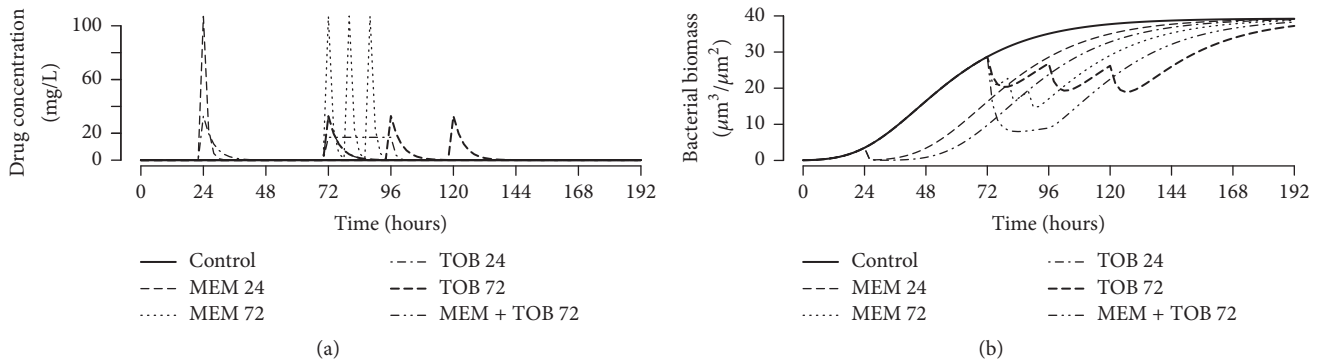


FIGURE 8: Real Data Example 3: simulation using the fitted model. (a) MEM and TOB administrations. (b) The corresponding predictions by the model using the estimates reported in Table 1.

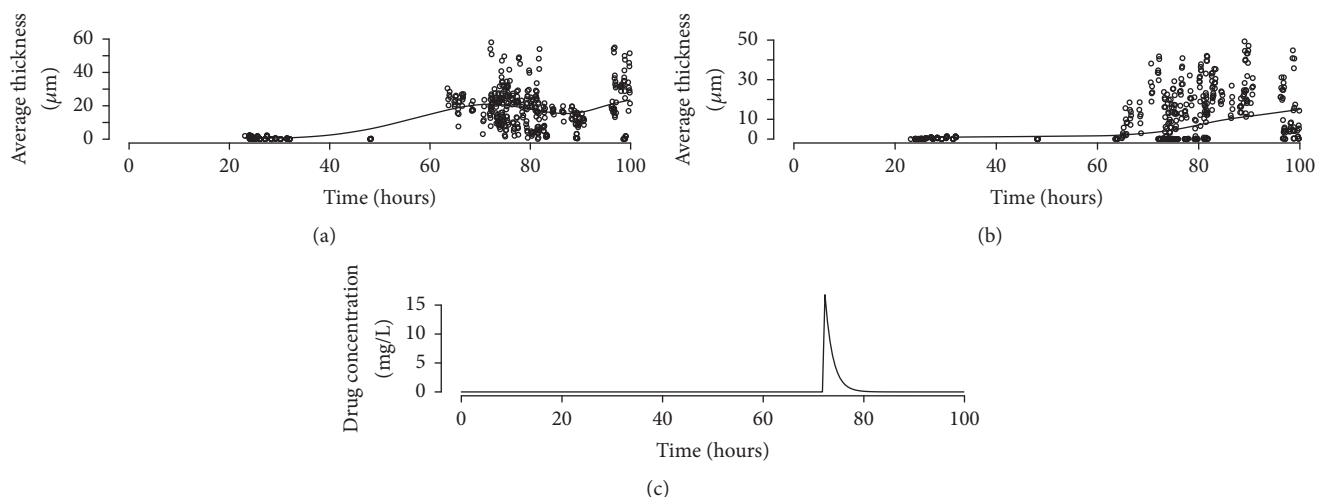


FIGURE 9: Real Data Example 4. (a) Live biofilm average thickness. (b) Dead biofilm average thickness. (c) Drug administration profile. All panels: circles, observations; solid line, model predictions.

Acknowledgments

This work was supported by Grant no. 1R01 AI097380-0121 from the National Institute of Allergy and Infectious Diseases of the National Institutes of Health.

References

- [1] L. Hall-Stoodley, P. Stoodley, S. Kathju et al., "Towards diagnostic guidelines for biofilm-associated infections," *FEMS Immunology & Medical Microbiology*, vol. 65, no. 2, pp. 127–145, 2012.
- [2] N. Høiby, T. Bjarnsholt, M. Givskov, S. Molin, and O. Ciofu, "Antibiotic resistance of bacterial biofilms," *International Journal of Antimicrobial Agents*, vol. 35, no. 4, pp. 322–332, 2010.
- [3] G. J. Tortora, B. R. Funke, and C. L. Case, *Microbiology*, Pearson/Benjamin Cummings, San Francisco, Calif, USA, 8th edition, 2004.
- [4] M. R. Parsek and T. Tolker-Nielsen, "Pattern formation in *Pseudomonas aeruginosa* biofilms," *Current Opinion in Microbiology*, vol. 11, no. 6, pp. 560–566, 2008.
- [5] M. Lipsitch and B. R. Levin, "The population dynamics of antimicrobial chemotherapy," *Antimicrobial Agents and Chemotherapy*, vol. 41, no. 2, pp. 363–373, 1997.
- [6] M. A. Beaumont, "Estimation of population growth or decline in genetically monitored populations," *Genetics*, vol. 164, no. 3, pp. 1139–1160, 2003.
- [7] J. J. Bull and C. O. Wilke, "Lethal mutagenesis of bacteria," *Genetics*, vol. 180, no. 2, pp. 1061–1070, 2008.
- [8] W. E. Olmstead and R. A. Handelsman, "Diffusion in a semi-infinite region with nonlinear surface dissipation," *SIAM Review*, vol. 18, no. 2, pp. 275–291, 1976.
- [9] N. G. Cogan, J. S. Gunn, and D. J. Wozniak, "Biofilms and infectious diseases: Biology to mathematics and back again," *FEMS Microbiology Letters*, vol. 322, no. 1, pp. 1–7, 2011.
- [10] I. Klapper and J. Dockery, "Mathematical description of microbial biofilms," *SIAM Review*, vol. 52, no. 2, pp. 221–265, 2010.
- [11] E. Ariens, "The mode of action of biologically active compounds," in *Molecular Pharmacology*, G. De Stevens, Ed., pp. 136–148, Academic Press, New York, NY, USA, 1964.
- [12] I. F. Troconiz, L. B. Sheiner, and D. Verotta, "Semiparametric models for antagonistic drug interactions," *Journal of Applied Physiology*, vol. 76, no. 5, pp. 2224–2233, 1994.
- [13] C. Csajka and D. Verotta, "Pharmacokinetic-pharmacodynamic modelling: History and perspectives," *Journal of Pharmacokinetics and Pharmacodynamics*, vol. 33, no. 3, pp. 227–279, 2006.
- [14] A. Heydorn, A. T. Nielsen, M. Hentzer et al., "Quantification of biofilm structures by the novel computer program COMSTAT," *Microbiology*, vol. 146, no. 10, pp. 2395–2407, 2000.
- [15] M. H. Zwietering, I. Jonenburger, F. M. Rombouts, and K. Van 'Triet, "Modeling of the bacterial growth curve," *Applied and Environmental Microbiology*, pp. 1875–1881, 1990.
- [16] R. Xiong, G. Xie, A. S. Edmondson, R. H. Linton, and M. A. Sheard, "Comparison of the Baranyi model with the modified Gompertz equation for modelling thermal inactivation of *Listeria monocytogenes* Scott A," *Food Microbiology*, vol. 16, no. 3, pp. 269–279, 1999.
- [17] R. C. McKellar, "A heterogeneous population model for the analysis of bacterial growth kinetics," *International Journal of Food Microbiology*, vol. 36, no. 2-3, pp. 179–186, 1997.
- [18] J. Clairambault, "Modelling physiological and pharmacological control on cell proliferation to optimise cancer treatments," *Mathematical Modelling of Natural Phenomena*, vol. 4, no. 3, pp. 12–67, 2009.
- [19] B. Gompertz, "On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies," *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*, vol. 115, pp. 513–583, 1825.
- [20] L. Norton, "A Gompertzian model of human breast cancer growth," *Cancer Research*, vol. 48, pp. 7067–7071, 1988.
- [21] F. Kozusko and Z. Bajzer, "Combining Gompertzian growth and cell population dynamics," *Mathematical Biosciences*, vol. 185, no. 2, pp. 153–167, 2003.

- [22] L. Von Bertalanffy, "Problems of organic growth," *Nature*, vol. 163, no. 4135, pp. 156–158, 1949.
- [23] S. M. Moskowicz, J. M. Foster, J. Emerson, and J. L. Burns, "Clinically feasible biofilm susceptibility assay for isolates of *Pseudomonas aeruginosa* from patients with cystic fibrosis," *Journal of Clinical Microbiology*, vol. 42, no. 5, pp. 1915–1922, 2004.
- [24] E. R. Carson, C. Cobelli, and L. Finkelstein, *The Mathematical Modeling of Metabolic and Endocrine Systems*, John Wiley & Sons, New York, NY, USA, 1983.
- [25] L. Finkelstein and E. R. Carson, *Mathematical Modeling of Dynamic and Biological Systems*, Research Studies Press, London, UK, 1985.
- [26] P. Craven and G. Wahba, "Smoothing noisy data with spline functions—estimating the correct degree of smoothing by the method of generalized cross-validation," *Numerische Mathematik*, vol. 31, no. 4, pp. 377–403, 1978.
- [27] M. J. Wade, J. Harmand, B. Benyahia et al., "Perspectives in mathematical modelling for microbial ecology," *Ecological Modelling*, vol. 321, pp. 64–74, 2016.
- [28] A. Juška, "Minimal models of growth and decline of microbial populations," *Journal of Theoretical Biology*, vol. 269, pp. 195–200, 2011.
- [29] J. B. Kaplan, "Biofilm dispersal: mechanisms, clinical implications, and potential therapeutic uses," *Journal of Dental Research*, vol. 89, no. 3, pp. 205–218, 2010.
- [30] L. Hall-Stoodley and P. Stoodley, "Biofilm formation and dispersal and the transmission of human pathogens," *Trends in Microbiology*, vol. 13, no. 1, pp. 7–10, 2005.
- [31] M. Davidian and D. M. Giltinan, *Nonlinear Models for Repeated Measurement Data*, Monographs on Statistics and Applied Probability, Chapman & Hall, 1995.
- [32] A. J. Boeckmann, S. L. Beal, and L. B. Sheiner, *NONMEM V Users Guide, Part VII; Conditional Estimation Methods*, vol. 1, Division of Clinical Pharmacology, University of California, San Francisco, Calif, USA, 2016.
- [33] M. J. Lindstrom and D. M. Bates, "Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data," *Journal of the American Statistical Association*, vol. 83, no. 404, pp. 1014–1022, 1988.
- [34] E. J. Hannan, "Rational transfer function approximation," *Statistical Science*, vol. 2, pp. 1029–1054, 1987.
- [35] H. Akaike, "A New Look at the Statistical Model Identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, 1974.
- [36] J. A. J. Haagensen, D. Verotta, L. Huang, A. Spormann, and K. Yang, "New in vitro model to study the effect of human simulated antibiotic concentrations on bacterial biofilms," *Antimicrobial Agents and Chemotherapy*, vol. 59, no. 7, pp. 4074–4081, 2015.
- [37] J. A. J. Haagensen, M. Klausen, R. K. Ernst et al., "Differentiation and distribution of colistin- and sodium dodecyl sulfate-tolerant cells in *Pseudomonas aeruginosa* biofilms," *Journal of Bacteriology*, vol. 189, no. 1, pp. 28–37, 2007.
- [38] J. L. Kuti, P. K. Dandekar, C. H. Nightingale, and D. P. Nicolau, "Use of Monte Carlo simulation to design an optimized pharmacodynamic dosing strategy for meropenem," *Clinical Pharmacology and Therapeutics*, vol. 43, no. 10, pp. 1116–1123, 2003.
- [39] D. J. Touw, A. J. Knox, and A. Smyth, "Population pharmacokinetics of tobramycin administered thrice daily and once daily in children and adults with cystic fibrosis," *Journal of Cystic Fibrosis*, vol. 6, no. 5, pp. 327–333, 2007.
- [40] L. Huang, J. Haagensen, D. Verotta, P. Lizak, F. Aweeka, and K. Yang, "Determination of meropenem in bacterial media by LC-MS/MS," *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*, vol. 961, pp. 71–76, 2014.
- [41] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2016.