Bayesian inversion for a biofilm model including quorum sensing

Leila Taghizadeh a,*, Ahmad Karimi a, Elisabeth Presterl c, Clemens Heitzinger a,b

a Institute for Analysis and Scientific Computing, Vienna University of Technology (TU Wien), Wiedner Hauptstraße 8–10, 1040 Vienna, Austria
b Institute for Analysis and Scientific Computing, Vienna University of Technology (TU Wien), Wiedner Hauptstraße 8–10, 1040 Vienna, Austria
c Department for Hospital Hygiene and Infection Control, Medical University of Vienna, Währinger Gürtel 18–20, 1090 Vienna, Austria

A R T I C L E I N F O

Keywords:
Partial differential equation
Growth and degradation of biofilms
Quorum sensing
Existence and uniqueness
Inverse problem
Bayesian inversion
Uncertainty quantification

A B S T R A C T

We propose a mathematical model based on a system of partial differential equations (PDEs) for biofilms. This model describes the time evolution of growth and degradation of biofilms which depend on environmental factors. The proposed model also includes quorum sensing (QS) and describes the cooperation among bacteria when they need to resist against external factors such as antibiotics. The applications include biofilms on teeth and medical implants, in drinking water, cooling water towers, food processing, oil recovery, paper manufacturing, and on ship hulls. We state existence and uniqueness of solutions of the proposed model and implement the mathematical model to discuss numerical simulations of biofilm growth and cooperation. We also determine the unknown parameters of the presented biofilm model by solving the corresponding inverse problem. To this end, we propose Bayesian inversion techniques and the delayed-rejection adaptive-Metropolis (DRAM) algorithm for the simultaneous extraction of multiple parameters from the measurements. These quantities cannot be determined directly from the experiments or from the computational model. Furthermore, we evaluate the presented model by comparing the simulations using the estimated parameter values with the measurement data. The results illustrate a very good agreement between the simulations and the measurements.

1. Introduction

Infection of material and devices implanted into patients’ tissues and bones is associated with considerable morbidity and costs [1–3]. The use of all kinds of implants, e.g., osteosyntheses, joint prostheses, cardiac valves and devices, percutaneous intravascular catheters, invasive monitoring to sustain life at intensive care units, and other implants is increasing. Dependent on the site of implantation, the infection rates range from 0.2% to 5% in orthopaedic and trauma surgery and up to 40% in artificial hearts [4]. Given the high incidence of fracture stabilization devices of two million per year, the number of implant infections amounts to up to 100 000 per year [3,5]. The major pathogens of implant related infections are Staphylococcus aureus and coagulase negative staphylococci, primarily Staphylococcus epidermidis [6], followed by enterobacteria, Pseudomonas aeruginosa, and enterooccus. These organisms have in common that they are difficult to eradicate by standard antibiotic therapy due to their intrinsic resistance and exposure to antimicrobials.

In implant surfaces, these organisms grow in biofilms and thus cause persistent or recurrent infections [7]. The simple definition of a biofilm is microorganisms attached to a surface. A more comprehensive definition is that a biofilm consists of a structured community of bacterial cells enclosed in a self-produced polymeric matrix and adherent to a surface. Biofilms are highly individual based on the characteristics of the microbe, environmental conditions, nutrients, implant surfaces, and host immune reaction in case of implant infections (see Fig. 1).

Biofilm associated infections are frequently resistant to conventional antimicrobial therapy, because the bacterial biofilm on the surface serves as a reservoir where bacteria are quasi inaccessible to antibiotics and the host defenses [8,9]. In the clinical routine, antibiotic susceptibility is tested by determining the minimal inhibitory concentration (MIC) of the antibiotic on free floating bacteria in the growth phase. A low concentration of the MIC indicates the susceptibility of the microorganism and it is a rough approximation of the efficacy of the treatment.

There are numerous models of medical biofilms. All models as well as all visualisation methods like staining or preparation for electron microscopy or confocal laser scanning microscopy (CLSM) have limitations. However, initial steps into modeling include simple models like
static biofilms on microtiter plates or coverslips or dynamic biofilms within flow-cells. CLSM is a good tool to visualize metabolic and/or structural changes within biofilms and a reliable starting point for more complex biofilm models [10].

In previous publications, we demonstrated that clinically achievable antibiotic concentrations do not reduce biofilms and bacterial growth. Increasing the antibiotic concentration may only reduce the biofilm thickness and reduce bacterial growth [11,12]. Moreover, we could demonstrate that changing environmental factors, such as increasing the environmental temperature or changing the composition of the implant material resulted in the reduction of biofilm mass and bacterial load [13,14].

Antibiotic therapy of these chronic biofilm associated infections is generally effective only when the infection is acute and the implanted material is removed. When it is necessary to keep the implanted material, the success of therapy is mediocre. Even the exchange of the infected implant material and the necessity to implant a new prosthesis is associated with higher recurrence rates. International recommendations for the treatment of implant infections propose the use of quinolones, high-dose beta-lactams, or daptomycin in combination with rifampicin [15]. However, antimicrobial resistance and intolerability is common in chronic infection. Thus, antibiotic therapy of these very complicated infections has to be tailored to the individual patient and the characteristics of the pathogen.

In order to understand the growth of biofilms and their response to antibiotic therapy in vivo, we have developed a mathematical model based on parabolic partial differential equations (PDE) that describes the time-dependent evolution of the size of the biofilm. The model contains parameters such as growth rate and cooperation that cannot be determined directly from experimental data. We therefore solve the corresponding inverse problem to estimate these important parameter values. In particular, we use Bayesian inversion and Markov-chain Monte-Carlo techniques as they are capable of dealing with ill-posed inverse problems. The ultimate goal is to provide validated and predictive models for biofilm formation under conditions relevant in the clinical setting. Such predictive models are invaluable in decision making for optimal treatment.

The rest of this paper is organized as follows. In Section 2, we introduce the proposed PDE model equations as well as the required initial and boundary conditions. In Section 3, existence and uniqueness of solutions to the proposed model equations are discussed. Simulation results of the growth of biofilms as the forward model under consideration are presented in Section 4. Section 5 is devoted to the experimental findings. In Section 6, we define the biofilms inverse problem and describe two different approaches, namely a genetic algorithm as a global, deterministic method and Bayesian inversion as a tool from uncertainty quantification. Numerical results of solving the biofilm inverse problem using these two approaches and hence determining the parameter values together with the model verification are presented in this section as well. Finally, conclusions are drawn in Section 7.

2. The biofilm model

Mathematical models come in many forms that can range from very simple empirical correlations to sophisticated and computationally intensive algorithms that describe three-dimensional biofilm morphology and activity. The model we propose captures two main characteristics of a biofilm: It tracks its growth and degradation depending on environmental factors and it includes the emergence of resistance (i.e., cooperation against environmental factors).

We first introduce the model equations and then discuss the meaning of the various terms and how they relate to different behaviors of bacteria forming a biofilm.

2.1. The model equations

The mathematical model is a system that consists of a parabolic partial differential equation, which is well-known as a reaction–diffusion equation describing biofilm growth, coupled to an ordinary differential equation (ODE) that describes cooperation among the bacteria including quorum sensing [16]. This type of system of equations is well-known as a reaction–diffusion-ODE model and has been applied to the modeling of, for example, interactions between cellular processes such as cell growth in mathematical biology [17]. As mentioned above, the model has two outputs: the concentration of bacteria, whose evolution in time shows the growth and degradation of the biofilm, and the cooperation of the bacteria in the biofilm exhibiting the strength of resistance and protection against environmental factors.

Our model equations are the reaction–diffusion-ODE system

\[
\begin{align*}
\partial_t u &= \nabla \cdot (AVu) + au(1 - u/\beta) - \gamma(t, x, y)(1 - v)u \quad \text{in } \mathbb{R}^+ \times D, \\
\partial_t v &= \rho \max(0, \arctan(\mu(q(u) - v)))v - k\gamma v^2 \quad \text{in } \mathbb{R}^+ \times D,
\end{align*}
\]

where

\[
q(u)(t, x, y) := \left(\frac{\partial t}{\partial t} \ast G(\cdot, \cdot)(t, x, y) = \int_{\mathbb{R}^2} u(t, \xi, \eta)G(x - \xi, y - \eta)d\xi d\eta \right) \quad \text{in } \mathbb{R}^+ \times D
\]

and

\[
G(x, y) := \frac{1}{2\pi\sigma^2} \exp \left( -\frac{x^2 + y^2}{2\sigma^2} \right) \quad \text{in } \mathbb{R}^+ \times D.
\]

Here \(u(t, x, y)\) is the concentration of bacteria. The variables \(x\) and \(y \in D\) denote position, where \(D\) is a bounded domain in \(\mathbb{R}^2\) with the boundary \(\partial D\) and \(t \in \mathbb{R}^+\) denotes time. Since each bacterium has a finite size, the value of \(u(t, x, y)\) corresponds to the thickness of the biofilm at position \((x, y)\). The variable \(\gamma(t, x, y)\) denotes cooperation. A value of \(\gamma = 0\) means no cooperation, while \(\gamma = 1\) means maximal cooperation.

We call \(q(u)\) the quorum functional. We define it as the convolution of the bacteria concentration \(u(t, x, y)\) with a two-dimensional Gaussian \(G(x, y)\) with variance \(\sigma^2\). It measures the population density in a neighborhood of \((x, y)\) at time \(t\) and is used in the quorum sensing model.
Finally, $A$, $\alpha$, $\beta$, $\rho$, $\kappa$, $\nu$ and $\mu$ denote positive constants and $\gamma(t, x, y)$ is a function from $\mathbb{R}_+^3 \times \mathbb{R}_+^2$ to $\mathbb{R}_+^3$. These constants depend on the bacterial strain and on environmental factors. They are predictive factors that will be determined and that will make it possible to predict the development of a patient’s biofilm thus allowing to choose the optimal treatment.

In the following, the rationale behind the model equations is explained.

The term $V \cdot (A \nu u)$ describes the spreading of the bacteria according to a diffusion process. Different strategies for the bacteria, given the same energy budget, correspond to different relative sizes of the constants $A$ and $\gamma$: Spreading faster (larger $A$) in a thinner biofilm that provides less protection because there is less cooperation (smaller $\gamma$) on the one hand, or spreading slower (smaller $A$) and building a better protected biofilm with more cooperation (larger $\gamma$) on the other hand.

The term $\nu$ describes the increase of the number of bacteria using a constant growth rate $\alpha > 0$. We assume that enough nutrients are present to sustain the growth of the biofilm. Hence this means that a reduction of the amount of nutrients left is not explicitly included in this model. (Of course, coupling the growth of the biofilm to a variable amount of nutrients is a possible model extension. Another model extension is to model the spreading of the bacteria depending on the local food concentration, and to model the food concentration by an additional equation.)

The term $-\alpha u/\beta$ implies that there is an upper limit for the thickness of the biofilm, $\beta > 0$ is a constant and its value best corresponding to measurements will be found in Section 6.1. This value also gives the maximum depth of the biofilm. If we consider a logistic ordinary differential equation

$$\partial_t u(t) = a u(t)(1 - u(t)/\beta),$$

then

$$\lim_{t \to \infty} u(t) = \beta$$

holds for the long-term limit.

The term $-\gamma(t)(1 - v)u$ with a function $\gamma : \mathbb{R}_+^3 \times \mathbb{R}_+^2 \to \mathbb{R}_+^3$ describes the degradation of the biofilm due to environmental factors and antimicrobial agents. The function $\gamma$ depends on the time of adding antibiotic substances. $\gamma$ is not only a constant, but a function, in order to be able to investigate the solution after an antibiotic substance has been added after the biofilm has grown for some time. In many situations, $\gamma$ is a piecewise constant function. The factor $1 - v$ means that the adverse effect of the environment on the biofilm is reduced whenever there is more cooperation represented by $v$. Of course, $1 - v$ could be replaced by any function of $v$ that vanishes for $v = 1$, that equals 1 for $v = 0$, and that is monotonically decreasing.

The variable $\nu(t, x, y)$ is proportional to the extent of cooperation among the bacteria in the biofilm. There is increasing empirical evidence that suggests that bacteria cooperate by producing exoproducts or public goods such as components of the matrix in the biofilm [18, 19]. Hence $\nu(t, x, y)$ denotes the strength of the resistance or protection that the bacteria develop against environmental influences. If their concentration $u$ is low, they do not use energy to develop resistance. However, if their local average concentration $q(u)$ is above a threshold $v > 0$, the quorum threshold (i.e., $q(u) - v > 0$), the bacteria start to cooperate and spend some of their excess energy on cooperative measures such as strengthening the matrix of the biofilm. $v$ is in the interval $[0, 1]$, where 0 denotes no cooperation and 1 means maximum cooperation. Otherwise the term including $q$ is zero.

The concept of quorum sensing has been included in biofilm models and its different aspects have been studied by many authors [20–22]. For example, in [22], the authors very recently formulated a mathematical model of QS as a stress response mechanism that increases resistance against antibiotics. In the current paper, we present a new model for biofilms describing their growth and degradation by means of a system of equations including a PDE for biofilm concentration and an ODE for biofilm cooperation. By the second equation, we define cooperativity as a new concept (or variable), and we investigate the concept of QS in the framework of the newly defined variable, i.e., the cooperation. Although QS in the sense of signaling is not studied here, we look at QS in the framework of the concept of cooperation among bacteria according to a quorum threshold by calculating a quorum function, i.e., the local average concentration of the bacteria, which measures the population density in a neighborhood of all points. There is much room for further extensions of the present model, e.g., by incorporating further equations for signaling.

2.2. The initial and boundary conditions

In order to complete the proposed model equations (1)–(3), we discuss the initial and boundary conditions. We consider the initial conditions

$$u_0(x, y) := u(t = 0, x, y) = \exp(-(x^2 + y^2)/10) \quad \forall (x, y) \in D,$$  

$$v_0(x, y) := \nu(t = 0, x, y) = \exp(-(x^2 + y^2)/10) \quad \forall (x, y) \in D.$$  

The Eqs. (4) mean that the biofilm starts to grow at the origin $(0, 0)$ with initial concentration $(0, 0)$.

Furthermore, if the solution $u$ is radially symmetric, then we transform the model equation to polar coordinates $(r, \phi)$ so that the derivatives with respect to the angle $\phi$ vanish. This yields the independent variables $r$ and $r$ and reduces the number of dimensions. We will discuss this situation in more detail below.

To supplement the model, for every point $(x, y)$ on the boundary $\partial D$, we use the zero Neumann boundary condition

$$\frac{\partial u}{\partial n} = 0 \quad \forall r \in \mathbb{R}^+ \quad \forall x, y \in \partial D,$$  

where $n$ denotes the unit outwards normal vector on the boundary.

3. Existence and uniqueness of solutions of the model

In this section, we study the existence and regularity of solutions to the initial–boundary value problem (1)–(5). To this end, we consider the initial–boundary value problem (IBVP)

$$\frac{\partial u}{\partial X} = d_\nu du + f(u, v) \quad \text{in} \quad \mathbb{R}^+ \times D,$$  

$$\frac{\partial v}{\partial X} = g(u, v) \quad \text{in} \quad \mathbb{R}^+ \times D,$$  

$$\frac{\partial u}{\partial \nu} = 0 \quad \text{on} \quad \mathbb{R}^+ \times \partial D,$$  

$$u(0, x, y) = u_0(x, y) \quad \text{in} \quad D,$$  

$$v(0, x, y) = v_0(x, y) \quad \text{in} \quad D.$$  

First some assumptions are required. Then, we state a result on local-in-time existence and uniqueness of solutions to the above IBVP.

Assumptions 1. We assume:

1. The domain $D \subset \mathbb{R}^d$ is bounded with sufficiently regular boundary $\partial D$.
2. The initial conditions $u_0$ and $v_0$ are bounded and symmetric functions on $D$.
3. The nonlinearities $f = f(u, v)$ and $g = g(u, v)$ are locally Lipschitz continuous functions on $u : \mathbb{R}^+ \times D \to \mathbb{R}^+$ and $v : \mathbb{R}^+ \times D \to (0, 1) \subset \mathbb{R}^+$.

Theorem 1 (Local-In-Time Existence and Uniqueness [17]). Under Assumptions 1, there exists $T = T(||u_0||_{\infty}, ||v_0||_{\infty}) > 0$ such that the initial–boundary value problem (6) has a unique local-in-time mild solution $(u, v) \in L^\infty([0, T], L^\infty(D))$.

We recall the definition of a mild solution and then give a brief sketch for the proof of Theorem 1.
Definition 3.1. A mild solution of the initial–boundary value problem (6) is a couple of measurable functions \( u : [0, T] \times D \rightarrow \mathbb{R}^+ \) and \( v : [0, T] \times D \rightarrow (0, 1] \subset \mathbb{R}^+ \) satisfying the system of integral equations

\[
\begin{align*}
&u(t, x, y) = S(t)u_0(x, y) + \int_0^t \int_{D} S(t-s)f(u(s, x, y), v(s, x, y))ds, \\
&v(t, x, y) = v_0(x, y) + \int_0^t \int_{D} g(u(s, x, y), v(s, x, y))ds,
\end{align*}
\] (7a)

where \( S \) is the semigroup of linear operators generated by the Laplacian with homogeneous Neumann boundary condition.

Proof. The system (1)–(5) satisfies all assumptions: It is straightforward to check that the initial conditions are bounded and symmetric, since they are Gaussian functions. Furthermore, the nonlinearities \( f \) and \( g \) in the model are locally Lipschitz continuous functions as can be seen using the mean-value theorem and due to the boundedness of the computational domain and the derivatives of the functions \( f \) and \( g \) on the domain. The function \( \tan^{-1} \) has a bounded derivative and the derivative of a polynomial (as a \( C^1 \) function) is continuous and thus bounded.

The proof is based on an application of the Banach fixed-point theorem in order to construct a local-in-time unique solution for system (7), as the nonlinearities \( f \) and \( g \) are locally Lipschitz continuous functions. Details of the approach and the proof of Theorem 1 for even more general reaction–diffusion equations can be found in [23, Theorem 1, p. 112]. The construction of nonnegative solutions of a particular reaction–diffusion model can also be found in [26].

Remark 1. To obtain more regular solutions, we can use more regular initial conditions, i.e., if \( u_0 \in H^2_0(D) := \{ u \in H^2(D) : u \cdot \nu = 0 \text{ on } \partial D \} \), then \( u(t) \in C([0, T]; H^2(0, 1) \times L^2(0, 1)) \). The proof is based on the theory of strongly continuous semigroups [25]. Furthermore, if \( u_0 \in C^{2+\eta}(D) \) and \( v_0 \in C^{\eta}(D) \) for some \( \eta \in (0, 1) \), and the compatibility condition \( \nabla \cdot (\nabla u_0) = 0 \) holds on \( \partial D \), then the mild solutions of the IBVP (6) are smooth and satisfy \( u \in C^{1+\eta/2+\eta}(0, T) \times D \) and \( v \in C^{1\eta}(0, T) \times D \).

For more details and studies of general reaction–diffusion–ODE systems in Hölder spaces, we refer the reader to [23, Theorem 1, p. 112] and [26].

4. The forward problem

In this section, we solve the IBVP (1)–(5) numerically in order to discuss its behavior and to observe the evolution of concentration and cooperation in time.

4.1. The model in polar coordinates

As mentioned, the solution \( u \) is radially symmetric for radially symmetric initial conditions. Therefore, we can transform the model equation to polar coordinates \( (\rho, \phi) \) and the derivatives with respect to the angle \( \phi \) vanish. This results in a problem in the independent variables \( t \) and \( \rho \), and thus reduces the number of dimensions.

The original model can be rewritten as

\[
\begin{align*}
\frac{\partial u}{\partial t} &= A \frac{\partial^2 u}{\partial \rho^2} + A \frac{\partial u}{\partial \rho} + \frac{\sigma}{\rho} (1 - u/\beta) - \gamma (1 - v)u, \\
\frac{\partial v}{\partial t} &= \rho \max(0, \arctan(\mu(q(u) - v)))v - \kappa v^2, 
\end{align*}
\] (8a)

where

\[
q(u)(t, \rho) := \left( u(t, \cdot) * r G(\cdot) \right)(\rho)
\] (9)

and

\[
G(\rho) := \frac{1}{2\pi \sigma^2} \exp\left(-\frac{\rho^2}{2\sigma^2}\right),
\] (10)

with the initial conditions

\[
\begin{align*}
u(t, 0, r) &= \exp(-r^2/10), \\
u(t, 0, r) &= \exp(-r^2/10),
\end{align*}
\] (11a)

which are radially symmetric, and with zero Neumann boundary condition for \( u \). We have implemented the PDE model (8)–(11) of growth and degradation of biofilms including quorum sensing by means of the method of lines (MOL) and present the results in the following section. MOL is a technique for solving partial differential equations by discretizing in all but one dimension and then integrating the semi-discrete problem as a system of ODEs. Here, the discretization in space is done by means of the finite-difference method, and the resulting ODE system in time is solved using a multistep solver based on numerical differentiation formulas (NDFs).

4.2. Numerical results of the biofilm forward problem

Here, we present simulation results for the model (8)–(11). We assume a circular dish with diameter \( 35 \) mm centered at the origin as the growth space and the computational domain for the biofilms, which are initially located at the center of the dish. Biofilm growth is monitored for about six hours. The goal is to calculate the area covered by the biofilm in every hour.

In these simulations, the initial biofilm was located at the center of the dish. We find the largest circle in the computational domain where the concentration is above a certain threshold and calculate the relative coverage using the area of this circle. We discuss numerical results of biofilms growth and cooperation using different parameter values. The default setup for the parameters is \( A = 1.01, a = 1.03, \beta = 0.74, v = 0.52, \kappa = 0.2, \rho = \beta = \sigma = \gamma = 1 \). In the numerical experiments, we study the effect of parameter variations on the solutions of the biofilm model, namely the concentration and the cooperation, while keeping the rest of the parameters constant with the default values.

Fig. 2 displays the resulting concentrations of biofilms using diffusion constants \( A = 0.02 \) and 0.8. These results show that when the concentration spreads faster (i.e., when the diffusion constant is larger), the biofilm is thinner and the covered area is larger. For a smaller diffusion constant, the biofilm is thicker and more concentrated with smaller coverage, as expected. This result agrees with our qualitative assessment in Section 2.1.

Fig. 3 shows concentrations of biofilms for three different growth rates \( a = 1.01, 1.03 \) and 1.05. In fact, if two populations of bacteria start to grow with the same diffusion constant but different growth rates, the one with higher growth rate creates a much more concentrated and thicker biofilm. Furthermore, the numerical results show larger coverage area for a bigger growth rate.

We observe the effect of different values of \( \beta \) on the concentration and the cooperation of the biofilms in Fig. 4. According to these results, using a larger value for the parameter results in a thicker biofilm and bigger coverage, however slight changes in the parameter value hardly affect the relative coverage.

Moreover, in Fig. 5, the effect of different values of the parameter \( \kappa \) is studied and the corresponding numerical results, i.e., the concentration of biofilm and the cooperation are illustrated. According to these results, the cooperation among the bacteria increases with larger \( \kappa \), and the coverage is almost constant.

According to the simulation results, uncertainty in the model parameters leads to different results. Now the question is what the actual parameter values are. In Section 6, we will answer this question by means of solving the biofilm inverse problem and identify all the unknown quantities of the model. Before we can do that, the experimental results are described in the next section.
5. Experimental results

In this section, we present the results of our experiments and observations for biofilm growth. In these experiments, we aimed to monitor the development of biofilms during 24 h. To this end, 1 mL of each reference (diluted in Brain Heart Infusion Medium $1.5 \cdot 10^7$) was cultivated in single 24-well Ibidi μ-Dishes (Ibidi Treat 1, 5 polymer coverslip, tissue culture treated; Ibidi GmbH, Planegg/Martinsried, Germany). Biofilms were grown at 37°C for 24 h on an orbital shaker. Every hour one well was taken off for further investigation. Furthermore, biofilms were washed two times in PBS and fixed with 4% glutaraldehyde. We observed bacteria in biofilms and the matrix (EPS) structure at different times of biofilm formation. To observe the dense DNA of the dead bacteria (Molecular Probes®; Thermo Fisher Scientific), propidium iodide was used. Polysaccharides, representing the most characteristic fraction of the extrapolymeric substances, were stained using concanavalin-A (ConA) (Sigma-Aldrich Corp, St. Louis, MO, USA). Fig. 6 displays the formation of S. epidermidis and S. aureus biofilms after six hours of incubation.

During the 24-h long observation of the reference strains, we found differences in the pattern of biofilm formation; S. aureus ATCC25923 aggregated and formed various grape-like aggregations of bacterial cells coated by single polysaccharides before spreading on the surface and forming a biofilm layer. In contrast, S. epidermidis biofilms started with scattered cells spreading over the surface until reaching confluence and their maximum thickness at 24 h without forming grapes or clusters [27].

Fig. 7 illustrates the area coverage by the two reference strains of biofilms, namely S. aureus ATCC 25923 and S. epidermidis DSM 3269 after 6 h. In these figures, we show how much area is covered by polysaccharide, mix, and DNA.

In the experiments, the relative coverage at time $t$ is defined as the relative area of the dish covered by the biofilm as measured by counting particles using the ImageJ software. ImageJ is an open source Java image processing program for automatic particle counting and analyzing.

In this work, we have not included the amount of available nutrients as a variable in our model, but we assume that enough nutrients
are present in order to enable biofilm growth. As already mentioned, coupling the growth of the biofilm to a variable amount of nutrients is a possible model extension. Further model extension are to model the spreading of the bacteria depending on the local food concentration and to model the food concentration by an additional equation. Furthermore, if we include the amount of nutrients in the model, new findings show that the growth of biofilms can even be affected by the communication between nearby bacterial communities [28]. The bacterial communities seem to use a time-sharing strategy in the sense that two distinct biofilms can synchronize their growth through electrical signals. When bacteria face limited nutrients, they alternate feeding periods and each community takes turns consuming nutrients in order to reduce competition and to maximize efficiency in consumption. This behavior extends communication among functional units such as cells to biofilms. It modifies cell-to-cell signaling mediated by ion channels [29–32] to signaling between two distinct biofilms.

6. The inverse problem

The coefficients and parameters used in the mathematical model are effective factors in the growth and cooperation of biofilms, but they have unknown and uncertain values in the sense that they can only
be determined by measurements that always come with measurement errors. These parameters cannot be controlled during the experiments. However, knowing the parameter values that determine the growth and cooperation of biofilms helps to react against biofilms after diagnosis in the sense that these parameter values give concrete information about the bacterial behavior. Understanding the bacterial behavior is also helpful to know which sort of antibiotics and how much of them should be used to kill the bacteria. Therefore any information about these effective factors improves the speed and quality of the treatment.

There are various methods for estimating the parameter values and solving the inverse problem. One can divide the statistical inference approaches into two categories: frequentist and Bayesian inference. The differences between these two approaches stem from the way the concept of probability is interpreted.

In the frequentist approach, the unknown parameter is assumed to be fixed and deterministic, and probabilities are defined as long-term frequencies of occurrences of an event. The event has to occur many times. Therefore, in this approach we collect data from a sample of the population and estimate its mean as the value which agrees best with the data. In contrast, in the Bayesian approach, the unknown parameters are assumed to be random variables. In this technique, probabilities are rooted in degrees of belief and logical support and can be used to represent uncertainties in any event, even in non-repeatable events. In Bayesian inference, we define probability distributions over possible parameter values and use data to update the distribution, which means that beliefs are updated in response to new evidence. The updating is done by applying Bayes’ theorem. In fact, the new information (e.g., experimental data) makes the probability distribution more focused around the true value of the unknown parameter. Thus confidence intervals can easily be calculated.

Here, our goal is to estimate effective quantities in the biofilm model $S.\text{epidermidis} \text{ DSM 3269 (left) and } S.\text{ aureus ATCC 25923 (right) (polysaccharide, mix, DNA) after } 6\text{ h based on experimental data.}$

![Image](image-url)

**Fig. 7.** Total area covered by the biofilms $S.\text{epidermidis DSM 3269 (left) and } S.\text{aureus ATCC 25923 (right) (polysaccharide, mix, DNA) after }6\text{ h based on experimental data.}$

6.1. Deterministic approach: Genetic algorithm

In this section, we use a genetic algorithm (GA) as a global-optimization method in order to estimate the parameters of interest in the biofilm inverse problem. The genetic-algorithm based approaches are deterministic methods in the sense that the parameters of interest are assumed to be deterministic variables, whose optimal values are found by means of minimizing a cost (fitness) function. This method must be equipped with a cost function, since it enables the method to rank the individuals. Here, the cost function is defined as the residual

$$\text{cost} = \|\text{cov} - \text{cov}_\text{exp}\|_2,$$

between the experimental data and numerical results of the forward problem, i.e., the coverage data obtained by using the MOL solver for the forward model. Here cov and cov$_\text{exp}$ denote the coverages obtained from the simulations and the experiments, respectively. In the following, the steps of the genetic algorithm are shown [33]:

1. **Initialization.** In this step, a random initial population is generated.
2. **Evaluation.** Each individual is evaluated using a fitness function.
3. **Selection.** Selection is a process in which the individuals which are suitable for the next generation are chosen. In this step, our model adopts tournament selection. This step is repeated until the number of individuals selected is equal to the desired population size. In order to ensure the propagation of elite individuals, elitism is used. This mechanism selects the individuals with the best fitness values and directly places them in the next generation, while the remaining individuals must go through the selection process.
4. **Crossover and Mutation.** Using crossover and mutation, children are produced from the parents. In mutation, children are produced by making random changes to a single parent, while in crossover this is done by combining the vector entries of a pair of parents.
5. **Evaluation.** Each individual is evaluated using the fitness function.
6. **Check termination criteria.** Steps 2 to 5 are repeated until the termination criteria are satisfied. Here the algorithm is stopped when the maximum number of generations has been exhausted or when the solution with the best fitness has not been changed from the previous generation.

Table 1 lists the GA setting and can be used for reproducing the results. The selection function defines how individuals are selected to become parents. The number of elite children shows the number of
individuals with the best fitness values in the current generation that are guaranteed to survive into the next generation. These individuals are called elite children. Moreover, crossover fraction is the fraction of individuals in the next generation, other than elite children, that are created by crossover. The remaining are generated by mutation.

In this algorithm, we have calculated the mean and the minimum of the obtained residual in the populations in 50 generations and we have shown the results in Fig. 8. As the number of generations increases, the individuals in the population crowd around the minimum point 0.

The GA yields the optimal parameter values in the biofilm model equation (1) by solving the optimization problem

\[(A, \alpha, \beta, \nu, \kappa, \rho, \sigma, \mu, \gamma) = \arg \min \text{cost},\]

where the cost function is defined in (12). The values found are shown in Table 2. We will use these values as the initial values for the Markov chains of the corresponding parameters in the stochastic approach in the next section. A near optimal choice of starting states of a Markov chain helps the chain to quickly converge around the true value of the parameters of interest.

As mentioned before, the genetic algorithm is deterministic in the sense that it does not take into account the uncertainty in the measurements. Therefore, in the next section, we consider a Bayesian framework that accounts for the uncertainty in the problem.

### 6.2. Bayesian PDE inversion

The second approach to solve the biofilm inverse problem is Bayesian PDE inversion [35–37]. In this method, we consider the unknown parameters as random variables and calculate the posterior probability density that reflects the distribution of the parameter values based on the observations. Therefore, in this method, not a single parameter value but its probability distribution is found. This is an advantage, since probability distribution conveys information how well the parameters can be determined.

In Bayesian techniques, prior knowledge about the parameters of interest is updated using measurements to obtain posterior information about the parameters. The connection between the parameter information is made by means of the well-known Bayes’ Theorem from probability theory but interpreted on top of a model equation.

We denote the probability space by \((\Omega, F, P)\), where \(\Omega\) is the sample space, \(F\) a \(\sigma\)-algebra of all events, and \(P\) a probability measure. It is assumed that all the random variables are absolutely continuous and that the unknown parameters \(q \in \mathbb{R}^p\) and the measured data \(y\) are realizations of the random variables \(Q\) and \(M\), respectively.

They are connected by the statistical model \(M = G(Q) + e\).

In this model, \(e\) is the measurement error, which is a mean-zero random variable, and \(G(Q)\) is the observation operator dependent on the random variable \(Q\) with realizations \(q = Q(\omega)\). Moreover, it is assumed that \(\pi_q(y), \pi(y|q)\), and \(\pi(y|q)\) are the probability density functions of the prior, posterior, and (data) sampling distributions, respectively. The density \(\pi(y|q)\) of the data provides information from the measured data to update the prior knowledge, and it is usually called the likelihood density function.

The goal of Bayesian inversion is to estimate the posterior probability density function \(\pi(y|q)\), which reflects the uncertainty about the quantity of interest \(q\) using measured data \(y\). Bayes’ Theorem for inverse problems can be stated as follows.

**Theorem 2** (Bayes’ Theorem for Inverse Problems [37,38]). Let \(\pi_q(y)\) be the prior probability density function for the realizations \(q\) of the random parameters \(Q\). Let \(y\) be a realization or measurement of the random observation variable \(M\). Then the posterior density of \(Q\) given the measurements \(y\) is

\[
\pi(y|q) = \frac{\pi(y|q)}{\int_{\Omega} \pi(y|q)\,dq} = \frac{\pi_q(y) \pi(y|q)}{\int_{\Omega} \pi_q(y) \pi(y|q)\,dq}.
\]

Calculating the integral in (14) is costly especially in the case of high-dimensional parameter spaces \(\mathbb{R}^p\). In order to find the posterior density without computing the costly integral in (14) and in order to efficiently sample the parameter space, Markov-chain Monte-Carlo (MCMC) methods are the method of choice. In these methods, a reversible Markov-chain is constructed whose stationary distribution is the posterior density.

Choosing the proposal scales in the MCMC algorithm is crucial as they affect the convergence speed of the Markov chain. To this end, we use the delayed-rejection adaptive-Metropolis algorithm (DRAM) [39], which is a variant of the Metropolis–Hastings algorithm that combines adaptivity [40] and delayed rejection [41]. The adaptive part updates the proposal covariance matrix with an optimal scale and ensures that information learned about the posterior distribution is remembered as the chain progresses. The delayed rejection updates the proposal scale when the proposed value is rejected to improve mixing and to avoid stagnation of the chain.

The adaptive Metropolis (AM) algorithm is a global adaptive strategy, where a recursive relation is used to update the proposal covariance matrix. In this algorithm, we take the Gaussian proposal centered at the current state of the chain \(q_k\) and update the chain covariance matrix at the \(k\)th step using

\[V_k = \sigma^2 P \text{Cov}(q_k, q_{k-1}, \ldots, q_1) + \epsilon I_p.\]
where $s_p$ is a design parameter and depends only on the dimension $p$ of the parameter space. This parameter is specified as $s_p := 2.38^2/p$ as the common choice for Gaussian targets and proposals \cite{39}, as it optimizes the mixing properties of the Metropolis–Hastings search in the case of Gaussians. Furthermore, $I_p$ denotes the $p$-dimensional identity matrix, and $\epsilon > 0$ is a very small constant to ensure that $V_k$ is not singular theoretically, and in most cases it can be set to zero.

The adaptive Metropolis algorithm employs the recursive relation

$$V_{k+1} := \frac{k - 1}{k} V_k + \frac{s_p^2}{k} \left( k \bar{q}_{k-1} \bar{q}_{k-1}^T - (k + 1) \bar{q}_k \bar{q}_k^T + \epsilon I_p \right)$$

to update the proposal covariance matrix, where the sample mean $\bar{q}_k$ is calculated recursively by

$$\bar{q}_k = \frac{1}{k+1} \sum_{i=1}^{k} q_i$$

where $q_i$ is the $i$-th sample. The number of samples or iterations in the algorithm must be large enough to estimate the parameter after discarding a sufficiently long burn-in period at the beginning.

6.2.1. Numerical results of the Bayesian inversion for the biofilm inverse problem

In this section, to provide a better insight into the biofilm model, we quantify uncertain values of the model by means of the Bayesian inversion method and by comparing and analyzing measurements and simulations. Bayesian inference has been already applied in order to parameter identification in different models related to biofilms \cite{42,43}. Here the goal is to extract information as much as possible using the measurements. However the amount of recovered information depends on many factors including the size of available experimental data. We have used the DRAM algorithm as an adaptive MCMC method in the context of the Bayesian inversion method for this analysis. To start the Markov chains, we use the optimal parameter values found by the genetic algorithm, which are summarized in Table 2 in Section 6.1 to speed up the convergence of the generated chains. We also use the measured area covered by the $S.\text{epidermidis}$ biofilm (mix) according to Section 5.

Fig. 9 illustrates the 9D Bayesian estimation results and it displays the marginal histograms of the resulted posterior distribution of nine parameters of the biofilm model together with their mean values. The means of the estimated posterior distributions have a good agreement with the optimal values found by the GA in Section 6.1. The results give us confidence intervals for each of the unknown quantities as well.

The correlation between some of the model parameter pairs is illustrated in Fig. 10 by showing two dimensional histograms of the posterior distribution of the pairs calculated by the DRAM algorithm. For the rest of the pairs, similar histograms have been obtained.
parameter estimation. The model evaluation

In order to verify the response of the presented biofilm model, we compare the simulated coverage using the optimal parameter values with the experimental data. Fig. 11 illustrates this assessment and shows a very good agreement between the simulations and the measurements, which proves the efficiency of the presented mathematical model as well as the robustness of the applied inverse methods for parameter estimation.

Algorithm 1 The DRAM algorithm

Initialization:
Choose the first state of the chain $q_0$ such that $\pi_0(q_0) > 0$.
Choose the number $N_{\text{samples}}$ of samples or iterations (often $N_{\text{samples}} > 10^3$).
Choose the parameter $\epsilon$.
Choose the initial proposal covariance matrix $V_0$ (diagonal or symmetric).
Choose the factor $\gamma_1$ (often $\gamma_1 := 1/5$) for the second-stage proposal distribution.

for $k = 1 : N_{\text{samples}}$ do

1. (Adaptivity:) The covariance matrix $V_k$ in the $k$-th step is updated by (15).

2. A first-stage proposal $q^*$ is generated from $J(q^*|q_{k-1}) := N(q_{k-1}, V_k)$.

3. The new value $q^*$ is accepted with probability

$$a(q^*|q_{k-1}) = \min \left(1, \frac{\pi(q^*)}{\pi(q_{k-1})} \frac{J(q_{k-1}|q^*)}{J(q^*|q_{k-1})} \right). \quad (16)$$

4. If the new state is accepted, we set $q_k = q^*$.
   Otherwise:
   (a) (Delayed rejection:) A second-stage proposal $q^{**}$ is generated from proposal density

$$J_2(q^{**}|q_{k-1}, q^*) := N(q_{k-1}, \gamma_1^2 V_k). \quad (17)$$

   where $V_k$ is the adapted covariance matrix.
   (b) The new value $q^{**}$ is accepted with probability

$$a_2(q^{**}|q_{k-1}, q^*) := \min \left(1, \frac{\pi(q^{**})}{\pi(q_{k-1})} \frac{J(q^*|q_{k-1})(1 - a(q^*|q_{k-1}))}{J(q^{**}|q_{k-1})(1 - a(q^{**}|q_{k-1}))} \right). \quad (18)$$

(c) If the new state is accepted, we set $q_k := q^{**}$, otherwise $q_k := q_{k-1}$.

end for

6.3. The model evaluation

7. Conclusions

In this work, we presented a reaction–diffusion-ODE system as the model equations for biofilms including the cooperation among the bacteria. This is the first time – to the best of our knowledge – that cooperation of bacteria in biofilms is introduced and quantified. The model includes also the concept of quorum sensing, which is presented in the framework of the newly defined quantity, i.e., the cooperation in the sense that when local average mass of biofilms is above a quorum threshold, the bacteria start to cooperate, which leads to higher resistance against antibiotics. We also presented local-in-time existence and uniqueness results as well as regularity of solutions to the presented model type.

To provide better insight into the model, we applied a multi-parameter Bayesian analysis to quantify the unknown parameter values of the model. We proposed the DRAM algorithm in the context of Markov-chain Monte-Carlo methods to extract multiple parameters of the presented biofilm model and estimated posterior distribution of the unknown quantities and their confidence intervals. Moreover, in order to speed up the method, we used a genetic algorithm to determine the starting values of the Markov chains. Carefully choosing the starting state ensures that the chain converges to the true value more quickly. Furthermore, we assessed and validated the presented biofilm model using the simulation results with the estimated optimal parameter values and the measurements. The results show a very good agreement between the model response and the experimental data. The mathematical model equations in conjunction with Bayesian PDE inversion make it possible to assign biologically and medically relevant parameter codes.
values to various species of bacteria. This procedure therefore allows us to quantify and to compare the behavior and the strategies of different species of bacteria.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work was supported by FWF (Austrian Science Fund) START project Y660 PDE Models for Nanotechnology.

References