Spatio-temporal Model Checking for 3D Individual-Based Biofilm Simulations

Bowen Li[1,2], Jayathilake Pahala Gedara[3], Yuqing Xia[4], Thomas P. Curtis[4], and Paolo Zuliani[1,2(✉)]

[1] School of Computing, Newcastle University, Newcastle upon Tyne, UK
[2] Interdisciplinary Computing and Complex bioSystems (ICOS) Research Group, Newcastle University, Newcastle upon Tyne, UK
{bowen.li2,paolo.zuliani}@newcastle.ac.uk
[3] Department of Oncology, University of Oxford, Oxford, UK
jayathilake.pahalagedara@oncology.ox.ac.uk
[4] School of Engineering, Newcastle University, Newcastle upon Tyne, UK
{yuqing.xia,tom.curtis}@newcastle.ac.uk

Abstract. Individual-based microbial modelling (IbM) is a bottom-up approach to study how the heterogeneity of individual microorganisms and their local interactions influence the behaviour of microbial communities. In IbM, microbes are represented as particles endowed with a set of biological and physical attributes. These attributes are affected by both intra- and extra-cellular processes resulting in the emergence of complex spatial and temporal behaviours, such as the morphology of microbial colonies. However, the quantitative and qualitative analysis of such behaviours is difficult and often relies on visual inspection of large quantities of simulation data or on the implementation of sophisticated algorithms for data analysis. In this work, we aim to alleviate the problem by applying SSTL (Signal Spatio-Temporal Logic) model checking to formally analyse the spatial and temporal properties of 3D microbial simulations (so-called traces). Complex behaviours can be then described by simple logical formulas and automatically verified by a model checker. We apply SSTL to analyse several outstanding spatio-temporal behaviours regarding biofilm systems, including biofilm surface dynamics, their detachment and deformation under fluid conditions.

Keywords: Spatio-temporal model checking · Individual-based modelling · SSTL · NUFEB · Biofilm

1 Introduction

Spatial bio-modelling and simulation are powerful methods for understanding complex structural characteristics of biological systems. The approach uses mathematical equations and computers to mimic, simulate, and predict the system behaviour in an explicit and efficient way. In the realm of microbial ecology, spatial modelling of microbial communities is typically constructed in two ways.

J. Bowles et al. (Eds.): DataMod 2021, LNCS 13268, pp. 157–174, 2022.
https://doi.org/10.1007/978-3-031-16011-0_11

Population-based Models (PbMs) use partial differential equations to directly describe population changes over time and space [11]. They are continuum models where time, space, and microbial density are continuous variables (based on a mesh-represented structure) rather than discrete variables. An alternative to PbMs are *Individual-based Models* (IbMs) [10,12] which have gathered considerable attention due to their ability to precisely capture how the heterogeneity of individual organisms and local interactions influence emergent behaviour of microbial communities. Unlike PbMs, conventional IbMs combines both continuum and discrete methods by representing environmental conditions such as soluble nutrients as continuum fields, and individual microbes as discrete particles. In addition, each microbe has its own set of biological and physical attributes, such as growth rate, mass, position, and diameter. The collective action of each individual allows one to explore hypotheses relating local changes to attributes at the population or community level. As an example, changing the amount of nutrients in a growth environment can significantly affect both the thickness and surface area of the resulting microbial colony due to the heterogeneity of microbial metabolism [2].

While IbMs and PbMs offer powerful frameworks to facilitate bio-modelling in space and time, the post-hoc analysis of spatio-temporal properties of simulations can be challenging. Such analysis is often achieved by either manual (visual or textual) inspection of simulation traces, or by developing sophisticated bespoke algorithms for data processing. For example, to quantify the geometrical characteristics of a simulated 3D biofilm colony (surface area, roughness, average height, *etc.*), one may have to implement a set of analysis algorithms based on discrete approximations of specific simulation domains, as third-party functions of existing software [10]. This becomes one of the major barriers preventing scientists to gain a rigorous understanding of complex behaviours from model simulations.

To alleviate the problem, we propose the use of *spatio-temporal model checking* to specify and formally verify biological characteristics in space and time. Unlike traditional model checking where the analysis focuses on temporal evolution of system, spatio-temporal model checking allows reasoning about both time and space, with topology as a mathematical framework. In our case, spatio-temporal properties such as biofilm surface structure over time, are expressed by *Signal Spatio-Temporal Logic* (SSTL) [7,18,19]. SSTL is a linear-time temporal logic to describe behaviours of traces generated from simulations (or even measured from real systems). The logic integrates the temporal modalities of STL (*Signal Temporal Logic*) [15] with two spatial operators: *somewhere* and *surrounded* which enable specifying properties over discrete space models. Given a SSTL property, SSTL model checking automatically checks its satisfaction by exhaustively exploring all the data points (w.r.t, space and time) in the simulated trace, in order to identify spatial patterns and structures of interest over a time series. To date, spatial and spatio-temporal logics have been successfully applied to various systems, examples include identifying vehicular movement in public transport systems [6], monitoring mobile ad-hoc sensor network [3], identifying

diffusion pattern [19], and specifying spatio-temporal patters in particle-based simulation [22].

In this paper, we apply SSTL model checking to analyse the dynamics of 3D individual-based biofilm simulations. Biofilms are communities of microorganisms encased in a self-produced extracellular matrix where microbes stick to each other or to a surface. This structural complexity provides both biological and mechanical stability for the biofilms against environmental stress such as nutrient limitation or shear forces due to fluid flow [16]. Understanding the change of biofilm structural characteristics in response to changing environments can therefore yield essential information to design and maintain biofilm-related applications, such as waster water treatment or bioremediation where the spatial dynamics of biofilms dramatically affects their ability to remove toxic pollutants [23,24]. In this work, we first use an IbM solver to model and simulate two biofilm systems: biofilm growth in quiescent environment, and biofilm deformation and detachment under fluid force. Then we show how to use SSTL model checking to analyse various spatial and temporal aspects of the biofilm systems. In particular, we evaluate the effect of nutrient availability on biofilm surface structure, and the effect of fluid strength (shear rate) on biofilm streamer formation and detachment.

2 Methods

This section gives a brief overview of the computational methods as well as the software that we used for the modelling, simulation and model checking of biofilm systems – more details can be found in [10,14,19,20].

2.1 Individual-based Model

A mechanical individual-based model developed in our previous work is applied for modelling biofilm systems [10]. The model combines fundamental biological processes and physical interactions to simulate the growth of microbes and their response to fluid flow at the micro-scale.

Figure 1 gives an overview of the model framework. Individual microbes are represented as rigid spheres, with each microbe having a set of attributes, including position, diameter, outer-diameter, biomass, outer-mass, velocity, force, growth rate, yield coefficient, and genotype. Some of the attributes vary among individuals and can change through time, others are constant throughout a simulation (*e.g.*,yield coefficient and genotype). Outer-diameter and outer-mass represent an EPS (Extracellular Polymeric Substances) shell. EPS are biopolymers that play a major role in keeping the mechanical stability of biofilms. They are initially accumulated as an extra shell around the microbes before being excreted to the environment. Microbial functional classes (or genotypes, such as heterotrophs, ammonia oxidizing bacteria, nitrifying oxidizing bacteria, *etc.*) are groups of one or more individual microbes that share the same characteristics

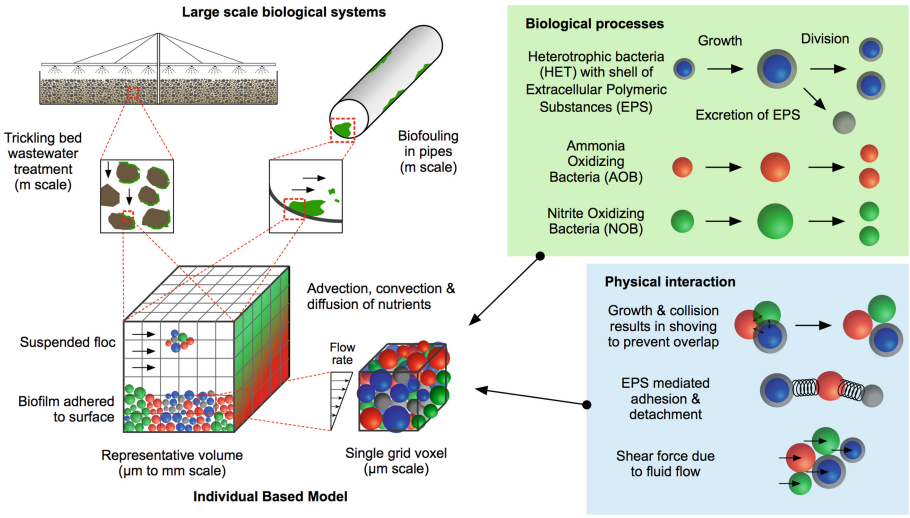

Fig. 1. Overview of the individual-based biofilm model. The micro-scale computational domain represents a small volume of the large-scale biological systems. The attributes of the microbes in the domain are govern by biological, chemical, and physical processes. Figure adopted from [10]

or parameters. For the sake of simplification, a mono-species biofilm is considered in this work that consists of heterotrophic bacteria (HET) and their EPS production. Moreover, the excreted EPS are represented as spheres rather than a continuum matrix structure.

The computational domain is the environment where microbes reside and the biological, physical and chemical processes take place. It is defined as a micro-scale 3D rectangular box. Within the domain, chemical properties such as nutrient concentration or nutrient consumption rate are represented as continuous fields. To resolve their dynamics over time and space, the domain is discretised into Cartesian grid elements so that the values can be calculated at each discrete grid on the meshed geometry. The style of domain boundary can be defined as either periodic or fixed. The former allows microbes to cross the boundary, and re-appear on the opposite side of the domain, while a fixed wall prevents microbes to interact across the boundary or move from one side of the domain to the other. Microbial attributes are governed by both inter-cellular and intra-cellular processes. We consider the following processes that capture the essential behaviours of biofilms and their response to fluid flow.

Microbial Growth. An individual microbe grows and its mass and outer-mass increase by consuming nutrients in the grid where the microbe resides. The growth of heterotrophic bacteria is calculated based on the Monod equation described in [21]. The new mass and outer mass are then used to update diameter and outer-diameter of the microbe, respectively.

Microbial Division. Cell division is the result of microbial growth and is considered as an instantaneous process. Division occurs if the diameter of a microbe reaches 1.35 µm; the cell then divides into two daughter cells. The total mass of the two daughter cells is always conserved from the parent cell. One daughter cell is (uniformly) randomly assigned between 40% and 60% of the parent cell's mass, and the remaining mass is assigned to the other daughter cell. Moreover, one daughter cell takes the position of the parent cell while the other daughter cell is randomly placed next to the first one.

EPS Production. Heterotrophs can secrete EPS into their neighbouring environment. Initially, EPS is accumulated as an extra shell around the secreting microbe. When the relative thickness of the EPS shell of the microbe exceeds a certain threshold value (outer-diameter/diameter >1.3), around half (uniformly random ratio between 0.4–0.6) of the EPS mass excretes as a separate EPS particle and is (uniformly) randomly placed next to the microbe.

Mechanical Relaxation. When microbes grow and divide, the system may deviate from mechanical equilibrium (*i.e.*, non-zero net force on microbes) due to microbe overlap or collision. Hence, mechanical relaxation is required to update the location of the microbes and minimise the stored mechanical energy of the system. Mechanical relaxation is carried out using the discrete element method, and the Newtonian equations of motion are solved for each microbe in a Lagrangian framework [9]. The model considers three forces: 1) The contact force is a pair-wise force exerted on the microbes to resolve the overlap problem at the individual level. The force equation is solved based on Hooke's law, as described in [5]; 2) The EPS adhesive force is also a pair-wise interaction imposed by EPS to attract nearby microbes. The force is modelled as a spring, with the spring coefficient being proportional to the combined EPS mass of the two individuals [10]; 3) The drag force is the interaction of fluid and particulate microbes, which is simplified by modelling one way coupling, *i.e.*, only the effect of the fluid on the microbes is considered, the flow field is not affected by the movement of microbes. In this work, the force is based on Stokes flow past a sphere [10].

Nutrient Mass Balance. Nutrient distribution within the 3D computational domain is calculated by solving the advection-diffusion-reaction equation for soluble substrates. The transport equation is discretised on a Marker-And-Cell (MAC) uniform grid and the scalar is defined at the centres of the grid (cubic element). The temporal and spatial derivatives of the transport equation are discretised by Forward Euler and Central Finite Differences, respectively. The equation is solved for the steady state solution of the concentration fields.

2.2 Signal Spatio-Temporal Logic

The *Signal Spatio-Temporal Logic* (SSTL) is a spatial extension of the Signal Temporal Logic (STL) [15]. The logic allows specifying spatio-temporal properties over discrete space and continuous time series generated during the simulation of a (stochastic) complex system.

Discrete Space Model. In SSTL, the discrete space representation is modelled as a weighted undirected graph. Formally, a weighted undirected graph is a tuple $G = (L, E, w)$ where L is a finite set of nodes (locations), $E \subseteq L \times L$ is a finite set of edges (connections between nodes), and $w : E \to R_{>0}$ is a function that returns the positive weight of each edge. In addition, the discrete space is also equipped with a metric, which is a function that gives the shortest weighted path distance between each pair of element in L, *i.e.*,the shortest path between any two different locations.

Signal and Trace. A *spatio-temporal signal* is a function $\overrightarrow{s} : \mathbb{T} \times L \to \mathbb{E}$ where $\mathbb{T}$ is a dense real interval representing time, L is the set of locations, and the domain of evaluation $\mathbb{E}$ is a subset of $\mathbb{R}^* = R \cup \{+\infty, -\infty\}$. Signals can be described in either a qualitative or a quantitative way: those with $\mathbb{E} = \mathbb{B} = \{0, 1\}$ are qualitative Boolean signals, whereas signals with $\mathbb{E} = \mathbb{R}^*$ are quantitative real signals. A *spatio-temporal trace* is a function $\mathbf{x} : \mathbb{T} \times L \to \mathbb{R}^n$, s.t. $\mathbf{x}(t, \ell) := (\mathbf{x}_1(t, \ell)), ..., \mathbf{x}_n(t, \ell) \in \mathbb{D} \subseteq \mathbb{R}^n$ where each $\mathbf{x}_i : \mathbb{T} \times L \to \mathbb{D}_i \subseteq \mathbb{R}$, for $i = 1, ..., n$ is the projection on the i^{th} coordinate/variable. Intuitively, each trace is a unique simulation trajectory containing both temporal and spatial information of the simulated system.

SSTL. The syntax of SSTL is given by:

$$\varphi \;:=\; \mu \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \mathcal{U}_{[t_1,t_2]} \varphi_2 \mid \diamondsuit_{[d_1,d_2]} \varphi \mid \varphi_1 \mathcal{S}_{[d1,d2]} \varphi_2$$

where μ is an atomic predicate, negation and conjunction are the standard Boolean connectives, and $\mathcal{U}_{[t_1,t_2]}$ is the temporal *until* operator, where $[t_1, t_2]$ is a real positive closed interval with $t_1 < t_2$ representing time. A trace satisfies the until formula if φ_2 is satisfied at some time point within the interval $[t_1, t_2]$ and φ_1 is true up until that point. Additional temporal operators can be derived as syntactic sugar:

- the *eventually* operator $\mathcal{F}$ where $\mathcal{F}_{[t_1,t_2]}\varphi := \mathit{true}\ U_{[t_1,t_2]}\varphi$, and
- the *always* operator $\mathcal{G}$ where $\mathcal{G}_{[t_1,t_2]}\varphi := \neg\mathcal{F}_{[t_1,t_2]}\neg\varphi$.

Intuitively, $\mathcal{F}_{[t_1,t_2]}\varphi$ expresses that φ is eventually satisfied at some time point in the $[t_1, t_2]$ interval, whereas a trace satisfies $\mathcal{G}_{[t_1,t_2]}\varphi$ if φ is true for every time point in $[t_1, t_2]$.

The spatial operators $\diamondsuit_{[d_1,d_2]}$ and $\mathcal{S}_{[d1,d2]}$ are *somewhere* and *surrounded*, respectively. The former means that φ holds in a location reachable from the current one with a distance between d_1 and d_2, whereas the latter is satisfied by locations in a φ_1-region, and surrounded by φ_2-region at a distance between d_1 and d_2. SSTL has both classical Boolean semantics and quantitative semantics. The former returns true or false depending on the satisfaction of a SSTL formula, whereas the quantitative semantics returns a real value that 'measures' how robustly a formula is satisfied (or not). The formal SSTL semantics and the algorithms are detailed in [19].

2.3 Software

We use the NUFEB[1] software [14] for the modelling and simulation of (individual-based) biofilm system, and the jSSTL[2] library [20] for the specification and verification of SSTL properties of simulation traces produced by NUFEB.

NUFEB [14] (Newcastle University Frontiers in Engineering Biology) is a 3D, open-source, massively parallel simulator for individual-based modelling of microbial communities. The tool is built on top of the state-of-the-art software LAMMPS (Large-scale Atomic Molecular Massively Parallel Simulator) [26] extended with IbM features. NUFEB allows flexible microbial model development from a wide range of biological, chemical and physical processes, as well as individual microbes types via an input script. The tool supports parallel computing on both CPU and GPU facilities based on domain decomposition, which enables simulating large number of microbes (beyond millions of individuals). During a simulation, NUFEB can output any state variable of microbes or grids. The output results can be stored into various formats for visualisation or analysis.

In this work, we extend NUFEB with new functions for coupling with SSTL. In particular, the SSTL space model is based on the existing NUFEB mesh structure. SSTL-based attributes such as volume fraction will be calculated at each discrete cubic grid (so-called *SSTL grid*). A new output format is implemented for dumping the space model as well as simulation traces during simulation. A new post-processing routine is also implemented to read and visualise model checking results from the SSTL model checker.

jSSTL [20] is a Java-based tool for SSTL property specification and model checking. The tool takes three types of input written in CSV or tabular based ASCII files: a SSTL formula file, a SSTL space model file, and a spatio-temporal trace file. The latter two are obtained from (NUFEB) simulations, while the formula file is specified by the user. jSSTL can compute both the Boolean and the quantitative spatio-temporal semantics of a SSTL formula at each time point and in each location. The tool also provides a simple user interface as an Eclipse plug-in to specify and verify SSTL properties.

jSSTL utilises the Floyd-Warshall algorithm to compute the shortest path for each location pair in a weighted directed space model. The space model in NUFEB is a simple lattice-like mesh structure restricted to orthogonal box (see, for example, the computational box in Fig. 1). We therefore extend jSSTL with a new algorithm specifically for computing the distance matrix of our NUFEB space model. The algorithm takes the indexes of each node pair and directly compute the distance (*i.e.*, the number of SSTL grids between the two nodes) based on a 3D box. Our algorithm takes less than three seconds to build the distance matrix of a $30 \times 12 \times 24$ mesh – the default algorithm takes instead

[1] https://github.com/nufeb.
[2] https://github.com/Quanticol/jsstl.

more than 20 min. All the codes used and developed in this work are publicly available[3].

3 Results

In this section, we address SSTL model checking on spatio-temporal emergent properties of biofilm systems. We first apply the technique to study the surface morphology of a biofilm growing in a quiescent medium. Then we use SSTL to analyse biofilm streamer formation and detachment in a fluid environment.

3.1 Biofilm Surface Morphology

Biofilms can sense nutrient concentration gradients by adjusting their surface structure. Understanding biofilm morphology in relevant environmental conditions is therefore essential to predict biofilm effects in many practical applications. For example, biofilm morphology is thought to be crucial for the emergence of mutations conferring antimicrobial resistance. Again, biofilms with irregular surface (thus large surface area) can significantly improve their performance in wastewater treatment processes. In this section, we use SSTL to evaluate the influence of nutrient gradients on biofilm structure in a quiescent growth medium.

Table 1. Key parameters and IbM processes used in the biofilm growth model

Parameters and settings	Value	Unit
Parameters		
Dimensions	$100 \times 40 \times 80$	μm
Cartesian grid elements	$30 \times 12 \times 24$	Grids
SSTL grid elements (Property 1)	$30 \times 12 \times 24$	Grids
Initial microbes	40	
SSTL time point (dt)	1000	Seconds
Simulation time	9.5	Days
Substrate bulk concentration	0.1–0.3	g/m^3
IbM processes		
Biology: Microbial growth, division, EPS production		
Chemistry: nutrient mass balance		
Physics: contact force, EPS adhesion		

Biofilm Growth Model. A mono-species biofilm model (HETs and their EPS production) is developed for modelling biofilm formation and its spatial

[3] https://github.com/shellbw/NUFEB-sstl.

dynamics. Table 1 highlights the key parameters and the IbM processes used in the model. Initially, 40 heterotrophs are inoculated on the substratum of a $100 \times 40 \times 80\,\mu m$ computational domain. We assume that the bulk environment is situated at the top and supplies a constant bulk nutrient concentration to the computational domain. A nutrient concentration gradient can be therefore formed due to the diffusion as well as the consumption by microbes. Two independent simulations are then carried out to simulate biofilm growth in different bulk nutrient concentrations ($0.1\,g/m^3$ and $0.3\,g/m^3$). In the nutrient-limited environment, a mushroom-shaped biofilm structure results from the competition between microbes at the biofilm surface (Fig. 2(a)). Biofilms growing in a nutrient-rich environment form instead a smooth surface (Fig. 2(b)).

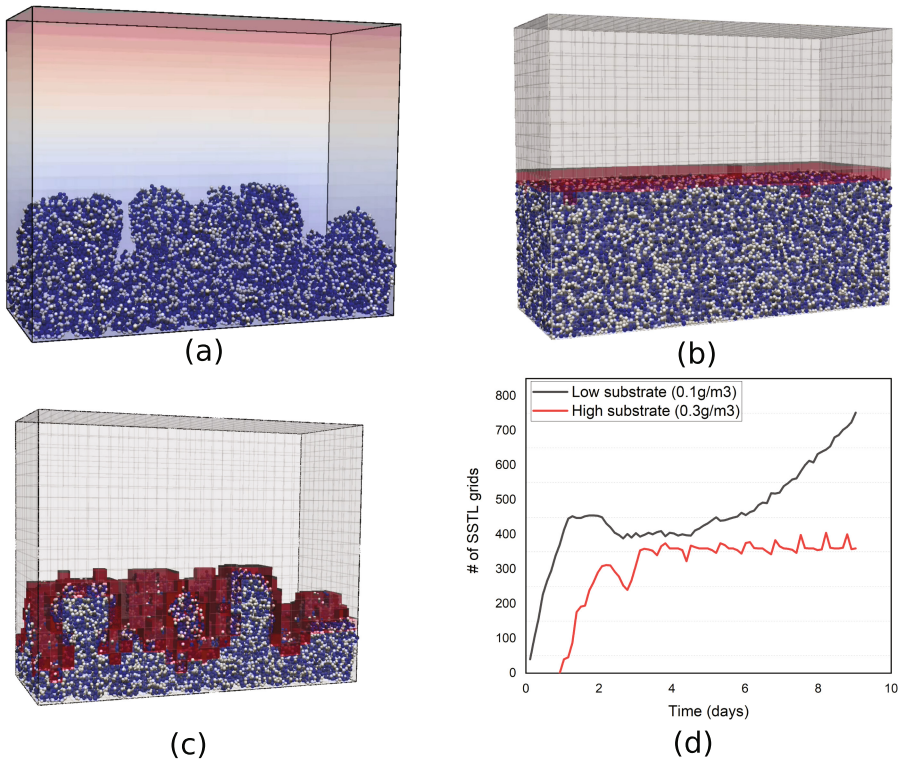

Fig. 2. Simulation and model checking results of biofilm formation at 9.5 days: (a) biofilm growing in nutrient-limited condition (max number of microbes $= 4.2 \times 10^4$; CPU time = 4 mins (four cores)), (b) biofilm growing in nutrient-rich condition (max number of microbes $= 8.1 \times 10^4$; CPU time = 11 mins (four cores)) with the SSTL grids satisfying Property 1 highlighted, (c) satisfaction of Property 1 for the biofilm simulation in nutrient-limited condition, and (d) number of SSTL grids satisfying Property 1 under different growth conditions and over time.

Model checking. The SSTL space model is generated at the beginning of each simulation, and its structure is same as the NUFEB mesh structure ($30 \times 12 \times 24$ grids), *i.e.*, 24 layers with 360 grids in each layer. The spatio-temporal trace $\mathbf{x}$, generated during the simulation, consists of a set of signals of microbial volume fraction *VolFrac* at each SSTL grid and at each time point (1,000 s), where *VolFrac* is defined as the total volume of microbes in the SSTL grid divided by the volume of the grid. The following SSTL formula characterises biofilm surface structure:

Property 1 (surface structure):

$$\mathcal{F}_{[t_i,t_i]}[\mathit{VolFrac} > 0 \wedge \diamondsuit_{[0,1]}(\mathit{VolFrac} = 0)]. \tag{1}$$

The above formula states that "*an SSTL grid is considered biofilm surface at the time point t_i if there exists at least one microbe in the grid and at least one of its neighbour grid does not have microbes inside*", where t_i is i^{th} simulation time point with $t_0 = 0, t_1 = dt, t_2 = 2 \times dt, ..., t_{800} = 800 \times dt$. SSTL allows nesting temporal and spatial operators to express complex structure dynamics. In the formula, we use the *somewhere* operator $\diamondsuit$ to identify the edge of a biofilm by evaluating the volume fractions *VolFrac* in each target grid as well as its neighbouring grids, while the *eventually* operator describes the dynamics of the biofilm surface *over time*.

Figures 2(b) and (c) illustrate the model checking result with all the SSTL grids satisfying the formula highlighted. It is clear that the formula is capable of identifying which grids belong to the biofilm surface and which do not. By plotting the number of satisfied SSTL grids over time (Fig. 2(d)), we are able to quantitatively analyse the biofilm surface dynamics. In both cases, the satisfied grids increase rapidly at early stage of the simulations, indicating the substratum coverage by biofilm from the 40 initial microbes. Then, the number of surface grids of the biofilm growing in the nutrient-rich condition reaches a stable state (between 360 and 380 grids), which suggests a tendency towards flat surface structure. The fluctuations are due to the formation of small bumps during growth. For the biofilm growing in the nutrient-limited condition, the surface grids increase continuously without reaching a stable state. At 9.5 days, the number of surface grids is more than twice of the layer grids (360) indicating a high biofilm surface porosity.

Given the model checking result, it is also possible to estimate additional biofilm surface properties such as biofilm surface area (*i.e.*,the number of satisfied grids multiplied by the single surface area of the grid). In the above example, the surface areas at 9.5 day are 1,225 μm^2 ($368 \times 3.33\,\mu m^2$) and 2,497 μm^2 ($750 \times 3.33\,\mu m^2$) in the two systems, respectively.

3.2 Biofilm Deformation and Detachment

Biofilms attached to surfaces under fluid flow behave in a complex way due to fluid-microbe and microbe-microbe interactions. Understanding how biofilms interact with the fluid can help unravel their survival mechanisms, which is essential for biofilm removal or preservation in practical applications. For example,

the accumulation of biofilms in industrial pipelines may lead to biocorrosion, while their removal can be achieved by the application of hydrodynamic shear forces [1,8]. Here we apply IbM and SSTL to model and analyse two important spatial emergent behaviours of hydrodynamic biofilms: streamer formation and detachment. A biofilm streamer is a filament-like biofilm structure that can cause rapid and catastrophic clogging in biomedical systems. Biofilm detachment, on the other hand, is essential for removing biofilms.

Table 2. Key parameters and IbM processes used in the hydrodynamic biofilm model

Parameters and settings	Value	Unit
Parameters		
Dimensions	$200 \times 40 \times 100$	μm
Cartesian grid elements	$24 \times 8 \times 20$	Grids
SSTL grid elements (Property 2 & 4)	$15 \times 8 \times 20$	Grids
SSTL grid elements (Property 3)	$1 \times 8 \times 20$	Grids
Simulation time	4×10^5	Seconds
SSTL time point	1800	Seconds
Shear rate	$0.15 - 0.3$	s^{-1}
IbM processes		
Physics: contact force, EPS adhesion, shear force		

Biofilm Hydrodynamic Model. To model hydrodynamic biofilms, we apply shear force to a pre-grown biofilm that consists of heterotrophs and their EPS production. For the sake of simplification, the model does not consider biological activities during the fluid stage. The presence of EPS imposes adhesion to the microbes whereas the shear force drives microbe motion along the flow direction (+x). We simulate the model with various shear rates γ in order to study the effect of shear rate on the biofilm structure (Table 2). Figures 3(a)-(c) illustrate the biofilm dynamics at the rate $\gamma = 0.2\,s^{-1}$. In the early stage, the biofilm forms short streamers, and small microbial clusters detach from the head of the streamer due to cohesive failures. As the fluid continues to flow, the top of the biofilm is highly elongated. Although large chunks of detached microbes can be observed at this stage, the streamer continuously grows to a significant length and maintains a stable condition.

Model Checking. We now show how to use SSTL model checking to analyse the formation of biofilm streamers and to quantify their spatial properties. The space model is defined as a $15 \times 8 \times 20$ mesh structure which covers the downstream area and excludes the biofilm body, as shown in Fig. 3. The following formula is used to evaluate whether a SSTL grid is part of a biofilm streamer:

Property 2 (Streamer Formation):

$$\mathcal{F}_{[t_i,t_i]}\mathbf{G}_{[0,t_{\max}]}(\mathit{VolFrac} > 0). \tag{2}$$

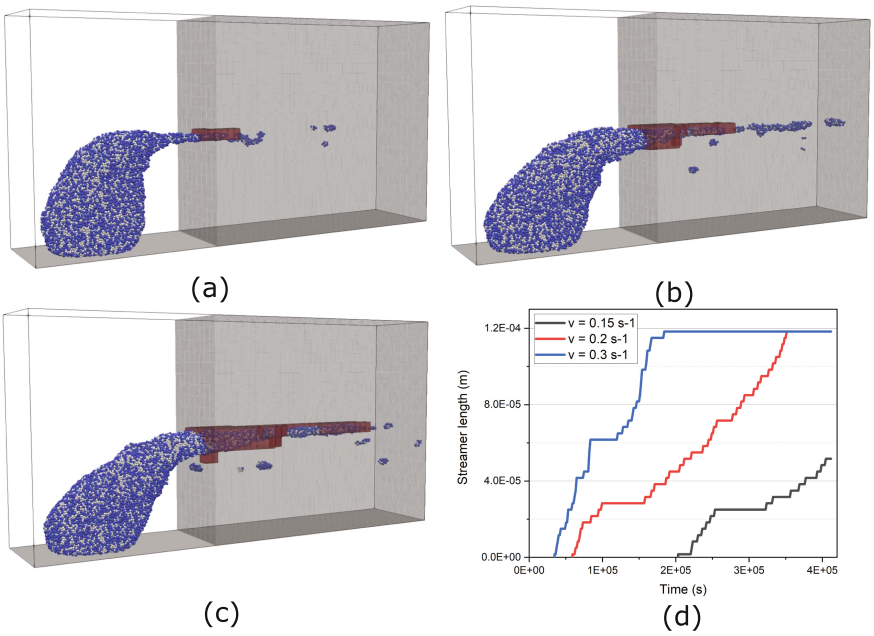

Fig. 3. Simulation of biofilm under shear force and satisfaction of Property 2 with $\gamma = 0.2\,\mathrm{s}^{-1}$ and $t_{\max} = 36{,}000$, at (a) $t_i = 1 \times 10^5$ s, (b) $t_i = 2 \times 10^5$ s, and (c) $t_i = 3 \times 10^5$ s; (d) Streamer length temporal dynamics under different shear rates. Max number of microbes $= 4.6 \times 10^4$ (simulation start); min number of microbes $= 3.2 \times 10^4$ (simulation end); CPU time = 32 mins (four cores)

The formula intuitively means that "*an SSTL grid is considered biofilm streamer at time point t_i if the grid is occupied by microbes continuously for the next $t_{\max}$ time units*". The formula is checked at each SSTL grid over a series of simulation time points t_i, with $t_{\max} = 36{,}000$ (*i.e.*,10 h to ensure the continuity of microbes passing grids). Figures 3 (a)-(c) visualise the model checking results for the simulation trace $\mathbf{x}$ generated for the $\gamma = 0.2\,\mathrm{s}^{-1}$ case. It can be seen that the satisfied grids (highlighted) are in agreement with the streamer structure. The grids occupied by the detached clusters do not satisfy the property as there will be a time point when the clusters leave such grids. In Fig. 3(d), we evaluate the effect of shear rates on streamer formation by plotting the satisfied grid with largest coordinate in x direction over time. As expected, when the shear rate increases the streamer formation rate also increases. For example, when $\gamma = 0.3\,\mathrm{s}^{-1}$ the streamer length reaches 120 µm after 1.6×10^5 s, whereas the streamer takes twice of the time to reach the same length when $\gamma = 0.2\,\mathrm{s}^{-1}$.

Property 2 combines the temporal operator *always* and the geometry of SSTL grids to identify the spatial structure of streamers. This expression is different from Property 1 where the biofilm surface structure is described by a spatial operator. A similar idea can be also applied to detect biofilm detachment events.

In such case, we define the space model as a single layer wall located at the downstream side as shown in Fig. 4(a). We use the following SSTL (temporal) formula to check whether SSTL grids are occupied by detached clusters.

Property 3 (Biofilm Detachment):

$$\mathcal{F}_{[t_i,t_i]}[\mathit{VolFrac} > 0 \wedge \mathcal{F}_{[0,t_{\max}]}(\mathit{VolFrac} = 0)]. \tag{3}$$

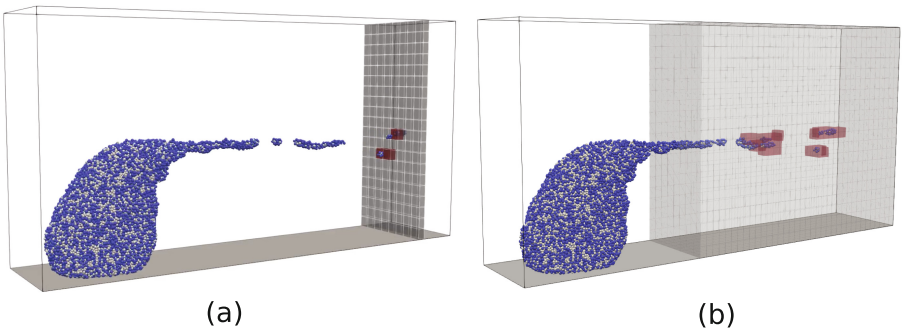

Fig. 4. (a) Satisfaction of Property 3 with $\gamma = 0.2\,\mathrm{s}^{-1}$, $t_i = 1.5{\times}10^5$ s, and $t_{\max} = 3{,}600$ (b) Satisfaction of Property 4 with $t_i = 1.5 \times 10^5$ s, $[\mathrm{d}_1, \mathrm{d}_2] = [0, 3]$

The formula states that "*an SSTL grid is considered detached cluster at time point t_i if the grid is occupied by microbes, but it will become free in the next $t_{\max}$ time units*". Figure 4(a) illustrates an example of the model checking result. At each time point t_i, the formula is satisfied at any grid in the wall if its volume fraction is greater than zero but will become zero in the future. The second argument makes sure microbes crossing the grid are (part of) detached clusters rather than continuous streamer. The use of wall-like grid structures allows us to analyse the frequency of detachment events over time. In Fig. 5, we show the effect of shear rates on detachment frequency by plotting the Boolean satisfaction of Property 3 with respect to all grids at each time point (*i.e.*,for each time point in $\mathbf{x}$, record true if there exists at least one grid satisfying the formula, otherwise record false). The result shows that the detachment frequency increases as the shear increases, thus indicating a positive correlation between biofilm cohesive failure and fluid strength. This is in agreement with both experimental observations [25] and simulations using other approaches [27]. Moreover, by recalling the streamer length result shown in Fig. 3 (d), we can conclude that biofilm deformation and detachment are closely related, as the occurrence of detachment is consistent with the appearance of streamers.

The above property can identify whether incoming microbes are (or are part of) a detached cluster, but it is unable to accurately quantify physical attributes of clusters, such as volume or shape. Inspired by [19] where the authors use the

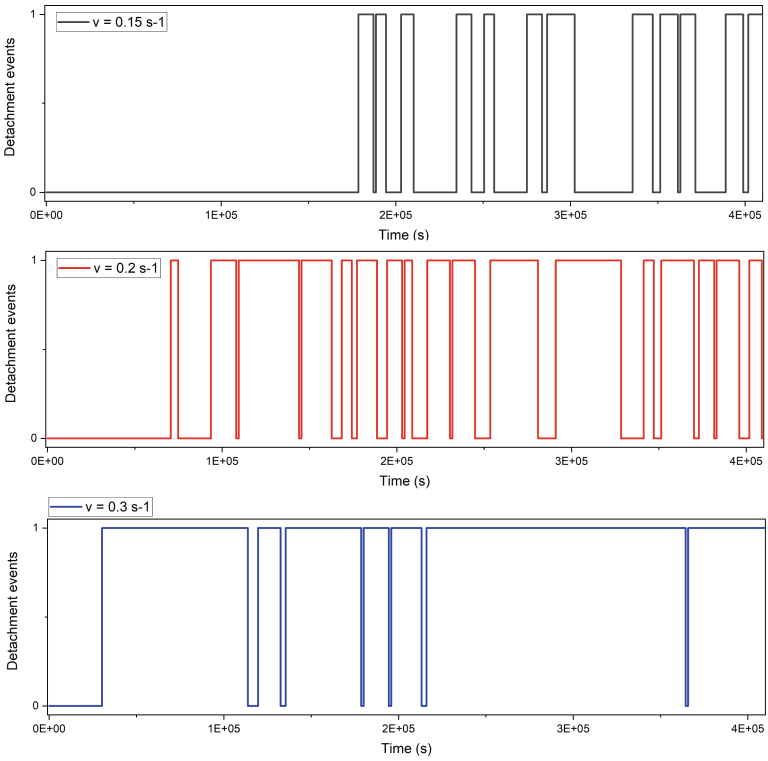

Fig. 5. Boolean evaluation of Property 3: detachment under different shear rates

surrounded operator $\mathcal{S}$ to characterise diffusion pattern in 2D, we introduce the following SSTL formula to identify the geometry of any detached cluster.

Property 4 (Detached Cluster):

$$\mathcal{F}_{[t_i,t_i]}[(\mathit{VolFrac} > 0)\mathcal{S}_{[d_1,d_2]}(\mathit{VolFrac} = 0)]. \tag{4}$$

Informally, the formula checks *"whether there exists a sub-region of the grids such that all the grids of the sub-region are occupied by microbes (i.e., VolFrac* > 0*). Furthermore, the sub-region is surrounded by grids without microbes inside (i.e., VolFrac* $= 0$*)"*. Note that the use of distance bounds $[d_1, d_2]$ in the surrounded operator allows us to constrain the cluster size. For example, Fig. 4 (b) reports the Boolean satisfaction of the formula in each grid for the $\gamma = 0.2\,\mathrm{s}^{-1}$ simulation, with $[d_1, d_2] = [0, 3]$. As shown in the figure, only the grids containing detached clusters satisfy the formula, whereas those containing continuous streamers do not. It is worth to note that the granularity of the SSTL mesh structure is important in spatial model checking. In this case, for example, a more precise cluster structure can be captured when a fine mesh is used. Such accuracy, however, can be significantly offset by the *state space explosion prob-*

lem when handling large 3D space models, and it is therefore beyond the scope of the current work.

3.3 Performance

Several performance statistics of the SSTL model checking are given in Table 3. As previously mentioned, the SSTL distance matrix needs to be built just once and takes a few seconds only. The verification of properties involving spatial operators, in particular the *surrounded* operator in Property 4, can be computationally more expensive compared to properties with temporal operators only (Property 2 and 3). All the simulations and model checking analyses reported in this paper were carried out on a Linux system with a 3.4 GHz Intel Core i5 processor and 12 GB RAM.

Table 3. CPU time and memory usage for building distance matrices and model checking Properties 1–4 (single core)

	Property 1	Property 2	Property 3	Property 4
Distance matrix (s)	2.1	N/A	N/A	2.8
Memory usage (MB)	789	N/A	N/A	823
Model checking (s)	18.5	35.3	47.1	1,161
Memory usage (MB)	1,357	1,430	1,329	2,513

4 Conclusion

In this paper we utilised SSTL model checking to identify dynamic spatio-temporal behaviours arising from individual-based simulations. We first extended our previous developed IbM (Individual-based Model) solver NUFEB to support additional output data for SSTL (Spatio-Temporal Logic) model checking. We added a new, more efficient algorithm in the SSTL model checker jSSTL specifically for computing the distance matrix of the NUFEB mesh structure. Then we used SSTL to specify complex properties of two biofilm systems including dynamics of biofilm surface morphology, biofilm streamer formation and biofilm detachment under shear force. The model checking results demonstrated that SSTL can capture and analyse such behaviours.

The advantages of using SSTL for post-hoc analysis of 3D IbM simulations are two-fold. First, the approach is capable of expressing complex spatio-temporal properties in a relatively simple and succinct logic language that can be model checked in a robust and automatic way. Such a formal verification is not only more reliable than manual inspection of simulation traces, but it can also greatly reduce the complexity of post-processing simulation data. Second, while in this work we integrated the NUFEB solver with jSSTL, the model checker

could be combined with minor modifications with other IbM tools such as iDynoMiCS [13], Simbiotics [17], and BioDynaMo [4], since they are all equipped with a Cartesian grid mesh structure which can be directly used with SSTL.

In the current work, we focused on off-line (post-hoc) verification of SSTL properties. In the future, it would be interesting to explore the use of real-time SSTL verification, where SSTL can be interfaced with NUFEB or other IbM simulators and alert the user if a SSTL property is violated (or satisfied) during a simulation. Such runtime verification is particularly important for monitoring large individual-based simulations (which can run for days or weeks on high-performance systems) by providing instant feedback on the correctness of the model. Moreover, SSTL can quantify the degree with which a simulation trace satisfies or violates a property. Such "robustness" information could be useful to assess quantitatively the severity of unexpected behaviour in a simulation, thereby increasing the confidence in the analysis of the simulation traces.

Acknowledgements. This work was supported by the Impact Acceleration Account award 'BioHPC: Simulating Microbial Communities on High-Performance Computers' to Newcastle University, funded by the EPSRC (UK).

References

1. Abe, Y., Skali-Lami, S., Block, J.C., Francius, G.: Cohesiveness and hydrodynamic properties of young drinking water biofilms. Water Res. **46**(4), 1155–1166 (2012)
2. Allan, V., Callow, M., Macaskie, L., Paterson-Beedle, M.: Effect of nutrient limitation on biofilm formation and phosphatase activity of a citrobacter sp. Microbiology. **148**, 277–288 (2002). https://doi.org/10.1099/00221287-148-1-277
3. Bartocci, E., Bortolussi, L., Loreti, M., Nenzi, L.: Monitoring mobile and spatially distributed cyber-physical systems, pp. 146–155 (2017). https://doi.org/10.1145/3127041.3127050
4. Breitwieser, L., et al.: BioDynaMo: an agent-based simulation platform for scalable computational biology research (2020). https://doi.org/10.1101/2020.06.08.139949
5. Brilliantov, N., Spahn, F., Hertzsch, J.M., Poschel, T.: Model for collisions in granular gases. Phys. Rev. E. **53**, 5382 (2002). https://doi.org/10.1103/PhysRevE.53.5382
6. Ciancia, V., Gilmore, S., Grilletti, G., Latella, D., Loreti, M., Massink, M.: Spatio-temporal model checking of vehicular movement in public transport systems. Int. J. Softw. Tools Technol. Transfer **20**(3), 289–311 (2018). https://doi.org/10.1007/s10009-018-0483-8
7. Ciancia, V., Latella, D., Loreti, M., Massink, M.: Specifying and verifying properties of space. In: Diaz, J., Lanese, I., Sangiorgi, D. (eds.) TCS 2014. LNCS, vol. 8705, pp. 222–235. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-44602-7_18
8. Donlan, R.: Biofilms: microbial life on surfaces. Emerg. Infect. Dis. **8**, 881–90 (2002). https://doi.org/10.3201/eid0809.020063
9. Ghaboussi, J., Barbosa, R.: Three-dimensional discrete element method for granular materials. Int. J. Numer. Anal. Meth. Geomech. **14**(7), 451–472 (1990). https://doi.org/10.1002/nag.1610140702

10. Jayathilake, P.G., et al.: A mechanistic individual-based model of microbial communities. PLoS ONE **12**(8), 1–26 (2017). https://doi.org/10.1371/journal.pone.0181965
11. Klapper, I., Dockery, J.: Mathematical description of microbial biofilms. SIAM Rev. **52**, 221–265 (2010). https://doi.org/10.1137/080739720
12. Kreft, J.U., Picioreanu, C., Wimpenny, J.W.T., van Loosdrecht, M.C.M.: Individual-based modelling of biofilms. Microbiology **147**(11), 2897–2912 (2001). https://doi.org/10.1099/00221287-147-11-2897
13. Lardon, L., et al.: iDynoMiCS: next-generation individual-based modelling of biofilms. Environ. Microbiol. **13**, 2416–34 (2011). https://doi.org/10.1111/j.1462-2920.2011.02414.x
14. Li, B., et al.: NUFEB: a massively parallel simulator for individual-based modelling of microbial communities. PLoS Comput. Biol. **15**, e1007125 (2019). https://doi.org/10.1371/journal.pcbi.1007125
15. Maler, O., Nickovic, D.: Monitoring temporal properties of continuous signals. In: Lakhnech, Y., Yovine, S. (eds.) FORMATS/FTRTFT -2004. LNCS, vol. 3253, pp. 152–166. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-30206-3_12
16. Nadell, C., Drescher, K., Wingreen, N., Bassler, B.: Extracellular matrix structure governs invasion resistance in bacterial biofilms. ISME J. **9**, 1700–1709 (2015). https://doi.org/10.1038/ismej.2014.246
17. Naylor, J., et al.: Simbiotics: a multiscale integrative platform for 3D modeling of bacterial populations. ACS Synth. Biol. **6**, 11941–120 (2017). https://doi.org/10.1021/acssynbio.6b00315
18. Nenzi, L., Bortolussi, L.: Specifying and monitoring properties of stochastic spatio-temporal systems in signal temporal logic. EAN Endors. Trans. Cloud Syst. **19**, e4 (2015). https://doi.org/10.4108/icst.valuetools.2014.258183
19. Nenzi, L., Bortolussi, L., Ciancia, V., Loreti, M., Massink, M.: Qualitative and quantitative monitoring of spatio-temporal properties with SSTL. Log. Methods Comput. Sci. **14**(4) (2018). https://doi.org/10.23638/LMCS-14(4:2)2018
20. Nenzi, L., Bortolussi, L., Loreti, M.: jSSTL - A Tool to Monitor Spatio-Temporal Properties (2017). https://doi.org/10.4108/eai.25-10-2016.2266978
21. Ofiţeru, I., Bellucci, M., Picioreanu, C., Lavric, V., Curtis, T.: Multi-scale modelling of bioreactor-separator system for wastewater treatment with two-dimensional activated sludge floc dynamics. Water Res. **50**, 382–395 (2013). https://doi.org/10.1016/j.watres.2013.10.053
22. Ruscheinski, A., Wolpers, A., Henning, P., Warnke, T., Haack, F., Uhrmacher, A.: Pragmatic logic-based spatio-temporal pattern checking in particle-based models, pp. 2245–2256 (2020). https://doi.org/10.1109/WSC48552.2020.9383908
23. Sehar, S., Naz, I.: Role of the biofilms in wastewater treatment. In: Dhanasekaran, D., Thajuddin, N. (eds.) Microbial Biofilms, IntechOpen, Rijeka (2016). https://doi.org/10.5772/63499
24. Singh, R., Paul, D., Jain, R.: Biofilms: Implications in bioremediation. Trends Microbiol. **14**, 389–397 (2006). https://doi.org/10.1016/j.tim.2006.07.001
25. Stoodley, P., Wilson, S., Hall-Stoodley, L., Boyle, J.D., Lappin-Scott, H.M., Costerton, J.W.: Growth and detachment of cell clusters from mature mixed-species biofilms. Appl. Environ. Microbiol. **67**(12), 5608–5613 (2001). https://doi.org/10.1128/AEM.67.12.5608-5613.2001

26. Thompson, A.P., et al.: LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, Meso, and continuum scales. Comp. Phys. Comm. **271**, 108171 (2022). https://doi.org/10.1016/j.cpc.2021.108171
27. Xia, Y., Jayathilake, P.G., Li, B., Zuliani, P., Chen, J.: CFD-DEM modelling of biofilm streamer oscillations and their cohesive failure in fluid flow. Biotechnol. Bioeng. **118**, 918–929 (2020). https://doi.org/10.1002/bit.27619