Individual-based modeling unveils pattern formation and population dynamics in microbial aggregates

Cockx, Bastiaan Julien Reinaart

Publication date:
2024

Document Version
Publisher's PDF, also known as Version of record

Individual-based modeling unveils pattern formation and population dynamics in microbial aggregates

Bastiaan Julien Reinaart Cockx

PhD Thesis
January 2024

DTU Sustain

Department of Environmental and Resource Engineering
Technical University of Denmark

Individual-based modeling unveils pattern formation and population dynamics in microbial aggregates

Bastiaan Julien Reinaart Cockx

PhD Thesis, January 2024

Address: DTU Sustain
Department of Environmental and Resource Engineering
Technical University of Denmark
Bygningstorvet, Building 115
2800 Kgs. Lyngby
Denmark

Phone reception: +45 4525 1600

Homepage: https://www.sustain.dtu.dk
E-mail: info@sustain.dtu.dk

Preface

This thesis is based on the work carried out at the Technical University of Denmark, Department of Environmental and Resource Engineering. The project was supported by the Integrated Water Technology (InWaTech) project, which promotes collaborative research between DTU and KAIST, and was performed under the supervision of Professor Barth F. Smets and Associate Professor Borja Valverde Pérez.

The thesis is composed of a synopsis and three research papers. The synopsis introduces the subject "*Individual-based modeling unveils pattern formation and population dynamics in microbial aggregates*", it provides a general overview of the carried out research, while it provides additional detail for the work that is not covered in the three included research papers.

I **Bastiaan J R Cockx**, Tim Foster, Robert J Clegg, Kieran Alden, Sankalp Arya, Dov J Stekel, Barth F Smets, Jan-Ulrich Kreft, 2023. Is it selfish to be filamentous in biofilms? Individual-based modeling links microbial growth strategies with morphology using the new and modular iDynoMiCS 2.0. *PLOS computational biology, in revision.*

II **Bastiaan J R Cockx**, Jan-Ulrich Kreft, Barth F Smets. Force mediated spatial structure in granular biofilm assembly. *Manuscript in preparation*

III Tim Foster, **Bastiaan J R Cockx**, Robert J Clegg, Shirin Moossavi, Kieran Alden, Tariq H Iqbal, Barth F Smets, Jan-Ulrich Kreft. eGUT, a platform enabling agent-based modelling of the gut microbiome and its spatial interactions with the host mucosa. *Manuscript in preparation.*

In addition, the following publication, not included in this thesis, was also concluded during the PhD study.

• Boris Parra, **Bastiaan Cockx**, Veronika T. Lutz, Lone Brøndsted, Barth F. Smets, Arnaud Dechesne. Isolation and Characterization of Novel Plasmid-specific Phages Infecting Bacteria Carrying Diverse Conjugative Plasmids. *Microbiology Spectrum, accepted.*

Acknowledgements

This PhD has been a long journey, and despite it taking longer than I had anticipated, I am grateful for having had the opportunity to meet so many great people. I would like to thank my supervisor Barth F. Smets for giving me the opportunity to attend the DTU PhD school, for his continued support and for never giving up on me. Furthermore, I would like to thank my supervisor Borja Valverde Pérez for proactively stepping in when it was needed and for supporting me through my last mile. I would like to thank Arnaud Dechesne, for being a great and supportive colleague and providing me with excellent feedback on my writing. My special thanks and gratitude go out to Jan-Ulrich Kreft, he has supported me throughout all these years. His patience, expertise and kindness have been fundamental in helping me overcome the challenges I've faced during this project. I am not sure whether I could have overcome them without him by my side.

I would like to thank Robert, Sankalp, Stefan and Tim who have made it a joy to work closely on the development of iDynoMiCS 2.0 and eGUT, and my other co-authors for all their hard work, ideas, inspiration, and support with the manuscripts covered in this PhD thesis. I would like to thank the Integrated Water Technology (InWaTech) project and DTU Sustain for their financial and organizational support and for arranging an excellent PhD school. For creating an environment in which I could develop my skills and expertise which will be a great asset in my future career. It has been an enormous privilege to meet with all the fantastic colleagues and friends I got to know through DTU. They are the people that make it great to be at DTU Sustain. I would like to thank Steffen Foss Hansen who has made it his priority to take away my stress and anxieties. I would like to thank our laboratory technicians, administrative staff, IT support and our maintenance and cleaning staff for their help, but mostly for always greeting me with a smile and creating a pleasant and welcoming work environment. Thanks to all the other colleagues, students and friends I got to know through the METlab research group and DTU Sustain. Alex, Antia, Arda, Boris, Camiel, Camilla, Carlos, Daniela, Dongah, Elena, Elham, Estelle, Fynn, Gizem, Jan, Jane, Lene, Liguan, Maria, Marlene, Marta, Martin, Pauly, Pedram, Qingxian, Ricarda, Sara, Sike, Tal, Uli, Vaibhav and Yunji it would not have been the same without you.

Finally, I would like to thank my family and friends outside of DTU, who have always been there for me, supported me and brought me welcome distractions.

Summary

Microbes are commonly found in aggregated form, where a community of various microbial species fill a variety of ecological niches. They play an essential role in global nutrient cycles, human and animal health, and in numerous industrial applications. Microbial aggregates, in many cases referred to as biofilms, have a dynamic life cycle. The cycle starts when planktonic cells start attaching to a surface or other cells. In this so called reversible attachment phase microbial surface properties such as the hydrophobicity of the cell play a key role. After this the nascent biofilm enters the irreversible attachment phase, more cells continue to be embedded in the growing aggregate, and phenotypic changes associated with microbial life in a biofilm, including the production of large quantities of Extracellular Polymeric Substances (EPS), start to occur. As the aggregate grows, spatial structures such as microcolonies and stratification start to emerge. Mature biofilms may get into a final phase in which EPS degradation occurs and cells disperse into the environment.

The spatial and community structure of a microbial aggregate result from a combination of physical processes such as matter transport and mechanical stresses acting upon the microbial aggregate, and the individual traits and behavior of the microbes embedded in the aggregate. Due to the complexity of multi-species microbial aggregates and our limited ability to observe the dynamic development of internal structures and biochemical conditions in microbial aggregates, modeling approaches have played a substantial role in biofilm research. The first mathematical and computational biofilm models, developed over half a century ago, clarified how, through diffusion limitation, chemical gradients form within microbial aggregates that can in turn, lead to the stratification of microbial species in the aggregate. Later, Individual-based Models (IbM) showed how individual traits and variability can affect the development of the microbial aggregate and how subtle differences can lead to vastly different outcomes.

Over two decades of IbM assisted biofilm research has proven fruitful and the approach is increasingly being used in contemporary research. An increasing number of modeling platforms facilitate the development of IbMs. However, current drawbacks and limitations restrain us from using the methodology in a broader palette of studies. These drawbacks include the limited scale of computationally feasible model systems, the limited availability of individual based observations required to parameterize the model systems, the expertise

required to develop an IbM, and limitations in the ability to accurately represent some important microbial traits or model systems with currently available modeling tools. Recent efforts by multiple research consortia have brought substantial improvements in some of these limitations, but an integrated approach that addresses all these limitations simultaneously is still missing.

Central to my thesis is the development of such an integrated approach, and thereby facilitating a wide range of biofilm research to help explain how individual microbial traits lead to emergent properties of the microbial aggregate. This goal can be subdivided into the following objectives:

1. Enabling a larger range of dynamic individual-based characteristics and behaviors for a better representation of various microbial traits.
2. Widening the range of (bio-)chemical sub-models to provide a better representation of microbial metabolisms.
3. Improving the physical representation and interaction models available.
4. Closing the gap between experimental and modeling work facilitating a better integration of experiments and modeling and reducing the required expertise to construct a model.
5. Facilitating model parameterization where individual based experimental observations may be lacking.
6. Improving computational efficiency to allow for models at larger scales.

To achieve this goal, I have worked with a dedicated group of collaborators on the development of the novel individual-based modeling framework iDynoMiCS 2.0 and the closely related gut modeling software eGUT. The latter expands on iDynoMiCS 2.0 with an epithelium and mucus facilitating gut lumen and mucosa models. iDynoMiCS 2.0 introduces a new structure where modeled microbes, often referred to as agents, can be assembled from orthogonal modules. Any characteristic, including the morphology, biochemical and biophysical behavior, can dynamically change in response to external or internal cues. These cues include solute or signal molecule concentrations, the internal availability of storage molecules, or even stochastic processes. This allows for a better representation of dynamic microbial traits. A new biochemical submodel allows to express kinetic models through arithmetic functions. These arithmetic functions may include conditions in agent's local environment such as solute or signal molecule concentrations, and the internal conditions or properties of the agent. This massively expands the range of (bio-)chemical submodels available as any model that can be expressed arithmetically can be used. The new mechanical interaction and microbial morphology sub-models introduce rod-shaped and filamentous microbes to the modeling platform and,

like the (bio-)chemical models, allow the modeler to express these interactions through arithmetic functions.

By removing the necessity to provide abstract software or model parameters and enabling biologists to formulate their system in their own language rather than computer code, the gap between experimental and modeling work is reduced. This was achieved with iDynoMiCS 2.0 by the development of self-analyzing and optimizing algorithms, a set of default parameters where they make sense and relying on understandable parameters with biological meaning where they are needed. A graphical user interface further helps lower the barrier of entry into individual-based modeling. The integration of quantitative tools for spatial structure analysis provides insight into the structural development of microbial aggregates. By quantifying the spatial structure, it becomes easier to compare model results with experimental observations. Sensitivity analysis tools, utilizing Morris screening, reveal what state variables (properties of the modeled biofilm) are sensitive to the model inputs and thus what properties of the biofilm may change resulting from changing properties of individual microbes in the model. In this way it becomes easier to link emergent properties of the biofilm to microbial traits. A genetic algorithm was implemented to help estimate the model parameters. Despite the long evaluation times, stochastic processes, and localized variability, the genetic algorithm proves to be effective in optimizing an individual based biofilm model. The combination of tools facilitates Pattern-Oriented Modeling (POM) with iDynoMiCS 2.0. POM is a modeling approach that aims to identify and replicate the patterns that characterize the modeled system observed at different levels of the system's organization and scale. With the modeling platforms developed for this thesis, it is now possible to identify these characteristic multi-level patterns, and to infer individual based properties from them. POM can help mitigate the issues caused by the limited availability of individual based observations. In many cases reasonable initial parameter estimation can be made using existing empirical models. A focused effort addressing software bottlenecks enables biofilm simulations with over 10 million agents. This is over two orders of magnitudes more than iDynoMiCS 2.0's predecessor. A rigorous testing process verifies the effectiveness and numerical accuracy of the solvers and algorithms used in the platforms.

This thesis further includes multiple models utilizing iDynoMiCS 2.0 and eGUT, which provide interesting insights into the ecological processes and the dynamic community development of microbial aggregates. This includes a

model studying the interactions of yield- and rate strategists with either filamentous or spherical morphologies. The model reveals a strong competitive advantage for filament forming agents under nutrient limiting conditions. The next model explores spatial pattern formation during the initial aggregation process under various hydrological conditions and shows how differences in agent surface properties can lead to cell sorting. Agents with large Gibbs free energy of adhesion (ΔG_{adh}) form the center of an aggregate, surrounded by agents with small ΔG_{adh}. Further, a partial nitritation anammox model was developed. This model was used to identify how the growth characteristics of simulated microbes affect the structure and shape of the emergent biofilm. Finally, a multi-compartment gut model is developed using eGUT. The model simulates the microbial community composition and spatial structure of the proximal, transverse and distal colon. The insights into the spatial structure development of the mucosa are particularly valuable as it is often not possible to obtain this *in vivo*. Additional models are presented in the included manuscripts and/or available through the open source software repositories. The work presented in this thesis allows for a more accurate representation of microbes embedded in microbial aggregates in computational models. This improved representation enables these models to better capture the processes and interactions within biofilms and thereby predict patterns and phenomena that could not be captured with previous approaches.

Dansk sammenfatning

Mikrober findes almindeligvis i aggregatform, hvor en samling af forskellige mikrobielle arter fylder forskellige økologiske nicheområder. De spiller en essentiel rolle i globale næringscyklusser, menneskers og dyrs sundhed samt i talrige industrielle anvendelser. Mikrobielle aggregater, i mange tilfælde benævnt som biofilm, har en dynamisk livscyklus. Cyklussen begynder, når planktoniske celler begynder at binde sig til en overflade eller andre celler. I denne såkaldte reversible tilknytningsfase spiller mikrobielle overfladeegenskaber, som cellers hydrofobicitet, en nøglerolle. Derefter går den nydannede biofilm over i den irreversible tilknytningsfase, hvor flere celler fortsætter med at blive indlejret i det voksende aggregat, og fænotypiske ændringer, der er forbundet med mikrobiel liv i en biofilm, herunder produktion af store mængder ekstracellulære polymere stoffer (EPS), begynder at forekomme. Når aggregatet vokser, begynder rumlige strukturer som mikrokolonier og stratifikation at opstå. Modne biofilm kan nå en sidste fase, hvor nedbrydning af EPS forekommer, og celler spredes ud i miljøet. Den rumlige og samfundsmæssige struktur af et mikrobielt aggregat skyldes en kombination af fysiske processer som stoftransport og mekaniske belastninger, der virker på det mikrobielle aggregat, og de individuelle træk og adfærd hos mikroberne indlejret i aggregatet.

På grund af kompleksiteten af multi-arts mikrobielle aggregater og vores begrænsede evne til at observere den dynamiske udvikling af interne strukturer og biokemiske forhold i mikrobielle aggregater, har modelleringsmetoder spillet en væsentlig rolle i biofilmforskning. De første matematiske og beregningsmæssige biofilmmodeller, udviklet for over et halvt århundrede siden, afklarede, hvordan kemiske gradienter dannes inden for mikrobielle aggregater, der igen kan føre til stratifikation af mikrobielle arter i aggregatet. Senere viste Individual-based Models (IbM), hvordan individuelle træk og variation kan påvirke udviklingen af det mikrobielle aggregat, og hvordan subtile forskelle kan føre til helt forskellige resultater.

Mere end to årtier med IbM-assisteret biofilmforskning har vist sig frugtbar, og tilgangen bruges i stigende grad i moderne forskning. Et stigende antal modelleringsplatforme muliggører udviklingen af IbM'er. Imidlertid begrænser nuværende ulemper og begrænsninger os fra at bruge metoden i en bredere vifte af studier. Disse ulemper inkluderer den begrænsede skala af beregningsegnede modelsystemer, den begrænsede tilgængelighed af individuelle observationer, der kræves til parametrisering af modelsystemer,

den ekspertise, der kræves for at udvikle en IbM, og begrænsninger i evnen til præcist at repræsentere nogle vigtige mikrobielle træk eller modelsystemer med de i øjeblikket tilgængelige modelleringsværktøjer. Nylige bestræbelser fra flere forskningskonsortier har bragt betydelige forbedringer i nogle af disse begrænsninger, men en integreret tilgang, der adresserer alle disse begrænsninger samtidigt, mangler stadig.

Centralt i min afhandling er udviklingen af en sådan integreret tilgang og dermed lette en bred vifte af biofilmforskning for at forklare, hvordan individuelle mikrobielle træk fører til emergente egenskaber af det mikrobielle aggregat. Dette mål kan opdeles i følgende mål:

1. Muliggøre en større række dynamiske individbaserede karakteristika og adfærd for en bedre repræsentation af forskellige mikrobielle træk.
2. Udvide omfanget af (bio-)kemiske delmodeller for at give en bedre repræsentation af mikrobielle stofskifter.
3. Forbedre den fysiske repræsentation og de tilgængelige interaktionsmodeller.
4. Lukke kløften mellem eksperimentelt og modelarbejde ved at lette en bedre integration af eksperimenter og modellering og reducere den nødvendige ekspertise til at konstruere en model.
5. Muliggøre modelparametrisering, hvor individbaserede eksperimentelle observationer kan mangle.
6. Forbedre beregningseffektiviteten for at tillade modeller i større skalaer.

For at opnå dette mål har jeg arbejdet sammen med en dedikeret gruppe af samarbejdspartnere om udviklingen af den nye individbaserede modelleringsramme iDynoMiCS 2.0 og den nært beslægtede tarmmodelleringssoftware eGUT. Sidstnævnte udvider iDynoMiCS 2.0 med et epitel og slimhinde, der tillader modeller af tarmens lumen og slimhinde. iDynoMiCS 2.0 introducerer en ny struktur, hvor modellerede mikrober, ofte benævnt som agenter, kan samles fra ortogonale moduler. Enhver karakteristika, herunder morfologi, biokemisk og biologisk adfærd, kan dynamisk ændres som respons på eksterne eller interne signaler. Disse signaler inkluderer koncentrationer af opløsningsmidler eller signalmolekyler, den interne tilgængelighed af opbevaringsmolekyler eller endda stokastiske processer. Dette muliggør en bedre repræsentation af dynamiske mikrobielle træk. En ny biokemisk delmodel tillader at udtrykke kinetiske modeller gennem aritmetiske funktioner. Disse aritmetiske funktioner kan omfatte betingelser i agentens lokale miljø, såsom koncentrationer af opløsningsmidler eller signalmolekyler, og agentens interne forhold eller egenskaber. Dette

udvider kraftigt rækken af (bio-)kemiske delmodeller, da enhver model, der kan udtrykkes aritmetisk, kan bruges. De nye mekaniske interaktions- og mikrobielle morfologi-delmodeller introducerer stangformede og filamentøse mikrober til modelleringsplatformen og tillader, ligesom (bio-)kemiske modeller, at modelleren kan udtrykke disse interaktioner gennem aritmetiske funktioner.

Ved at fjerne nødvendigheden af at levere abstrakte software- eller modelparametre og muliggøre, at biologer formulerer deres system på deres eget sprog i stedet for computerkode, reduceres kløften mellem eksperimentelt og modelarbejde. Dette blev opnået med iDynoMiCS 2.0 ved udvikling af selvanalyserende og optimerende algoritmer, en række standardparametre, hvor de giver mening, og ved at stole på forståelige parametre med biologisk betydning, hvor de er nødvendige. En grafisk brugergrænseflade hjælper yderligere med at sænke barrieren for at kunne benytte den individbaseret modellering. Integrationen af kvantitative værktøjer til rumlig strukturanalyse giver indsigt i udviklingen af mikrobielle aggregaters struktur. Ved at kvantificere den rumlige struktur bliver det lettere at sammenligne modelresultater med eksperimentelle observationer. Værktøjer til følsomhedsanalyse, der anvender Morris-screening, afslører, hvilke tilstandsvariable (egenskaber ved den modellerede biofilm) der er følsomme over for modelindgange, og dermed hvilke egenskaber af biofilmen der kan ændre sig som følge af ændrede egenskaber af individuelle mikrober i modellen. På denne måde bliver det lettere at forbinde emergente egenskaber af biofilmen til mikrobielle træk. En genetisk algoritme blev implementeret for at hjælpe med at estimere modelparametre. På trods af de lange evalueringstider, stokastiske processer og lokaliseret variabilitet viser den genetiske algoritme sig effektiv i at optimere en individbaseret biofilmmodel. Kombinationen af værktøjer gør det muligt at modellere mønsterorienteret (POM) med iDynoMiCS 2.0. POM er en modelleringsmetode, der sigter mod at identificere og replikere de mønstre, der karakteriserer det modellerede system observeret på forskellige niveauer af systemets organisation og skala. Med de modelleringsplatforme, der er udviklet til denne afhandling, er det nu muligt at identificere disse karakteristiske flerlagsmønstre og udlede individuelle baserede egenskaber fra dem. POM kan hjælpe med at afbøde problemerne forårsaget af den begrænsede tilgængelighed af individuelle baserede observationer. I mange tilfælde kan der foretages rimelige initiale estimater af parametre ved hjælp af eksisterende empiriske modeller. En fokuseret indsats for at tackle softwareflaskehalse muliggør

biofilmsimulationer med over 10 millioner agenter. Dette er over to størrelsesordener mere end iDynoMiCS 2.0's forgænger. En omhyggelig testproces bekræfter effektiviteten og numeriske nøjagtighed af de løsere og algoritmer, der bruges i platformene.

Denne afhandling inkluderer yderligere flere modeller, der anvender iDynoMiCS 2.0 og eGUT, hvilket giver interessante indblik i økologiske processer og den dynamiske fællesskabsudvikling af mikrobielle aggregater. Dette inkluderer en model, der studerer interaktionerne mellem udbytte- og ratestrateger med enten filamentøse eller sfæriske morfologier. Modellen afslører en stærk konkurrencefordel for agenter, der danner filamenter under betingelser med næringsbegrænsning. Den næste model udforsker rumlig mønsterdannelse under den indledende aggregationsproces under forskellige hydrologiske forhold og viser, hvordan forskelle i agenternes overfladeegenskaber kan føre til celle-sortering. Agenter med stor fri energi af adhæsion (ΔG_{adh}) danner centrum af et aggregat, omgivet af agenter med lille ΔG_{adh}. Derudover blev der udviklet en model for partial nitritation anammox. Denne model blev brugt til at identificere, hvordan vækstegenskaberne for simulerede mikrober påvirker strukturen og formen af den fremkomne biofilm. Endelig udvikles en multi-kammer tarmmodel ved hjælp af eGUT. Modellen simulerer sammensætningen af mikrobiel samfund og rumlig struktur i den proksimale, transversale og distale tyktarm. Indsigter i udviklingen af slimhindens rumlige struktur er særlig værdifulde, da det ofte ikke er muligt at opnå dette *in vivo*. Yderligere modeller præsenteres i de medfølgende manuskripter og/eller er tilgængelige gennem open source software repositorier. Arbejdet præsenteret i denne afhandling tillader en mere nøjagtig repræsentation af mikrober indlejret i mikrobielle aggregater i beregningsmodeller. Denne forbedrede repræsentation muliggør, at disse modeller bedre kan fange processerne og interaktionerne inden for biofilm og dermed forudsige mønstre og fænomener, der ikke kunne fanges med tidligere tilgange.

Table of contents

1 Introduction and objectives

Microbial communities are omnipresent and play important roles in human and planetary health. Antimicrobial resistant biofilms can hinder effective treatment of microbial infections. The human microbiome found on the skin, oral cavity, respiratory- and gastronomic tract plays a vital role in human health and wellbeing. Environmental microbial communities play an essential role in global elemental cycles and in the breakdown of environmental toxins and pollutants (Costerton et al., 1987; Davey and O'Toole, 2000; Wanner et al., 2006). Microbial communities are widely used for various processes including water and residuals treatment, feed and food preservation and processing and plant growth promotion. These microbial communities are often found to form self-organized assemblages such as surface adhering biofilms (Flemming and Wuertz, 2019) (Costerton et al., 1995), loosely associated flocs or more tightly bound granular biofilms. Microbes have evolved and diversified over a long evolutionary history and harbor fascinating characteristics and behaviors. Within a single environment microbial diversity can be astonishing. For example in activated sludge flocs, thousands of unique species can be identified (Gonzalez-Gil et al., 2015; Peces et al., 2022). The complexity coming from the high diversity within microbial communities, which in turn are subject to spatio-temporal dynamics, can make it difficult to predict all the internal processes and the development of these communities (Widder et al., 2016).

Microbial life in aggregated form is vastly different from a planktonic lifestyle. Possibly one of the most obvious differences is the ability of aggregated cells to form distinct spatial structures. Multiple processes drive this spatial structure. Microbes can metabolize a wide variety of chemical compounds. In microbial aggregates, concentration gradients form resulting from the constant consumption and diffusion of these compounds in and out of the aggregate. This in turn can result in stratification often found in these microbial assemblages. Microbial growth and limited mobility within the aggregate can result in micro colonies of clonal populations (Davey and O'Toole, 2000). In partial nitritation anammox granules (figure 1.1), used in nitrogen removal systems utilizing microbial communities of ammonia oxidizing bacteria (AOB), anammox bacteria (AMX), and the typically unwanted nitrite oxidizing bacteria (NOB), both stratification and the formation of small clusters of bacteria can be observed. Microbes embedded in an aggregate show different characteristics, behavior and gene expression compared to planktonic cells (Parsek and

Fuqua, 2004). Microbes have been found to increase the production of extracellular polysaccharides (EPS) in biofilms, an important substance commonly associated with biofilms. Also genes involved in metabolism have been found to be upregulated in microbial aggregates (Becker et al., 2001). Microbes have developed intricate communication mechanisms such as quorum sensing, a mechanism involved in the self-regulation microbial communities, aiding their survival in certain environments (Lupp et al., 2003). Biofilms are often associated with increased anti-microbial resistance. Multiple mechanisms and properties of the biofilm aid this resistance (Stewart and William Costerton, 2001). At least part of this increased resistance can be directly linked to phenotype changes of the aggregated microbes. Phenotypic heterogeneity is even observed among individuals in single species microbial populations (Ackermann, 2013; Zimmermann et al., 2015). This makes life in microbial aggregates a complex symphony of dynamic processes and mechanisms occurring simultaneously.

Observations from classic laboratory experiments utilizing suspended cultures may not be applicable to the behavior of the same microbes in biofilms, since microbes within aggregates experience different selective pressures and environmental conditions from microbes with a planktonic lifestyle, and because they express different phenotypes and employ different mechanisms within an aggregate. Researchers have developed new approaches to overcome this issue (Costerton et al., 1987; Parsek and Fuqua, 2004). Specialized microscopy methods such as Fluorescence In Situ Hybridization (FISH) coupled to Confocal Laser Scanning Microscopy (CSLM) have provided insight in the spatial structure and species distribution in microbial aggregates, but they also come with important downsides. Detailed analysis of the internal structure of microbial aggregates are typically disruptive, making it impossible to follow the dynamic development of the spatial structures in these aggregates.

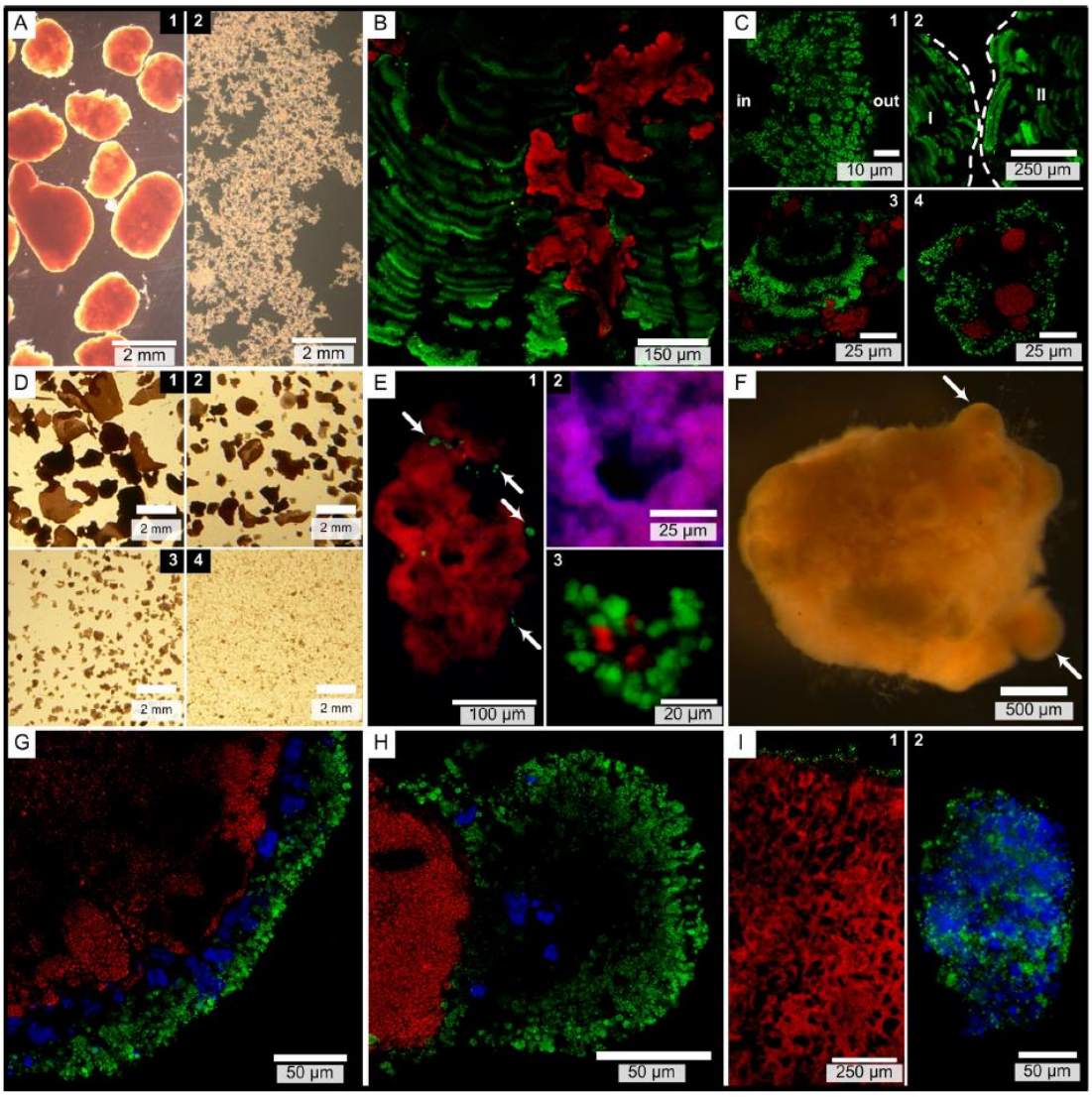


Figure 1.1. Micrographs of microbial aggregates, reprint from Vlaeminck et al. (2010)[1]. FISH with CLSM (B,C,G,H,I1) or epifluorescence microscope (E,I2) were used to determine the spatial organization of the aggregate, with FISH probes for AOB (green), NOB (blue) and AMX (red). Due to metabolic transformations of resources diffusing into the granules and the co-localization of sibling cells leading to clonal populations, the aggregates become spatially structured.

[1] This figure was published in Applied and environmental microbiology, 76, Vlaeminck, SE, Terada A, Smets BF, De Clippeleir H, Schaubroeck T, Bolca S, Demeestere L, Mast J, Boon N, Carballa M, Verstraete W, Aggregate size and architecture determine microbial activity balance for one-stage partial nitritation and anammox, 900-909, Copyright American Society for Microbiology (2010).

Mathematical and computational models have helped to gain a better understanding of these dynamic processes, the nature of distinct spatial structures and the ecology of microbial aggregates. The first mathematical biofilm models were developed in the 1970s (Wanner et al., 2006). Initially, these models considered the biofilm as a one dimensional continuum. Although simple, these initial models help explain the formation of chemical gradients and stratification of species in microbial aggregates (Kissel et al., 1984; Rittmann and Manem, 1992; Wanner and Gujer, 1986). Later two- and three-dimensional models were used to help understand the formation of more complex spatial structures. This includes the formation of finger-like structures in biofilms, resulting from a positive feedback loop where cells in an extrusion from the biofilm have better access to nutrients and thus can grow faster enlarging the extrusion (Dockery and Klapper, 2001; Picioreanu et al., 1998a; van Loosdrecht et al., 2002).

Cellular Automata (CA) and Individual-based Models (IbM) of biofilms first started to appear in the 1990s (Wimpenny and Colasanti, 1997; Picioreanu et al., 1998b; Kreft et al., 1998, 2001). These models discretized biomass into separate entities, restricted to a spatial-lattice in case of CA or unrestricted in case of IbMs, rather than modeling biomass as a concentration in a continuous biofilm. These models follow a bottom-up approach where the biomass entities, representing either small groups of cells or single cells, adhere to a certain set of rules or behaviors. These entities are sometimes referred to as biomass units, agents or individuals. With these models the properties of the microbial aggregate including spatial patterns, community composition, productivity and resilience are emergent properties resulting from the collective of properties and behaviors of the agents and interactions with their environment. Early CA models were used to simulate model micro colony formation (Schindler and Rataj, 1992) and different biofilm structures (Wimpenny and Colasanti, 1997). At present IbMs have become the more common approach among the two, as CA models come with a number of drawbacks concerning biomass spreading including over dispersion of microbial species and model outcomes that are sensitive to coordinate system (Eberl et al., 2001; Picioreanu et al., 2004). Individual based models are particularly good at capturing the effects of local interactions, individuality and adaptive behavior on the spatio-temporal dynamics of microbial aggregates. This includes stochastic events such as dispersal and community assembly, mutations, up- and down-regulation of genes or horizontal gene transfer. For example, the Individual based *Anabaena* spp.

model by Hellweger et al. (2016) incorporated gene expression and differentiation of *Anabaena* spp. individuals within a filament, and reproduced nearly all of the patterns observed *in vitro*.

IbMs ability of capturing fine grained spatial pattern formation, such as the co-localization of sibling cells forming micro colonies, allowed Kreft (2004) to demonstrate how this phenomenon can explain altruistic behavior in biofilms. This type of microstructure is also essential in understanding and more accurately predicting the behavior of engineered systems. Picioreanu et al. (2016) showed how the formation of micro colonies in nitrifying microbial aggregates affects local oxygen availability, the perceived microbial affinity for oxygen and consequently the growth rate of microbes embedded in the aggregate. Hence, the spatial structure of the microbial aggregates plays an important role in the microbial community composition and can play a key role in optimizing engineered systems. An example of this is the suppression of NOB in partial nitritation anammox nitrogen removal systems, where NOB is unwanted due to its competition for oxygen with AOB and for nitrite with AMX (Picioreanu et al., 2016).

The first generation of IbMs used a simple shoving algorithm to resolve overlap between agents. Later more detailed mechanical interaction models were developed. Janulevicius et al. (2010) used an individual-based model to show how striking circular patterns in myxobacteria aggregates can be explained by the swimming behavior of these bacteria. Celler et al. (2014) used an individual-based *Diatoma* model to explain the intricate structures of interconnected filaments observed in naturally occurring *Diatoma* biofilms. Others coupled individual-based modeling with computational fluid dynamics (CFD) to simulate biofilm detachment and deformation processes (Xavier et al., 2005; Martin et al., 2013; Jayathilake et al., 2017).

Individual-based modeling can provide useful abstractions from a more complex biological reality. Schluter et al. (2015) coupled experiments with EPS producing *V. cholera* in combination with a mutant variant lacking this ability, with an IbM. This was done to demonstrate how the adhesive properties of EPS affect microbial competition within the biofilm. It may not always be possible to obtain a mutant with the desired trait variation when studying other properties or behaviors of microbes, and if the mutant is obtained it may change in multiple aspects. This makes it difficult to distinguish what the primary causes of changes at the community level are. In such cases individual-based modeling can provide insight on questions in fundamental microbial ecology, as it allows us to study the effects of a single property in isolation. Setting up an

IbM, using one of the IbM platforms that have come available, is typically also far less time consuming than obtaining mutant strains, and thus IbMs can be an effective tool to inform later experimental work. Once a good computational model is established, it becomes trivial to run thousands of combinations of *in situ* experiments for a given model microbial aggregate. The position, environmental conditions and lineage of any given simulated microbe can be tracked over time, providing insight in dynamic developments within the simulated microbial aggregate.

The development of an IbM platform requires expert knowledge and is a time consuming effort, especially if it is developed to facilitate a wide variety of processes and answer a wide variety of research questions. A collaborative effort of multiple leading research groups in the field resulted in inception of the open source modeling platform (individual-based Dynamics of Microbial Communities Simulator) iDynoMiCS (Lardon et al., 2011). The platform combines features of previous models and facilitates the use of individual-based modeling of microbial communities for scientists with less programming experience. To date the platform is used in numerous scientific studies and is one of the most cited (243)[2] general purpose individual-based microbial community modeling platforms, only following its predecessor BacSim (508)[2]. The platform has received multiple extensions, such as the ability to simulate horizontal gene transfer (Seoane et al., 2010), a process that plays a central role in the spread of anti-microbial resistance genes, and has influenced the design of other modeling platforms (Bogdanowski et al., 2022; Breitwieser et al., 2022; Goñi-Moreno and Amos, 2015; Gorochowski et al., 2012; Jang et al., 2012; Li et al., 2019).

Now, over a decade after the introduction of iDynoMiCS, there is a wide variety of individual-based modeling platforms that focus on microbes and can support a range of specific models. Platforms similar to iDynoMiCS include Biocellion (Kang et al., 2014), Simbiotics (Naylor et al., 2017), BacArena (Bauer et al., 2017), NUFEB (Li et al., 2019), ACBM (Karimian and Motamedian, 2020) and McComedy (Bogdanowski et al., 2022). These platforms focus on microbial growth, metabolism, and mass transport and assess

[2] Number of citations according to google scholar.

how different growth strategies or metabolic interactions affect the fitness of species growing in a biofilm or how they impact systems-level outcomes in wastewater treatment systems or bioreactors. iDynoMiCS, NUFEB and Simbiotics can all model growth using equations originating from enzyme kinetics that determine reaction rates from substrate concentrations, such as Monod kinetics. NUFEB can additionally model agent growth based on thermodynamics, calculating the Gibb's free energy of catabolism (Gogulancea et al., 2019). Reaction rates and diffusion are coupled and solved using partial differential equation (PDE) solvers. These solvers are made efficient by taking advantage of a separation of timescales. The growth of microbes is on a much slower timescale than diffusion, and thus a pseudo steady state is reached rapidly after any changes occur within the system. BacArena and ACBM are unique in utilizing flux-balance analysis to estimate the metabolic flux based on their local solute concentrations. However, this is a computationally demanding process, making the platforms more restrictive in terms of model scale. BacArena only models agents in a fixed 2D grid, with one agent per grid cell, like cellular automata, while ACBM groups agents together when evaluating internal processes. The other biofilm modeling platforms simulate grid-free agents that evaluate internal processes on an individual basis. NUFEB and Simbiotics also allow adhesive forces to be modeled. In NUFEB and iDynoMiCS, agents are spherical, while in Simbiotics and ACBM they can be spherical or rod-shaped.

There is an additional group of platforms that originates from systems and synthetic biology and seek to discover how specific microbial community behaviors or phenomena can be achieved through the creation of synthetic microbial communities (Gorochowski et al., 2020): CellModeller (Rudge et al., 2012), BSim 2.0 (Matyjaszkiewicz et al., 2017) and gro (Gutiérrez et al., 2017). They can simulate microbial communities made up of rod-shaped microbial agents with specific metabolic, sensing and signaling properties. All three can simulate gene regulatory networks and diffusion of signaling molecules to explore and/or design synthetic microbial communities. While gro can only simulate 2D systems, CellModeller can simulate both 2D and 3D systems, BSim 2.0 can only simulate 3D systems. In models that use these platforms, growth kinetics are typically less important than gene regulation, hence growth is modeled as a simple rate, as in CellModeller, or a rate based on cell length, as in BSim 2.0. Gro allows growth to be based on Monod kinetics. CellModeller and gro do not include environmental constraints such as physical boundaries, thus agents tend to grow outwards to form round colonies. BSim 2.0, can model physical

spaces such as microfluidic chemostats where cells may grow and release diffusing signaling molecules.

Mathematical and computational modeling has delivered valuable insights into the dynamic processes occurring within microbial aggregates. Increasingly detailed models have improved our understanding of the ecological processes shaping these microbial assemblages and have informed the design of essential technologies and processes utilizing multispecies biofilms, in particular water treatment systems. The advent of accessible open source modeling tools enabled a broader group of researchers to utilize the method in their work. Although useful, these modeling tools still come with several limitations and drawbacks. This includes the limited scale of computationally feasible model systems, the limited availability of individual based observations required to parameterize the model systems, the expertise required to develop an IbM, and limitations in the ability to accurately represent some important microbial traits or model systems. Recent efforts by multiple research consortia have brought substantial improvements in some of these limitations, but an integrated approach that addresses all these limitations simultaneously is still missing. Ideally an individual based modeling platform should be able to capture all biologically significant processes and properties and thereby aid our understanding of microbial aggregates and ultimately improve our ability to construct predictive biofilm models.

1.1 Objectives and thesis structure

The goal of this thesis is to develop an integrated IbM approach that addresses the key limitations found in prior work. An approach that facilitates a wide range of biofilm research and helps explain how individual microbial traits lead to emergent properties of the microbial aggregate. This goal can be subdivided into the following objectives:

1. Enabling a larger range of dynamic individual-based characteristics and behaviors for a better representation of various microbial traits
2. Widening the range of (bio-)chemical sub-models to provide a better representation of microbial metabolisms.
3. Improving the physical representation and interaction models available.
4. Closing the gap between experimental and modeling work facilitating a better integration of experiments and modeling and reducing the required expertise to construct a model.
5. Facilitating model parameterization where individual based experimental observations may be lacking.

6 Improving computational efficiency to allow for models at larger scales.

Dynamic individual-based characteristics and behaviors

A key insight is that not only biofilms but also microbes themselves are dynamic in nature. The characteristics of microbes can change in many ways due to a variety of different causes. This can involve their metabolism, production of enzymes or EPS, their morphology, or other aspects such as motility. While some prior frameworks, such as iDynoMiCS 1, already allow for some dynamic traits, for example through metabolic switching, other changes in agent characteristics are typically not possible. By enabling higher degree of flexibility in agent characteristic a larger variety of dynamic phenomena could be captured through an IbM, including sporulation, dormancy, signaling (quorum sensing), genetic mutation, genetic switching, adaptation, microbial life cycles, ageing, chemotaxis and more.

(Bio-)chemical conversion and microbial metabolism

Although the Monod equation has been a staple in biofilm modeling, the kinetic model substantially simplifies the metabolism of microbes. Most IbM platforms only allow for kinetic models following a predefined structure such as Monod kinetics. Yet in some situations other kinetic models may be more suitable or even required, such as a modified Monod model adjusting for light-limitation in microalgae (Lee et al., 2015). Others have used thermodynamic based kinetic models to gain new insights (Gogulancea et al., 2019). By allowing the formulation of kinetic expression through an arithmetic function, it should become possible to use any type of kinetic model, and thus it should become possible to model organisms with metabolisms that cannot be accurately defined through Monod kinetics.

Microbial morphology and physical interaction

There is a lot of variety in microbial shape (Angert, 2005) and size (Schulz and Jørgensen, 2001). It has been argued that these properties are as important as other traits in ecological success (Kevin D. Young, 2006). Further, it is known that microbial aggregation involves drastic changes in microbial surface properties, allowing them to adhere to substrates or other microbes. First generation IbM platforms model microbes as spherical particles that avoid overlap through a shoving algorithm, but do not model microbial adherence. By integrating and further developing on the mechanical agent model by Janulevicius et al. (2010), Celler et al. (2014) and (Storck et al., 2014) a better representation of how

mechanical interactions and microbial morphology affect microbial interactions and structural development of microbial aggregates should become possible.

Closing the gap between experimental and modeling work

IbMs typically depend on many software and model parameters. This includes abstract concepts such as a shoving factor, which has no parallel in biology. Instead, IbM tools that enable biologists to formulate their system in their own language rather than computer code could bridge the gap between experimentalists and modelers (Hellweger et al., 2016). By implementing solvers that self-inspect convergence and automatically adjust in order produce outcomes within a reasonable error range, and by relying on parameters that can be obtained from experimental observations or literature, individual based modeling should become feasible for a broader range of researchers.

Model parameterization with limited experimental observations

The implementation of pattern analysis, sensitivity analysis and parameterization tools should make it possible to directly compare experimental and modeling results and enable model calibration on experimental data. Multi-scale pattern analysis facilitates Pattern-Oriented Modeling (POM). POM is a modeling approach that aims to identify and replicate the patterns that characterize the modeled system observed at different levels of the system's organization and scale. Identifying and quantifying these characteristic multi-level system patterns allows the inference individual based properties from them. This helps mitigate the issues caused by the limited availability of individual based observations. In many cases reasonable initial parameter estimation can be made using existing empiric models.

Computational efficiency

The spatial detail and complexion of IbMs limit the scale of the computational domain at which a simulation is feasible. Some spatial pattern formation may not be captured in a limited computational domain, it may further hinder modeling biofilms with substantial height, and it may force researchers to limit their model to two spatial dimensions, which in turn comes with a variety of side effects. The computational intense workload also makes large scale parameter sweeps with IbMs time consuming, which is probably one of the main reasons why many studies utilizing IbMs do not include detailed sensitivity

analysis and model calibration. By addressing software bottlenecks and inefficiencies, simulations at a much larger scale or model evaluation in a shorter duration of time should become feasible.

This thesis is organized in 7 chapters starting with this introduction reviewing the scientific context and prior biofilm modeling research. This thesis further includes 3 manuscripts which provide additional detail on the covered topics.

Chapter 2. *Individual-based biofilm modeling in microbial ecology and biotechnology.* Provides an overview of the inner workings of IbMs and highlights key differences between iDynoMiCS 2.0 from other platforms.

Chapter 3. *Verification of solvers and algorithms.* Provides a short overview of the verification process of iDynoMiCS 2.0. The platform was evaluated with known analytical solutions as well as compared to prior work.

Chapter 4. *The estimation of microbial growth characteristics.* Reviews thermodynamics based approaches and insights that can aid estimating initial parameters for biofilm models that include under characterized microbes.

Chapter 5. *Model analysis and parameterization.* Provides an overview of the analysis and parameterization strategy developed for IbMs. A basic partial nitritation/anammox model was used to provide context to the concept.

Chapter 6. *Pattern formation and population development in microbial aggregates.* Provides an overview of initial pattern formation and microbial ecology case studies using iDynoMiCS 2.0.

Chapter 7. Conclusions and perspectives

Manuscript I. *Is it selfish to be filamentous in biofilms? Individual-based modeling links microbial growth strategies with morphology using the new and modular iDynoMiCS 2.0.* Introduces the individual-based modeling platform iDynoMiCS 2.0, which addresses many key limitations.

Manuscript II. *Force mediated spatial structure in granular biofilm assembly.* Utilizes iDynoMiCS 2.0 to investigate the ecology and processes involved in microbial aggregates, including the role of differential adhesion on pattern formation in incipient aggregates.

Manuscript III. *eGUT, a platform enabling agent-based modelling of the gut microbiome and its spatial interactions with the host mucosa*. Combines the capabilities of iDynoMiCS 2.0 with a gut lumen and mucosa model to enable gut biofilm modeling.

2 Individual-based modeling of microbial aggregates

2.1 Components of biofilm models

Although microbial aggregates are found in many shapes and forms in a wide variety of environments, the chemical and physical processes and properties impacting the biofilm are of the same nature, and thus many biofilm models share a similarity in what properties and processes they describe. Biofilm models, including IbMs, describe the biofilm and its environment as a collection of numbers and mathematical equations.

Space is described through a Cartesian coordinate system and is typically limited by boundaries. This limited space is also known as the computational domain. IbMs describe biomass as individual agents which have properties and conduct programmed behavior. The complexity of these agents can be quite different from one model implementation to the other, but typical properties are position, mass, volumetric mass density, shape and chemical composition. Typical behavior includes chemical conversion, growth and cell division. These agents are located in an environment which also comprises a number of properties and processes. Most model implementations include local concentrations of chemical species, also referred to as solutes, and diffusivity, while some models may include other properties such as pH, temperature and even light intensity. Processes that are usually considered in IbMs are mass transport and shearing.

2.2 (Bio-)chemical conversion and transport processes

Advection and diffusion are the processes that drive matter transport in and around the biofilm. Some advective transport may occur in the biofilm, as observed in channel forming biofilms (Wimpenny and Colasanti, 1997), yet, generally the majority of matter transport in the biofilm is considered to be diffusion driven. Matter transport further away from the biofilm is typically advection driven. The intermediate zone transitions from diffusion dominated to advection dominated.

Some models have utilized computational fluid dynamics to model matter transport and mechanical stress over the biofilm. These simulations are com-

putationally intense however and are not feasible over the timescales considered in biofilm development. Therefore, the model system is typically simplified and modeled as separate regions (figure 2.1). This includes a region that is perfectly well-mixed called the bulk, which is the furthest away from the biofilm, and a region where mass transport is considered fully diffusion driven, this includes the biofilm and a so called diffusion boundary layer which separates the biofilm from the bulk. Models of surface adhering biofilm may additionally include a support region which represents the substratum the biofilm is attached to.

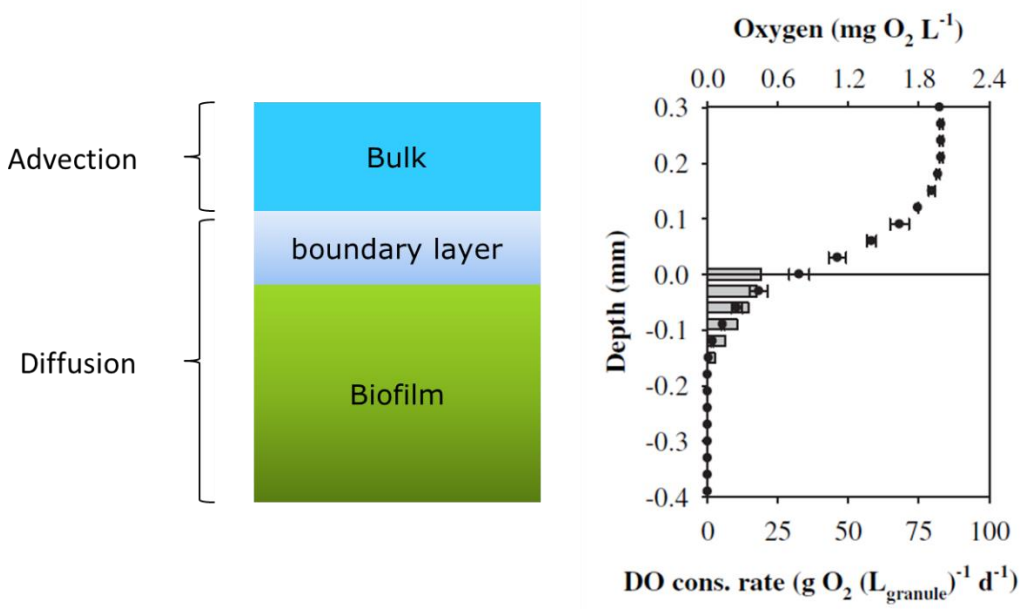


Figure 2.1. Schematic representations of the different regions in a typical biofilm model (Left), chemical species are assumed to be completely mixed within the bulk region where transport due to advection is dominant. Experimental observations of oxygen concentrations and consumption rates in a partial nitritation anammox granule (Right), reprint from Vázquez-Padín et al. (2010)[3], reveal typical concentration profiles for a rate limiting substrate.

[3] This figure was published in Water research, 44, Vázquez-Padín J, Mosquera-Corral A, Campos JL, Méndez R, Revsbech NP, Microbial community distribution and activity dynamics of granular biomass in a CANON reactor, 4359-4370, Copyright Elsevier (2010)

The bulk region is free of any chemical gradients. Temporal dynamics may still affect this compartment and it is therefore chemical concentrations in this compartment are often modeled using a classical bioreactor model:

$$\frac{d}{dt}S_s(t) = \sum_{inflows} D_i S_{s,i}(t) - \sum_{outflows} D_i S_s(t) + q_s(t) \quad 2.1$$

where $S_s(t)$ is the concentration of solute s at time t and D_i the dilution rate for a given inflow/outflow and the rate expression $q_s(t)$, which is the net sum of all local consumption and production rates for solute S. Alternatively, the concentrations in this region may be kept constant.

Due to the consumption of chemical species by microbes in the biofilm and diffusive fluxes occurring in and around the biofilm, chemical gradients form in the biofilm and in its proximity. The majority of IbMs model the concentration gradient in the proximity of the biofilm as a diffusion boundary layer (DBL). The DBL is a hypothetical layer assuming matter transport is fully driven by diffusion, which allows estimation of the chemical concentrations at the biofilm surface without simulating advection explicitly. The thickness of this layer can be determined experimentally by extrapolating the concentration gradient at the surface of the biofilm to the bulk concentration.

Alternatively, the Sherwood number (Sh) can be used to estimate the thickness of the diffusion boundary layer (Wanner et al., 2006):

$$Sh = \frac{k_m d_p}{\omega_s} \quad 2.2$$

Where k_m is the mass-transfer coefficient, d_p is the characteristic length and ω_s the diffusion coefficient of the liquid (usually water). The layer thickness (L_{DBL}) is defined as follows (Rittmann and McCarty, 2018):

$$L_{DBL} = \omega_s / k_m \quad 2.3$$

Empiric correlations using the Reynolds number (Re) and Schmidt number (Sc) have been established to estimate the Sherwood number, and thus also k_m, under various hydrodynamic conditions. For fluid flow over a flat plate the following correlation exists (Ruocco, 2018):

$$Sh = 0.332\, Sc^{1/3}\, Re^{1/2} \text{for} \quad Sc \gtrsim 0.5 \quad 2.4$$

For fluid flow around a rigid spherical particle the following equation can be used (Wanner et al., 2006):

$$Sh = 2 + 0.6\, Sc^{1/3}\, Re^{1/2} \quad 2.5$$

Picioreanu et al (2000a) developed a correlation that adjusts for the biofilm surface area, while other geometries or conditions may require different correlations. Many of these correlations are available in scientific literature (Horn and Lackner, 2014).

Microbial growth happens on a far larger time-scale than chemical reactions and diffusion (Picioreanu et al., 2000b). This allows biofilm models (including IbMs) to make a pseudo steady state assumption. This is to say that over a small enough time step the overall change in biomass is small enough to consider it constant over a short period of time. The model can approximate the chemical concentrations during that period by calculating steady state conditions, where Fickian diffusion and (bio-)chemical conversion balances. This simplification, known as the separation of timescales, can drastically reduce the computational demand of the simulation. The rate of change for each solute is now given by the elliptic Partial Differential Equation (PDE) that combines Fickian diffusion, given by the divergence div of the diffusional flux driven by the concentration gradient ∇S_s, and local reactions (Dyke, 1998).

$$\frac{\partial}{\partial t} S_s(x,t) = \mathrm{div}\big(\omega_s(x) \cdot \nabla S_s(x,t)\big) + q_s(x,t) \qquad 2.6$$

where $\boldsymbol{x}$ is the spatial position, ω_s is the local diffusivity, *and* $S_s(x,t)$ the local concentration of solute *s*. The gradient of a field, ∇, is commonly defined as:

$$\nabla = \hat{\imath}\frac{\partial}{\partial x} + \hat{\jmath}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z} \qquad 2.7$$

At the outer bounds of the computational domain any one of the following boundary conditions may be of effect; (1) An (impermeable) Neumann boundary with the following condition:

$$\nabla \mathrm{S}(x) \cdot \hat{n}|_{\Gamma} = 0 \qquad 2.8$$

The domain's solute concentration gradient $\nabla \mathrm{S}(x)$ normal (normal vector: $\hat{n}$) to the boundary Γ is 0. Neumann boundaries are used to represent hard surfaces. (2) A Dirichlet boundary where solute concentrations have a fixed value S_b:

$$\mathrm{S}(x)|_{\Gamma} = \mathrm{S}_b \qquad 2.9$$

Dirichlet boundaries are used to connect the bulk environment to the rest of the biofilm model. (3) A Periodic boundary that matches the concentration and gradient on opposite side of the computational domain:

$$S(x)|_{\Gamma min} = S(x)|_{\Gamma max} \quad 2.10$$

$$\nabla S(x)|_{\Gamma min} = \nabla S(x)|_{\Gamma max} \quad 2.11$$

Periodic boundaries allow the modeler to represent an overall larger biofilm by simulating a smaller representative segment.

The diffusion coefficient in the biofilm is typically different from that in the liquid the biofilm diffusivity ($\omega_{s,f}$) over liquid diffusivity ratio can range from 0.5 to 0.8, for small molecules $\omega_{s,f} = 0.8\omega_s$ is typically an appropriate approximation (Rittmann and McCarty, 2018).

2.3 Physical representation of microbes (agents)

Microbes in IbMs are represented at the level of single cells, rather than a biomass continuum as in classic biofilm models. The simulated cells are referred to as agents. In early IbMs these agents are represented as small circles or spheres that resolve overlap using a shoving algorithm (figure 2.2). These algorithms do not model intricate interaction forces (repulsion, attraction, agent attachment), but are still effective in decompressing a growing biofilm.

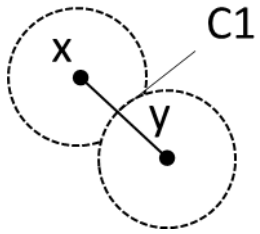


Figure 2.2. Shoving algorithms calculate a movement vector that is the inverse of the sum of overlap vectors from neighboring cells. Small differences in how these vectors are calculated exist. In BacSim this overlap vector is defined as follows: $R_0 = kR_x + R_y - C1$ where R_x and R_y are the radius of the focal cell and its neighbor, C1 is the Euclidean distance between the centers and k is a radius multiplier controlling the packing density of the biofilm.

Later models introduce new shapes and physical interactions agents may have (figure 2.3). These mass-spring models were first implemented by Janulevicius *et al.* (2010), Celler *et al.* (2014) and Storck *et al.* (2014). iDynoMiCS 2.0 builds further on this approach but is no longer limited by spring forces and therefore dubs it Force-based Mechanics (FbM).

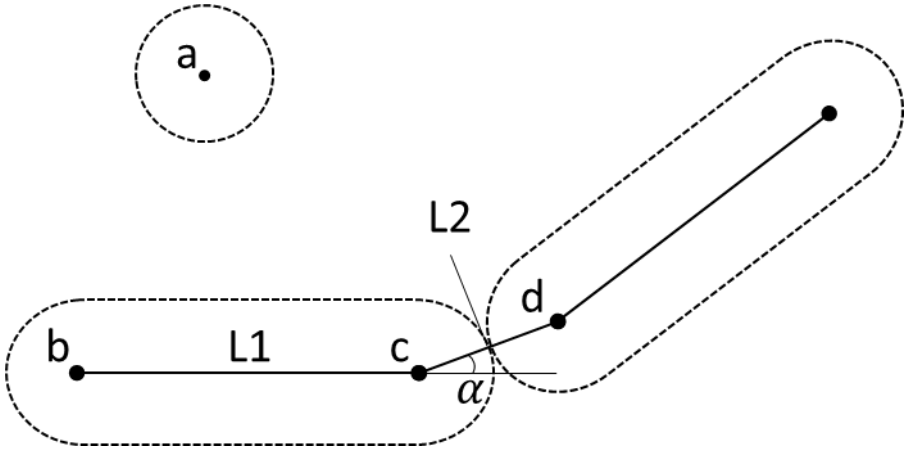


Figure 2.3 IbMs with advanced physical representation and mechanical interaction allow the modeling of agents with other morphologies such as rod-shaped bacilli or filamentous organisms. Multiple IbMs use a mass-spring system or similar approach to model these shapes. Dashed lines indicate sphere-swept volumes of 'dots' or line-segments. Dots are mass-points indicating position and orientation of agents. Solid lines indicate mechanical interactions between points (forces between points modeled as springs): The spring from b to c is responsible for the rigidity of rod-shaped agent, c and d connects two rod-shaped agents. α is the angle between two elements of a filament. This angle can be counteracted by a torsion spring applying forces on b, c and d. L1 and L2 are the moment arms.

Forces acting upon the agent may arise from multiple sources. This includes interaction forces such as collisions, structural forces such as the forces connecting the segments of a filament, motor forces, gravity buoyancy and (viscous) drag force.

2.3.1 Drag force and inertia

Drag force works in opposite direction of the movement of the cell relative to its environment. Inertia is the tendency of a moving object to maintain its speed and trajectory. Although this may be counter intuitive, inertia can usually be considered negligible on microbial scale (Purcell, 1977; Berg, 1993). In other words, the time it takes for a microbe to change its velocity due to a change in forces acting upon it is expected to be so small that it can be ignored. This is a logical conclusion when looking at the dimensionless Reynolds number (Re), which is interpreted as the ratio between inertial and viscous forces:

$$Re = \frac{F_{inertial}}{F_{viscous}} \qquad 2.12$$

For a spherical particle such as an agent, the Reynolds number can be calculated as follows:

$$Re_a = \frac{\rho_f \mathrm{v} d_a}{\mu_f} \quad 2.13$$

Where ρ_f and μ_f are the density and the dynamic viscosity of the fluid. v is the mean velocity of the object relative to the fluid, d_a the diameter of the agent. For typical motile microbes the particle Reynolds number will be around 10^{-4}, for non-motile cells this will be less. Under these conditions ($Re_a \leq 0.1$) Stokes' law can be used to calculate the drag force on the particle (Rhodes, 2008).

$$Re_a \leq 0.1 \rightarrow F_D = 3\pi d_a \mathrm{v} \mu_f \quad 2.14$$

A very small Reynolds number tells us that the inertial component is very small in comparison to the viscous component. Even in more extreme microbial scenarios this holds. We can use the drag force in combination with newton's second law of motion to calculate the deceleration of the agent:

$$\frac{dv}{dt} = \frac{-3\pi d_a \mathrm{v} \mu_f}{m} \quad 2.15$$

The derivative of this equation allows us to obtain the time it takes the agent to lose all of its velocity after any driving force is taken away and the amount of distance it travels during this time (See Box 2.1).

Box 2.1: The coasting time and distance after the fastest swimming microorganisms stops propelling itself.

The record holder of fastest swimming microorganism "*Candidatus* Ovobacter propellens". This prokaryote is able to swim at velocities up to 1000 $\mu m/s$ and has diameter of up to 5 µm (Fenchel and Thar, 2004). Even in this case we still find a Reynolds number of only $5 \cdot 10^{-3}$ during its fastest swim. Using Newton's second law coasting time and distance can be calculated. At a steady initial swimming speed of 1000 $\mu m/s$ in water, a cell density of 1 g/cm^3 we find that it takes only $1.39 \cdot 10^{-6}$ s and 13.9 Å for *Candidatus* Ovobacter propellens to lose all of its inertia after it stops propelling itself.

The amount of distance traveled as a consequence of inertia is minuscule, and therefore considered to be negligible (Purcell, 1977; Berg, 1993). This means the terminal velocity is a very good approximation of the actual velocity of the

microbe, and thus the system of equations can be simplified by ignoring inertia under low Reynolds number conditions:

$$\frac{dx}{dt} \approx \mathrm{v}_t(t) \tag{2.16}$$

The terminal velocity is reached if the drag force is in equilibrium with all other forces acting upon a mass-point of the agent $\sum F_p$.

$$\sum F_p - F_D = 0 \tag{2.17}$$

The terminal velocity can be expressed as follows:

$$\mathrm{v}_t = \frac{\sum F_p}{3\pi d_a \mu_f} \tag{2.18}$$

This simplification is reasonable for cells in agent-based biofilm models as in typical biofilms cell velocity or size is never expected to be large enough to break the low Reynolds number assumption. iDynoMiCS 2.0 distinguishes itself by implementing it, effectively halving the number of ordinary differential equations that need to be solved.

2.3.2 Repulsion, attachment, and motor force

The other forces that can act upon the agent may be quite different from one model to the other. These forces may result from collisions with other cells or surfaces, structural forces maintaining the length of rod shaped cells or the rigidity of filaments, attachment to other agents or the substratum, motor forces, etc. Where models utilizing shoving typically only consider expansion of the biofilm due to overlap and in some cases retraction of the biofilm because of decreasing local densities, force based implementation may include a wider variety of physical interactions. Initial mass spring models mostly rely on linear springs following Hooke's law:

$$F = k\delta l \tag{2.19}$$

Where k is the spring constant and δl the difference between the current length and the rest length. Janulevicius et al. (2010) also implemented angular springs and motor forces. iDynoMiCS 2.0 can model a wide variety of physical interactions, including springs and DLVO (Derjaguin, Landau, Verwey and Overbeek) interactions, as interactions can be expressed as a standard arithmetic expression.

By default agent collisions are modeled using the Hertz soft sphere model (Stevens and Hrenya, 2005):

$$F = \frac{4}{3}\sqrt{r_{eff}}\, E_{eff}\, \xi^{3/2} \qquad 2.20$$

Where r_{eff} is the effective radius, E_{eff} the effective Young's modulus and ξ the agent overlap.

2.4 Other processes

Microbial detachment and attachment are other processes many IbMs implement. Although it can be an important factor in the development of the biofilm, there does not seem to be a consensus on how to model these processes (Horn and Lackner, 2014). A variety of factors including biological, chemical and physical processes can play a role. These processes include sloughing, for example due to increased shear force on the biofilm during backwashing, or protozoan grazing of the biofilm. Other processes or changes may affect the strength of the EPS matrix, such as hydrolysis or sudden shifts in pH (Wang et al., 2012). Some models directly model shear stress on the biofilm resulting from fluid flow (Jayathilake et al., 2017). Horn and Lackner (2014) present a variety of biofilm detachment models of which a subset is also available in iDynoMiCS 2.0. iDynoMiCS 2.0 also implements an attachment algorithm that implements a 2D or 3D random walk that can be used to determine an attachment site for newly introduced agents (coming from the bulk or other sources).

Many more processes may be modeled as part of an IbM. This includes processes such as plasmid invasion (Seoane et al., 2010), damage repair and aging (Clegg et al., 2014; Wright et al., 2020), quorum sensing and chemotaxis (Sweeney et al., 2019) and more. Although discussing detailed model implementations of these processes goes beyond the scope of this section, a key takeaway is that IbMs are well suited to capture individual variability, local interactions, and adaptive behavior. IbMs can predict unique outcomes that may be lost in continuum models, where population averages are considered.

3 Verification of solvers and algorithms

To ensure the newly developed iDynoMiCS 2.0 platform functions correctly a thorough testing and verification procedure was performed. A detailed description of this procedure is included with the supporting information of *manuscript I*, while this chapter summarizes the procedure and reports the main findings. This procedure includes single-component testing, where individual solvers or algorithms are tested, and multi-component testing, where multiple solvers or algorithms are tested simultaneously. For the single-component tests simple test cases with known outcomes are compared to the outcomes of the iDynoMiCS 2.0 solvers. For multi-component tests no analytical- or otherwise absolute solutions exist, and the iDynoMiCS 2.0 outcomes are compared to those of other biofilm models.

3.1 Force-based Mechanics

Two components have been tested to verify the correct functioning of iDynoMiCS 2.0's Force-based Mechanics (FbM). The first part is the detection of collisions. In total, 36 scenarios were created, including different agent morphologies, agent-voxel collision detection and multiple domain boundary scenarios including agent-boundary collision, but also agent-agent/voxel collision through periodic boundaries. iDynoMiCS 2.0 identifies all collisions and misses of these scenarios correctly. The second part is the collision response. Here the FbM solver is used to decompress an initial over compressed state. The correct behavior here is that the solver converges to a state were agent to agent/surface forces minimized. Over the 1000 performed iterations, the maximum interaction force dropped from 275.1 fN to 0.08 fN (figure 3.1), and thus the solver works as expected.

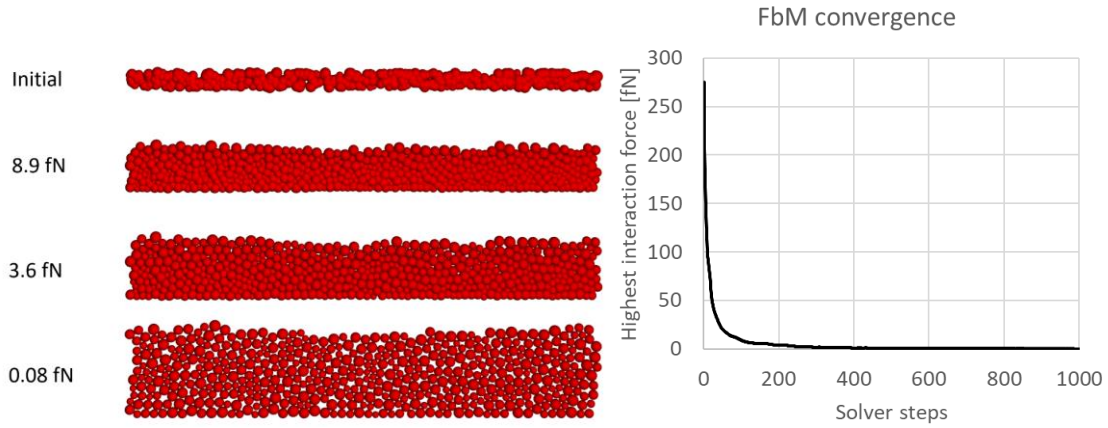


Figure 3.1. FbM leads to rapid relaxation of mechanical stress from an initially over-compressed state. Left panels from the top showing highest interaction force next to the biofilm structure: 275.1 fN for the initial state and 0.08 fN after 1,000 steps. The panel on the right shows exponential drop of the highest interaction force towards zero, demonstrating convergence of the FbM solver. This figure is reproduced from Manuscript I.

3.2 Well-mixed bioreactor model

The dynamics in the bulk region in iDynoMiCS 2.0 simulations can be described using a classical bioreactor model (Equation 2.1). The bioreactor model is solved using a basic Ordinary Differential Equation (ODE) solver. This solver was tested using two scenarios. The first scenario is a chemostat model with a single substrate (S) and a catalyst converting the substrate. This system can be described by the following differential equation:

$$\frac{dS}{dt} = \frac{QS_0}{V} - \frac{QS}{V} - \frac{mqS}{V} \qquad 3.1$$

where S_0 is the solute concentration in the inflowing medium, Q is the flow rate with dimension volume per time, V is the volume of the chemostat, t is time, q is the rate of solute consumption by the catalyst and m is the mass of the catalyst. The steady-state solution for this differential equation is:

$$S^* = \frac{QS_0}{Q + mq} \qquad 3.2$$

The solver converged to the exact steady state solution (figure 3.2 A).

The following scenario includes growing biomass. This can be implemented as a classic chemostat model with a deterministic outcome, or as an individual based model which introduces stochasticity. Both cases were implemented. The biomass or agents grow according to Monod kinetics and have a probability of being washed out of the chemostat creating a new steady state scenario where solute concentration, agent growth and wash out balance (A detailed

description of this model is included with supporting information of *Manuscript I*). The steady state solutions for the solute S^* and total agent mass P^* are:

$$S^* = \frac{{}^{Q}/_{V}\, K_S}{\mu_{max} - {}^{Q}/_{V}} \quad 3.3$$

$$P^* = Y(S_0 - S^*) \quad 3.4$$

Where μ_{max}, K_S and Y are the maximum specific growth rate, substrate affinity and the yield of the agent. The iDynoMiCS 2.0 simulation resulted in concentrations that closely approached those of the analytical solution, the differences in steady state concentration were 0.0008% and 0.0006% respectively figure 3.2 b and d. The stochastic model introduces a degree of variance to the model outcomes (figure 3.2 c). However, over 9 repetitions revealed that the model consistently converges to the same expected outcome, without any evident bias, with average steady state concentrations of S* = 6.64 $mg\ L^{-1}$ and P* = 498.40 $mg\ L^{-1}$ over the 9 repetitions.

3.3 Spatial reaction and diffusion model

To test (bio-)chemical reactions and diffusion in a spatial explicit domain, again two test scenarios were constructed. The first scenario includes a row of catalyst agents on an inert surface at the base of a spatial compartment and a diffusion boundary layer at a set distance over this row of catalyst agents. This matches a 1E system of a catalyzing surface (see the supporting information of *Manuscript I*) for which a steady state solution can be calculated:

$$S^* = \frac{D\ A\ S_0}{\Delta x\ m\ q + D\ A} \quad 3.5$$

where S* is the steady-state concentration at the catalyst surface, D is the diffusivity of the solute S, A is the surface area of the catalyst and Δx is the depth of the diffusion-dominated boundary layer. The iDynoMiCS 2.0 solvers converged to the predicted surface concentration of 1.8 mg/l (figure 3.2 d).

The second scenario includes a thicker layer of catalyst agents and is meant to study the concentration profile within the biomass region. For a homogeneously distributed catalyst following first order kinetics on an inert surface the following analytical model for the solute concentration can be used:

$$\frac{dS}{dt} = D\nabla^2 S - kS \quad 3.6$$

Where ∇^2 is the Laplacian operator and k is the reaction rate constant. The agents are placed on an inert surface where a gradient forms perpendicular from the surface, at the inert surface $x = 0$, there is no solute gradient $\frac{dS}{dx} = 0$. With a known solute concentration S_b at the top of the biofilm x_b, the steady state solute concentration in the catalyst layer can be found:

$$S^*(x) = S_b e^{-(x-x_b)\sqrt{k/D}} \frac{1 + e^{2x\sqrt{k/D}}}{1 + e^{2x_b\sqrt{k/D}}} \quad 3.7$$

This solution is further elaborated in the appendix. This problem can also be solved for a catalyst layer around an inert spherical particle (also included in the appendix).

Multiple tests were performed with this setup, both with diffusion boundary layer (Figure 3.2 e) as well as without (thus having the boundary concentration matching the bulk concentration). The solute concentration profiles through the catalyst region are as expected. Further, for the test excluding a boundary layer a unit-test was constructed, sampling the numeric solution every 2 μm throughout the catalyst region revealed a mean square error (MSE) of 0.0049.

The single-component tests demonstrated that the solutions were numerically correct with deviations of <0.02% in all test scenarios. Further, 3 multi-component tests were developed, these include a large scale stress test, an iDynoMiCS 2.0 implementation of benchmark problem 3, and a test case comparing iDynoMiCS 2.0 outcomes of yield and rate strategist competition with a prior BacSim model (Kreft, 2004). These tests use multiple model components simultaneously (hence the name) and are further distinctly different from the previous test cases as there are no analytical- or otherwise known correct solutions for these problems.

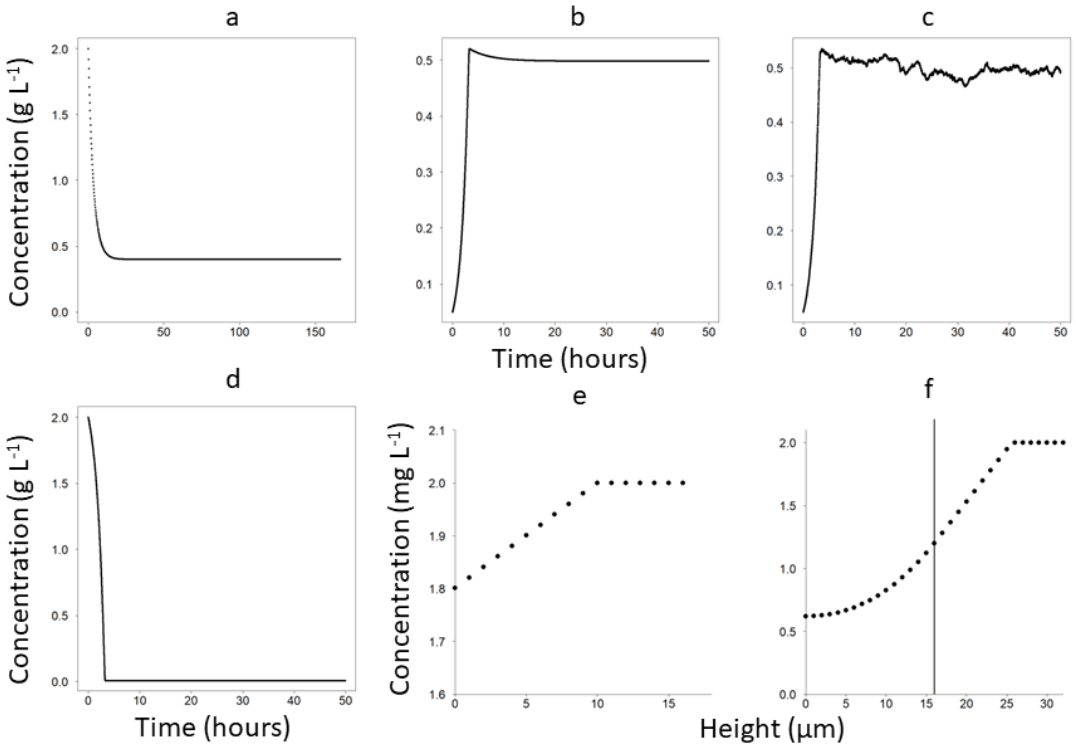


Figure 3.2. Results of numerical tests of the ODE and PDE solvers. A – Results from the non-growing chemostat population. Concentration tended towards the expected steady state of 0.4 $g\ l^{-1}$. B, c – Results from the growing chemostat population. Concentrations tended towards the expected steady states of 498.33 $mg\ L^{-1}$ for the biomass in the deterministic model (b) and a single repetition of the stochastic model (c). The expected solute concentration of 6.67 $mg\ L^{-1}$ was reached in both versions of the model. The graphs are nearly indistinguishable for the stochastic and deterministic model (d). (e) – Results from thin cell layer. Concentration at biofilm surface matched predicted concentration of 1.8 $mg\ L^{-1}$. (f) – Results from thick cell layer. The vertical line marks biofilm surface (x_b). The substrate concentration gradient was linear in the diffusion boundary layer above the biofilm surface and then decreased towards zero at the inert boundary at height 0, as expected.

3.4 Stress test

The stress test was developed to verify whether iDynoMiCS 2.0 can simulate large scale model systems. In these scenarios 1000 Ammonium Oxidizing Organisms (AOO) and 1000 Nitrite oxidizing Organisms (NOO) are randomly distributed over 500 µm by 500 µm to simulate a simple nitrifying biofilm. The agents produce EPS particles coupled to their growth rate. When the mass of an agent drops below 20% of its division mass it 'dies' and becomes inert. The model was adopted from Hubaux et al. (2015), and updated to include EPS production and cell death. The simulated biofilm reached over 10 million agents in 175 days of simulated time (figure 3.3), the simulation ran for 11.34 days on a single processor. This simulation reached approximately two orders

of magnitude more agents than the largest iDynoMiCS 1 simulations, currently NUFEB (Li et al., 2019) and iDynoMiCS 2.0 are the only two platforms shown to be capable of biofilm simulations at this scale, Biocellion (Kang et al., 2014) has been used in cell sorting simulations with over 1 billion agents, but there is no information on large biofilm simulations.

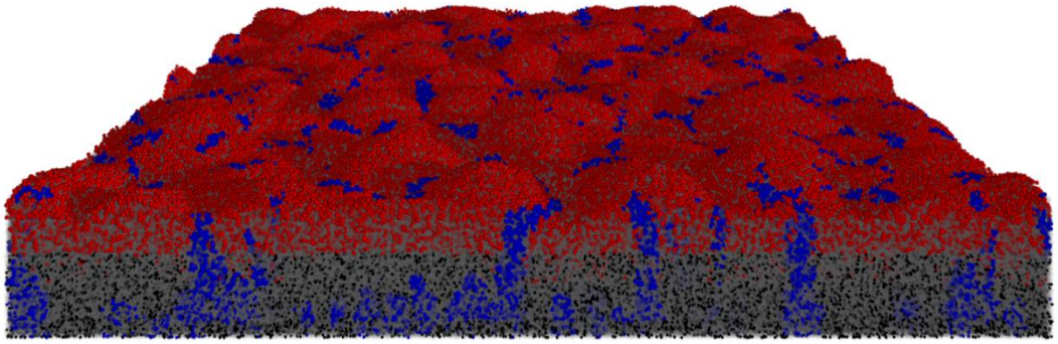


Figure 3.3. iDynoMiCS 2.0 is capable of simulating large 3D biofilms. A nitrifying biofilm was initiated with 1,000 AOO (red) and 1,000 NOO (blue) in a 500x500x500 µm domain. Both species produced EPS particles (gray semi-transparent). Agents that dropped below 20% of their division mass as a result of endogenous respiration (maintenance metabolism) became inactive (black). The 175-day biofilm contained 1.02×10^7 agents (bacteria and EPS particles). This figure is reproduced from Manuscript I.

3.5 Benchmark problem 3

To facilitate comparisons between different modeling approaches and establish the effects of different model designs, an International Water Association (IWA) task group was established. The task group developed Benchmark 3 (BM3), a model for microbial completion in biofilm systems. The model was implemented in a variety of modeling platforms to establish the effects of different model designs and simplifying assumptions on simulation outputs (Wanner et al., 2006). The benchmark problem comprises a model biofilm with heterotrophic bacteria oxidizing a source of Chemical Oxygen Demand (COD) and autotrophic bacteria oxidizing ammonia to nitrate. By implementing Benchmark 3 in iDynoMiCS 2.0, the platform could be compared to a variety of other model approaches, the comparison includes:

- W – a one-dimensional continuum biomass model run on the AQUASIM software (Reichert, 1994) and developed by Peter Reichert and Oskar Wanner (Reichert and Wanner, 1997; Wanner and Reichert, 1996).
- M1 – a variant of the W model with a fixed boundary-layer thickness by Eberhard Morgenroth (Morgenroth and Wilderer, 2000).

- DN – a two-dimensional cellular automaton model developed by Daniel Noguera and colleagues (Noguera et al., 2004).
- CP – a two-dimensional individual-based model, with biomass spreading via shoving, developed by Cristian Picioreanu and colleagues (Picioreanu et al., 2004).
- NUFEB – A three-dimensional individual-based model that uses a platform derived from a molecular dynamics simulator by Li et al. (Li et al., 2019).
- iDynoMiCS 1 – An individual-based model by Lardon et al. (2011) used here for 2D simulations. This platform is the precursor to the one described in this paper, and the implementation of BM3 is very similar.

Even though the comparison includes vastly different modeling approaches (continuum based, Cellular automaton and other individual based models), no major differences were observed in the resulting biofilm. The only significant difference observed is steady state COD concentration in the high ammonium test scenario for iDynoMiCS 2.0 using FbM (Hotelling's T^2 tests p-value of 0.0431).

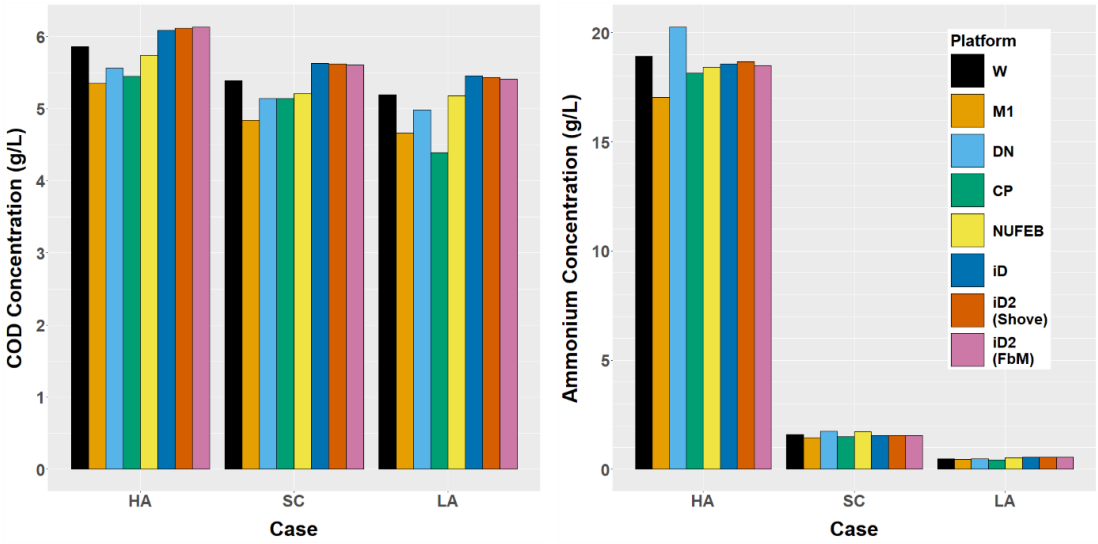


Figure 3.4. Comparing steady states in BM3. Steady state organic carbon (Chemical Oxygen Demand, COD) and ammonium concentrations in the bulk liquid for the three different BM3 cases (HA: High ammonium, SC: Standard case, LA: Low ammonium) across 7 model implementations (W: Wanner, M1: Morgenroth, DN: Dan Noguera, CP: Cristian Picioreanu, NUFEB: NUFEB, iD: iDynoMiCS 1, iD2: iDynoMiCS 2.0, either with shoving algorithm similar to iD or the new Force-based Mechanics). This figure is reproduced from Manuscript I.

3.6 Microbial competition

The final multi-component test is the implementation of a competition model by Kreft (2004), where a Yield Strategist (YS) with a high growth yield and low maximum growth rate competing against a Rate Strategist (RS) with a higher growth rate and lower yield in several initial configurations. In this model initial small differences can be amplified by positive feedback, leading to divergent outcomes. Our simulations revealed the same qualitative outcomes (figure 3.5), even though initial agent density thresholds that separate regions where different strategies win were somewhat shifted.

This effect may be explained by the difference in biomass spreading algorithms. Shoving, as used in BacSim, leads to more open spaces and local agent mixing, while iDynoMiCS 2.0's FbM, without the explicit modeling of EPS, results in denser biofilms. To keep the overall biofilm density similar, the agent density was reduced by 47% in iDynoMiCS 2.0. Shoving can model the effect of EPS production implicitly by generating space between agents, while EPS particles need to be included explicitly to generate space when using FbM simulations. With iDynoMiCS 2.0 both approaches can be used.

The extensive testing procedure performed on iDynoMiCS 2.0 confirms correct solver and algorithm behavior. The stress test demonstrates that iDynoMiCS 2.0 can be used for model scenarios at a far larger scale than its predecessor, while Benchmark 3 and the reproduction of the *biofilms promote altruism* model show that iDynoMiCS 2.0 can generally reproduce prior modeling studies even though small differences do exist.

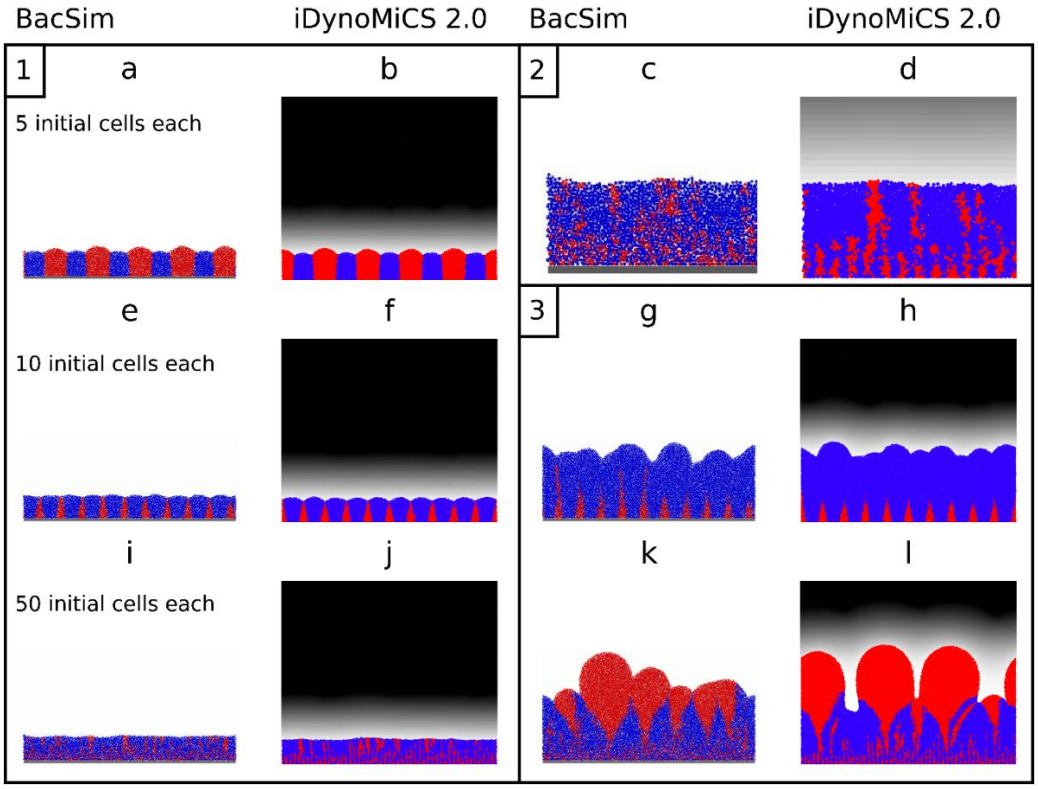


Figure 3.5. Biofilms promote altruism case study. Rate Strategist (RS, blue) and Yield Strategist (YS, red) competitions using the shoving algorithm in BacSim (Kreft, 2004) (reproduced from "Kreft J-U (2004) with permission) were replicated in iDynoMiCS 2.0 with its force-based mechanics. Cells were initially placed in alternating, equidistant positions with increasing density from 5 cells per strategy (Scenario 1: a-b), 10 cells each (Scenario 2: e-h) to 50 cells each (Scenario 3: i-l and c-d). iDynoMiCS 2.0 panels show local oxygen concentration as a linear gray-level gradient from zero oxygen (0 mg/L, white) to a maximum concentration (S_{ox_bulk} = 1 mg/L, black). Box 1 shows 3-week-old biofilms. Box 2 zooms into panels i and j. Box 3 shows 10-week-old biofilms developed from the 3-week-old biofilms shown in the same position on the left. This figure is reproduced from Manuscript I.

4 The estimation of microbial growth characteristics

Microbes are found in many different environments and are found to be able to utilize a broad array of resources. Typically, microbes are not found to be generalists, but rather filling specific ecological niches. Microbes perform biochemical conversions in order to generate the energy they need (catabolism) and in order to grow and reproduce, or to synthesize other beneficial products (anabolism). The characteristics of these biochemical conversions are central to many microbial models. Classic Monod kinetics is one of the simplest and commonly used models describing microbial growth:

$$\mu = \mu_{max} \frac{S}{K_s + S} \quad 4.1$$

The Monod equation describes the specific growth rate (μ) as function of the half-saturation coefficient (K_s) defined as the substrate concentration (S) at which the growth rate exactly half of the maximum growth rate ($\mu = \mu_{max}/2$). The equation describes the growth of an organism limited by a single limiting substrate. The substrate-to-biomass- or growth-yield Y_s is a stoichiometric parameter that discloses what amount of substrate is consumed per amount biomass produced. The Monod equation is an empiric relationship where μ_{max}, K_s and Y_s are typically determined experimentally. Alternatively, the biomass specific substrate utilization rate q_s can be expressed using the Monod equation:

$$q_s = q_s^{max} \frac{S}{K_s + S} \quad 4.2$$

Here q_s^{max} is the maximum substrate utilization rate per unit of biomass. Classic Monod kinetics ignores endogenous respiration, cellular maintenance or biomass decay which can constitute a non-negligible amount of the substrate consumed by the microbe, and thus skew the apparent growth-yield. To account for this effect many studies include a decay constant b_{decay}.

Ideally these parameters are determined experimentally, the required data is not always available. While kinetic parameters are typically determined with pure cultures of free-living cells, these cultures can prove hard if not impossible to obtain for some microbial species found in biofilms. Examples are the efforts made to quantify the growth parameters of multiple anammox and comammox species, where allot of research efforts has gone into obtaining en-

richment cultures used for kinetic quantifications, as obtaining truly pure cultures remained unsuccessful. Multiple methods and empirical relationships have been developed to estimate these parameters based on the thermodynamics of the microorganism. These methods may provide initial parameter estimates for microorganisms that are found in biofilms but have not sufficiently been characterized for modeling purposes. This chapter provides a brief overview of these methods and reflects on the use of these methods in relation to individual based biofilm models.

4.1 Calculation of microbial yield

Generally, three different methods to estimate microbial growth yield are used. This includes the thermodynamic electron equivalent method (McCarty, 1965; Rittmann and McCarty, 2018), which estimates the growth-yield based on a microbes energy transfer efficiency ε. The black box model to calculate biomass yields based on Gibbs energy dissipation (Heijnen et al., 1992), which estimates the growth-yield based on the amount of energy dissipated ΔG_{Dis} per amount of newly formed biomass. And the ATP-balancing method (Kleerebezem and Van Loosdrecht, 2010), which shares a lot of principles with the previous methods, but instead of utilizing an empiric ΔG_{Dis} relationship, it describes the growth-yield in terms of ATP-yield per amount of substrate consumed Y_{ATP}^{Cat} and biomass yield per amount of ATP consumed Y_{ATP}^{An}.

All three methods are based in thermodynamics and utilize the Gibbs free energy change of the biochemical conversions performed by the microbe. For any reaction this Gibbs free energy change can be expressed as follows:

$$\Delta G_r = \Delta G_r^0 + RT \sum_{i=1}^{n} v_{ir} \, ln a_i \qquad 4.3$$

Where n is the number of constituents in the reaction v_{ir} is the stoichiometric coefficient for a constituent in reaction r, a_i is the activity of the constituent and ΔG_r^0 is the free energy of formation, R is the universal gas constant and T the temperature.

4.1.1 Thermodynamic electron equivalent method

To estimate the heterotrophic yield, Rittmann calculates ΔG_s, the energy required to synthesize one unit of biomass from a given carbon substrate and uses this number to calculate how much equivalents of electron donor must be oxidized to supply this energy. ΔG_p is obtained by subtracting the electron-donor half-reaction free energy from the electron-acceptor half-reaction free energy

(Rittmann). ΔG_{pc} is obtained by multiplying the energy requirement per gram of cells with the electron equivalent of the cells (depending on the nitrogen source). The energy transfer efficiency ϵ corrects for energy lost with electron transfer, transfer efficiencies between 0.55 and 0.7 are typical (Rittmann and McCarty, 2018).

$$\Delta G_s = \frac{\Delta G_p}{\epsilon^n} + \frac{\Delta G_{pc}}{\epsilon} \quad 4.4$$

When energy is required for intermediate production $n = +1$ in case energy is released $n = -1$.

To sustain growth, A equivilents of electron donor must be oxidized.

$$\Delta G_s = -A\epsilon\Delta G_d \quad 4.5$$

Were ΔG_d is the amount of energy released per amount oxidized electron donor.

Thus:

$$A = -\frac{\frac{\Delta G_p}{\epsilon^n} + \frac{\Delta G_{pc}}{\epsilon}}{\epsilon\Delta G_d} \quad 4.6$$

The growth-yield is then calculated as follows:

$$Y_s = \frac{f_s^0 M_c}{n_e L} \quad 4.7$$

Where M_c is the specific weight of 1 Cmol biomass, n_e the number of electron equivalents per Cmol biomass, L electron donor mass per amount of electron equivalents, and f_s^0 the fraction of electrons used in cell sythesis:

$$f_s^0 = \frac{1}{1 + A} \quad 4.8$$

4.1.2 Energy dissipation method

Heijnen describes microbial growth in terms of an anabolic reaction An, which represents the conversion of substrates to produce 1 Cmol of biomass, and which requires an amount of energy ΔG_{An}. And a catabolic reaction Cat, which represents the conversion of substrates to generate an amount of energy ΔG_{Cat}. Finally, Heijnen describes an amount of energy dissipated ΔG_{Dis}.

The Gibbs free energy change associated with full metabolism normalized to 1 Cmol biomass is than written as follows:

$$\Delta G_{Met} = \lambda_{Cat}\Delta G_{Cat} + \Delta G_{An} + \Delta G_{Dis} \quad 4.9$$

Note that since ΔG_{Met}, ΔG_{An} and ΔG_{Dis} are normalized for 1 Cmol biomass they can be expressed in kJ/Cmol. The net energy change of the metabolism is assumed to be zero ($\Delta G_{Met} = 0$) and the ratio λ_{Cat} can be found by solving the energy balance. Heijnen proposed the following empiric equation to estimate ΔG_{Dis}:

$$\Delta G_{Dis} = 200 + 18(6 - C_l)^{1.8} + e^{((3.8-\gamma_D)^2)^{0.16}(3.6+0.4C_l)} \quad 4.10$$

Where C_l is the carbon chain length of the carbon source and γ_D the oxidation state. Heijnen further proposed that in cases with weak electron donors in the anabolic reaction, reversed electron transfer is required and $\Delta G_{Dis} = -3500$ kJ/Cmol can be used. The growth-yield is than defined as follows:

$$Y_s = \frac{1}{\lambda_{Cat}Y_s^{Cat} + Y_s^{An}} \quad 4.11$$

With Y_s^{Cat} and y_s^{An} are the stoichiometric coefficients of substrate s in the catabolic and anabolic reactions respectively.

4.1.3 ATP-balancing method

Rather than balancing the catabolic energy production with the anabolic energy consumption and energy dissipation, Kleerebezem and van Loosdrecht (2010) instead consider an ATP production and consumption balance:

$$\lambda_{Cat}Y_{ATP}^{Cat} + \lambda_{An}Y_{ATP}^{An} = 0 \quad 4.12$$

The growth yield is then expressed as follows:

$$Y_s = \frac{1}{y_s^{An} - Y_s^{Cat}\frac{Y_{ATP}^{An}}{Y_{ATP}^{Cat}}} \quad 4.13$$

4.2 Maintenance energy

Endogenous respiration, cellular maintenance or biomass decay can be expressed as an amount of substrate that needs to be consumed per hour per amount of biomass m_s, or in terms of decaying biomass per hour per amount of biomass. Since the amount of biomass that needs to grow to compensate the decayed biomass has a specific substrate requirement these values can be converted. Heijnen and Dijken (1992) found that for a large range of microbial

species the energy requirement for maintenance m_G, can be estimated with an Arrhenius function with an activation energy of 69 kJ/mol:

$$m_G = -4.5 \cdot exp[\frac{-69 \cdot 10^3}{R}(\frac{1}{T} - \frac{1}{298})] \quad 4.14$$

The estimation proved applicable in a wide range of temperatures ($5 - 75^{\circ}C$) where the type of electron donor and acceptor are of minor importance (Kleerebezem and Van Loosdrecht, 2010).

The energy requirement for maintenance can be written in terms of substrate requirement by dividing by the Gibbs free energy of the catabolic reaction.

$$m_s = \frac{m_G}{\Delta G_{cat}} \quad 4.15$$

Apart from energy requirements and biomass synthesis, microbes can consume substrates to produce other anabolic products. This can be expressed using the Herbert-Pirt substrate distribution relationship:

$$q_s = a\mu + bq_p + m_s \quad 4.16$$

Where q_p is the biomass specific product rate and a and b are stoichiometric parameters for substrate utilization for the growth and product reaction. The estimation of these parameters goes beyond the scope of this chapter, but a detailed procedure can be found in (von Stockar, 2013).

4.3 Maximum growth rate

The maximum specific substrate utilization- and microbial growth rate can be estimated using the maximum biomass specific electron-transfer rate q_e^{max}. Rittmann and McCarty (2018) claim that value should be about 1 electron equivalents per gram VSS (volatile suspended solids) per day at 20℃. The maximum substrate utilization rate is then obtained as follows:

$$\mathrm{q}_s^{max} = q_e^{max}/(1 - f_s^0) \quad 4.17$$

Which can be adjusted for different temperatures using the following relation:

$$\mathrm{q}_{s,T}^{max} = \mathrm{q}_{s,T^R}^{max}(1.07)^{(T-T^R)} \quad 4.18$$

Heijnen and Kleerebezem (2010) proposed the following correlation for q_e^{max}.

$$q_e^{max} = -3 \cdot exp[\frac{-69 \cdot 10^3}{R}(\frac{1}{T} - \frac{1}{298})] \quad 4.19$$

Which can be converted into a maximum specific growth rate using:

$$\mu^{max} = \frac{q_e^{max} \cdot \frac{-\Delta G_{cat}}{\gamma_D^*} - m_G}{\Delta G_{Dis}} \quad 4.20$$

Were γ_D^* is the number of electrons transferred in the catabolic reaction.

4.4 Substrate affinity

The substrate affinity, also known as the half-saturation constant K_s, is possibly the hardest parameter to predict. This becomes evident when reviewing the wide range of parameter values that can be found in literature, for example reported ammonia affinity for ammonia oxidizing microbes differ over 4 orders of magnitude (Kits et al., 2017). Part of this disparity may be explained by how they are obtained. When substrate affinities are determined with aggregated biomass, this effectively lumps together the substrate affinity with the transport processes occurring in the system, as the microbes will be exposed to lower substrate concentrations than measured in the bulk liquid due to transport limitations. In this case the constant is often referred to as the apparent affinity constant $K_s^{apparent}$. Transport limitations can still play a role in cultures of suspended cells, but the effect is expected to be smaller. Researchers have made efforts to compensate for these transport effects in order to be able to compare K_s values in different experimental systems (Shaw et al., 2015, 2013). Affinity constants without transport limitations are called intrinsic K_s. Intrinsic affinity constants should be used for IbMs, as matter transport is modeled explicitly in this case. It should be taken into account that also for intrinsic affinity constants, large variation in observed values exist.

For typical molecules involved in the microbial metabolisms, intrinsic K_s values are expected be < 1 mg/l. For an unknown species, under normal conditions, intrinsic K_s may be guessed to be in the range 0.1 to 0.001 mg/l based on K_s values of other species that perform the same conversion. Similar to previously discussed parameters, methods to estimate K_s based on thermodynamic properties have been proposed. Heijnen proposed affinity constants should be close to threshold concentrations that results in a ΔG_{Cat} of 15 kJ/mol, as this is the energy required to generate the required proton motive force (von Stockar, 2013). Alternatively K_s could be linked to the thermodynamic equilibrium constant K, (Liu et al., 2003) proposed $K_s = K^{-1}$ which results in:

$$K_s = e^{\Delta G_r^0 / RT} \quad 4.21$$

However, because of the large variation of observed K_s values, caution should be taken before using these values to parameterize a model. When the local substrate concentration is high relative to the K_s value, the impact of an error in K_s in the estimated growth or substrate utilization rate should remain relatively low.

4.5 Limitations and considerations

Studying microbial energetics is useful to identify the thermodynamics boundaries of microbial metabolisms, and to explore potential ecological niches. A great example of this is the total oxidation of ammonia by so called complete ammonia oxidizing '*comammox'* bacteria. Their ecological niche was first predicted to exist in environments with strong selective pressures for high yields (Costa et al., 2006). Based on a trade-off of between growth rate and yield dependent on the metabolic pathway length proposed by Pfeiffer and Bonhoeffer (2004). The niche was identified again in a study using a thermodynamics based model by van de Leemput et al. (Leemput et al., 2011), together with a niche for a still undiscovered organism performing nitrite dismutation to nitrate and dinitrogen gas. Until comammox bacteria were finally identified in various oligotrophic environments (Daims et al., 2015; van Kessel et al., 2015; Pinto et al., 2016; Palomo et al., 2016).

While the methods reviewed in this chapter can prove useful to estimate growth characteristics of un- or under quantified microbes, it is important to keep in mind that these are generalized *'rule of thumb'* methods that do not take the intricacies of the occurring biological processes into account. While from an evolutionary point of view, it makes sense that microbes optimize their metabolism to approach the best thermodynamic efficiency possible, microbes may make trade-offs, sacrificing one beneficial trait for another one.

One of these trade-offs is that between growth yield and growth rate, as observed with nitrification comammox bacteria or divided by ammonia- and nitrite oxidizing bacteria (AOB and NOB), a concept which is also addressed in more detail in *Manuscript I*.

Other microbes may make an anabolic investment to decrease their sensitivity for extreme environments (extremophiles), resist environmental toxins, or make other beneficial anabolic investments such as the production of EPS.

5 Model analysis and parameterization

It is important to be able to relate experimental observations with those found in modeling studies and vice versa. Historically there has been a strong focus on (bio-)chemical conversion processes and resulting solute concentrations, which is a metric that can easily be compared. Often the biofilm's spatial structure is a secondary concern, even though its impacts and importance are becoming more apparent. Most population level models simplify internal interactions, the spatial and community structure and thus they are typically less parameter rich compared to IbMs where these phenomena are explicitly modelled (spatial and community structure are emergent properties of IbMs). Models are often only calibrated using solute concentration data, resulting in models that poorly predict spatial structure. For many individual based models, a rigorous parameterization or calibration procedure is often not performed as their computational intense nature and relatively large number of input parameters can make this procedure very time consuming. The calibration of IbM is further complicated by model stochasticity, decision based behavior and very limited observation data.

This work implements analysis tools enabling direct comparisons between experimental and model observations not only on a chemical and population level, but also in terms of the spatial distribution of microbial species and the resulting biofilm structure. This work further implements sensitivity analysis and parameter optimization methods that are relatively computationally efficient and can handle the 'noisy' (stochasticity and decision based behavior) nature of IbMs. This includes the Elementary Effects (EE) method (Morris, 1991; Campolongo et al., 2007) for global sensitivity analysis and model parameter optimization procedure utilizing a Genetic Algorithm (GA) (Michalewicz and Schoenauer, 1996). These are two essential steps required for model calibration (van Waveren, 1999), subsequently uncertainty analysis and finally model validation is still required to finish the calibration procedure.

The full procedure enables the development of IbMs that can predict both chemical performance, population distributions as well as the spatial structure of biofilms. The detailed modeling and analysis of microbial aggregates has the potential to reveal more information about the biological system than can currently be obtained with laboratory techniques alone. Sensitivity analysis and parameter optimization can be used to improve the model and improve our understanding of the growth and behavior of biofilm dwelling microorganisms, even if these microorganisms have never been obtained in pure cultures. The

procedure should help reduce the gap between experimental and model-based biofilm work and facilitate pattern oriented modeling (POM) (Grimm et al., 2005), and facilitate parameter estimation informed by the emergent properties of the biofilm where individual observations may be difficult or even impossible to obtain.

5.1 Analysis of an individual based partial nitritation anammox biofilm model

An individual based partial nitritation anammox (PNA) model was constructed to test these procedures (figure 5.1). Rather than only looking at microbial guilds, where each guild is represented in the model as a single representative organism, this model incorporates two specific microbial strains for each guild: ammonia oxidizing bacteria (AOB), nitrite oxidizing bacteria (NOB) and anammox bacteria (AMX). Some important differences include: AOB1 has a strong affinity for oxygen, AOB2 has a strong affinity for ammonium, NOB1 has a strong affinity for nitrite, NOB2 has a strong affinity for oxygen, AMX1 is more sensitive to oxygen, but also has a stronger affinity for ammonium and nitrite than AMX2. The model is configured to simulate a chemostat with a continues in- and outflow. A detailed description of the model is included in the appendix (table 9.1 to 9.3).

The population composition (figure 5.2) is a model outcome, prevalent species, and extension characteristics of the biofilm can change considerably depending on the selective pressures exerted by the environment. In other words, an environment with an abundance of resources will select for fast growing organisms, whereas a famine environment will select for organisms with a better resource economy, and the biochemical kinetics of the biofilm will be different depending on the environment.

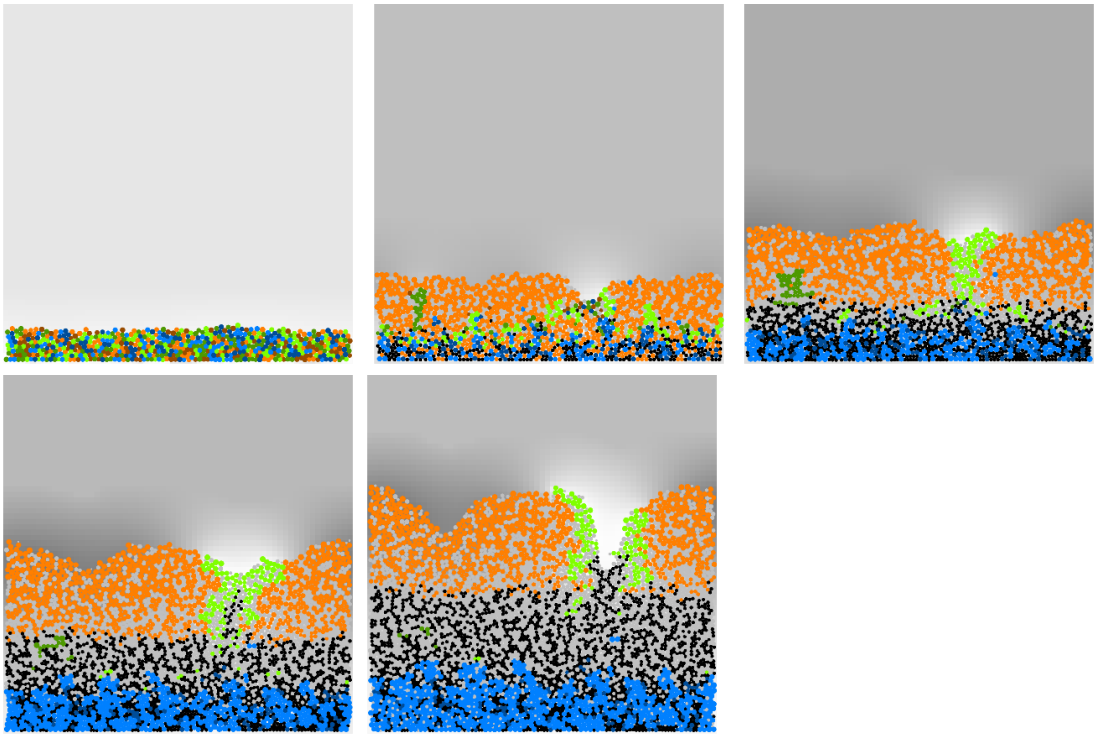


Figure 5.1. Partial nitritation anammox model. The initial model state contains 100 agents of each microbial agent species (no inert or EPS agents) randomly distributed over a solid bottom surface of the computational domain. The size of the computational domain is 128 by 256 µm, the top of the domain is truncated from the image panels. Panels show; the spatial biofilm structure upon model initiation (top left), after 50 days (top middle), after 100 days (top right), after 150 days (bottom left), after 200 days (bottom middle). The model includes 2 different strains or species of AOB (light and dark orange), NOB (light and dark green) and anammox bacteria (AMX) in light and dark blue. The simulation also includes EPS (gray) and inert cells (black). The background color indicates local nitrite concentrations, using a linear color gradient, where white indicates 0 mg N/L and black 0.5 mg N/L.

After 200 days AOB2 completely disappeared, suggesting strong competition for oxygen, while the NOB2 and AMX2 populations are slowly but steadily declining, suggesting strong competition for nitrite and/or ammonium and a limited effect from oxygen inhibition. The oxygen and ammonia concentrations drop quickly after the initial state as a population of AOB develops quickly. The AOB form much of the top layer of the biofilm where both ammonia and oxygen are available. The nitrite produced by the AOB helps to support the NOB and AMX populations. The total AMX population grows slowly but steadily at the base of the biofilm, where the oxygen concentration is lowest. A small population of NOB1 develops closer to the top of the biofilm where both nitrite and oxygen are available.

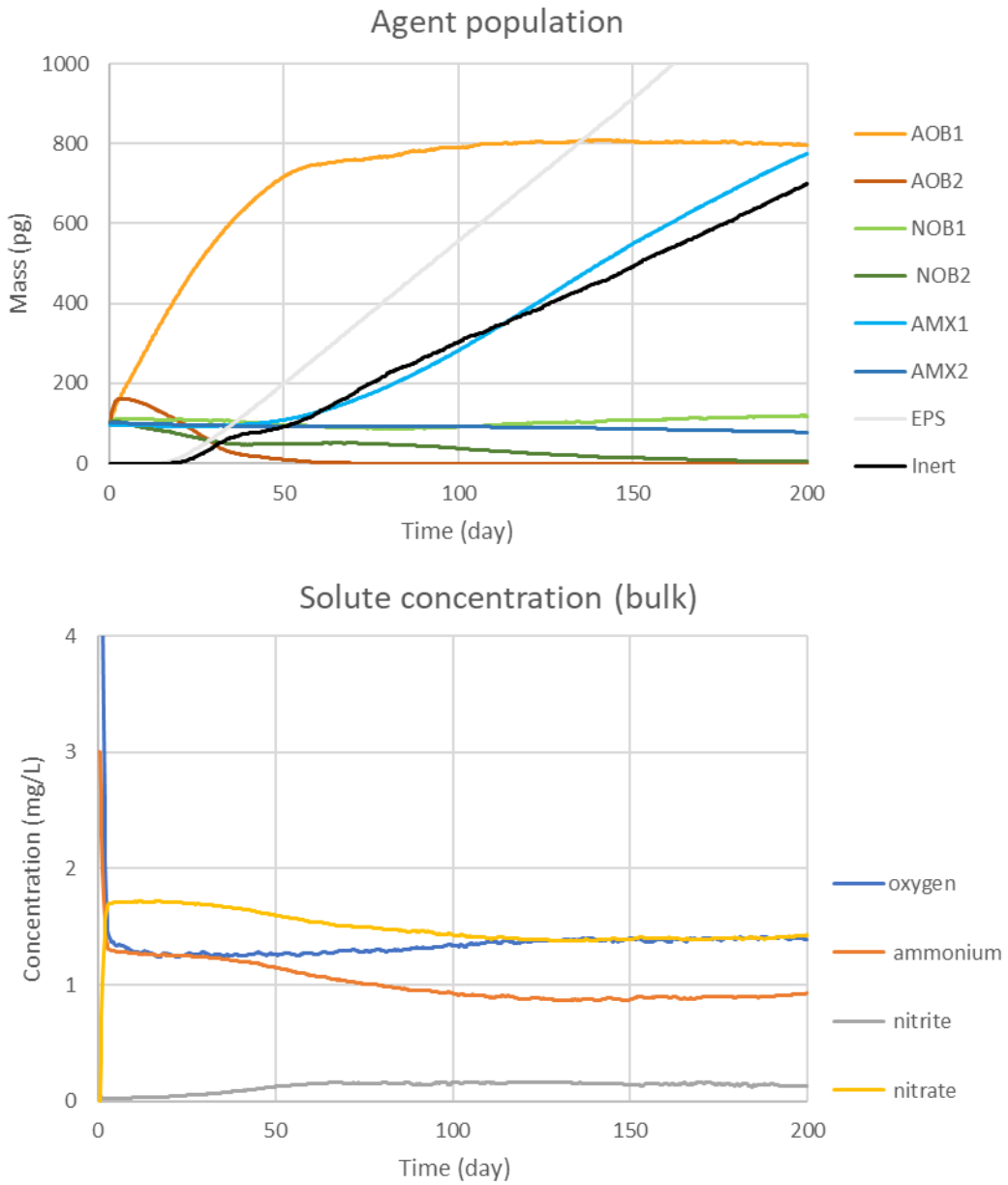


Figure 5.2. Partial nitritation anammox model. Population distribution over time (top) bulk solute concentration over time (bottom). The model includes 2 different strains or species of AOB (light and dark orange), NOB (light and dark green) and AMX in light and dark blue. The simulation also includes EPS (gray) and inert cells (black). The model includes 4 different solutes; oxygen, ammonium, nitrite and nitrate, carbon sources are considered to be non-limiting and are therefore not explicitly modeled.

5.2 Spatial analysis

The spatial analysis package allows further quantification of the biofilm structure, opening the opportunity for a more in-depth analysis. The spatial structure analysis relies on the same methodologies applied in microscopy image analy-

sis (Yang et al., 2000), thereby allowing for direct comparison between experimental and modeling work. The computational domain is discretized into voxels (small cubes aligned with the coordinate system). The voxel dimensions should be set to match the microscopy image pixel size. Here the voxel dimensions are set to 1 µm. For each voxel, collisions with agents are detected and a list of species or agents with specific traits of interest overlapping with the voxel is compiled. This allows for further spatial analysis on these traits (figure 9.1). The spatial properties of the entire biofilm can be quantified. Quantifiable parameters include the porosity, surface roughness and overall diffusion depth. It is further possible to include a more detailed analysis based on specific agent traits or species localization such as trait specific diffusion depth, trait co-localization and agent cluster size.

The (areal) porosity quantifies the amount of void space enclosed within the biofilm and is reported as a fraction. Throughout the simulation, the areal porosity always remained below 0.1%, thus the biofilm was dense with no to little enclosed void space. The surface roughness can be quantified using the fractal dimension. A perfectly smooth surface will result in a fractal dimension of 1, while a theoretical maximum fractal dimension of 2 would describe an extremely rough surface structure.

The fractal dimension (figure 5.3) of the initial state is close to 1, as can be expected for a flat biofilm. The fractal dimension increases, as the biofilm surface becomes more grooved, the fractal dimension increases to just over 1.09. Although the fractal dimension increases over time, it never gets close to the theoretical maximum of 2. The emergent biofilm surface profile is rougher than a flat surface, but the structure remains smooth overall. A higher fractal dimension would be expected for ‘tree-like’ surface structures.

The diffusion depth, measured in voxel lengths (1 µm), is the shortest distance of the agent to the surface of the biofilm. For the average diffusion depth this quantity is averaged over all agents or all agents belonging to a specific group. Diffusion depths can also be determined for a specific species or agents with a specific trait to provide information on their localization in the biofilm (figure 9.2).

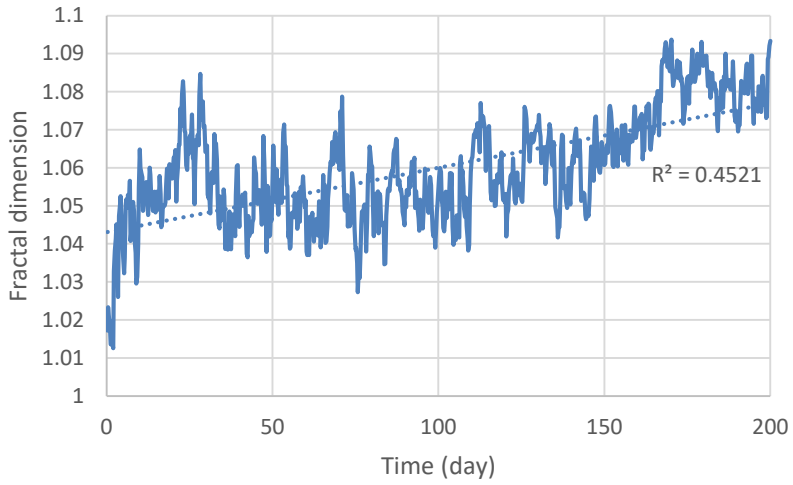


Figure 5.3. Biofilm fractal dimension over time. At the first time step, when the biofilm is almost perfectly flat a fractal dimension of 1.017 is observed. The fractal dimension quickly increases after, it fluctuates around an average of 1.06 and reaches a maximum of just over 1.094. A linear trendline reveals an increasing trend over the course of the simulation (R^2 0.4521).

In iDynoMiCS 2.0, the relative position of two groups of agents can be quantified with two metrics: the Average Diffusion Distance (ADD) and the Manders (co-localization) coefficient (Manders et al., 1993). The first is the shortest distance between an agent and an agent of the second group. The second is the number of voxels that collide with both groups divided by the number of voxels that collide only with one of the two groups. The Manders coefficient ranges from 0, agent group A is never found together with agent group B, to 1, where agent group A is always together with agent group B. Note that two Manders coefficients can be calculated per pair of agent groups. For example, agent group A may always be co-located with agent group B (coefficient of 1), while agent group B may be found both with or without agent group A.

One may argue that when agents of the two groups are in neighboring voxels this could still be considered co-localization, therefore two additional adjusted Manders coefficients are calculated. The first considers agents to be co-located if they are in directly adjacent voxels, and the second if they can be reached in 3 steps. The 3 coefficients are denoted with 1, 2 and 4 µm, after the maximum number of voxel lengths co-located agents can be removed from each other. A subset of co-localization plots for the PNA biofilm simulation are shown in figure 5.4.

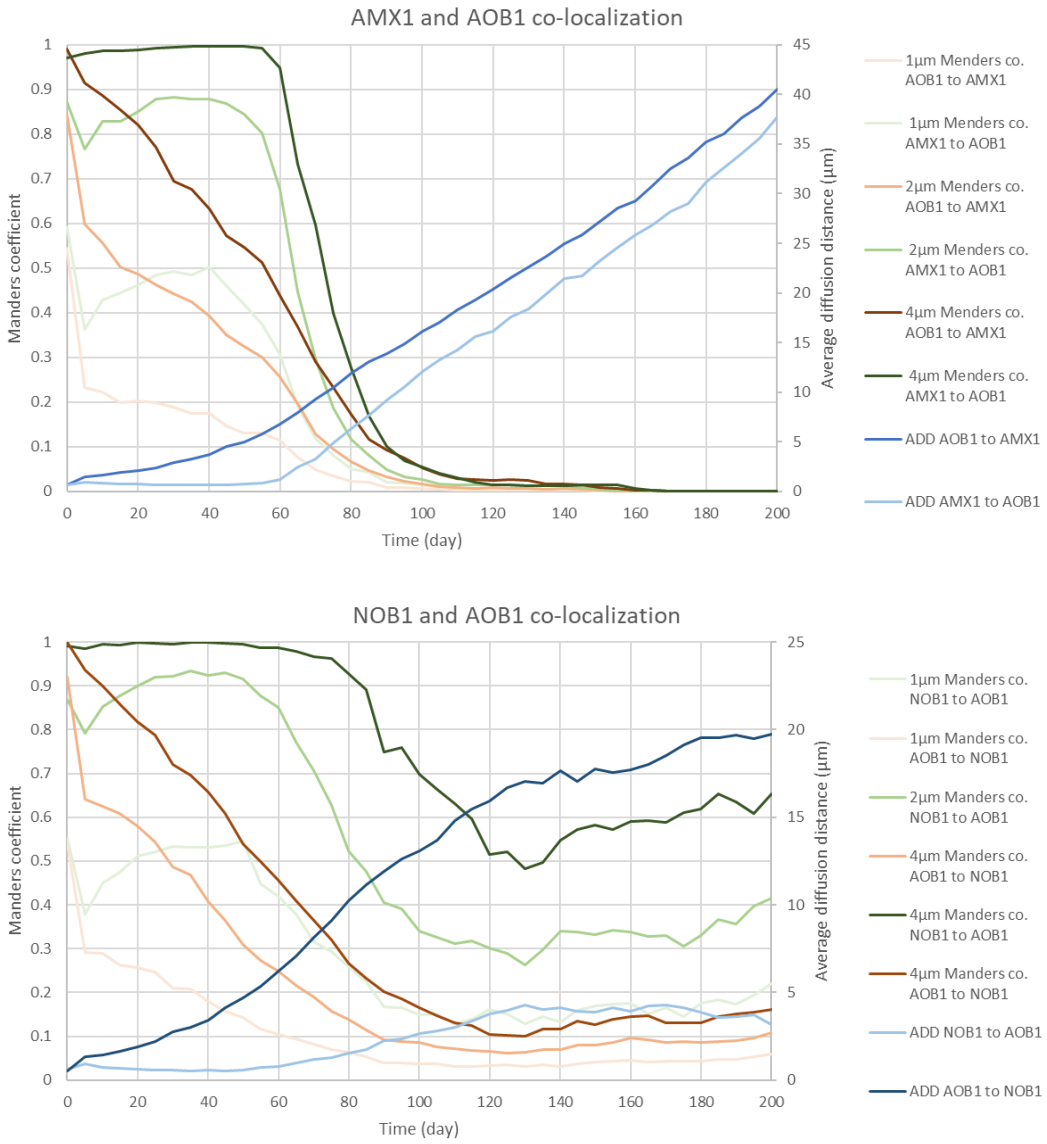


Figure 5.4. Co-localization of agent species. Agent species co-localization was quantified over the timespan of the simulation. With 6 unique agent species (ignoring inert and EPS agents), 15 unique pairs can be made, co-localization of two of those pairs are shown; AMX1-AOB1 (top) and NOB1-AOB1 (bottom). The plots show the Manders co-localization coefficient (left axis) and the ADD (right axis).

The first pair of agent groups includes the AMX1 and AOB1 agents. At the start of the simulation, there is a high degree of co-localization as at this stage the agents are randomly placed at the base of the computational domain. From this point, the co-localization coefficients start to drop, co-localization of AOB1 relative to AMX1 drops quicker than AMX1 relative to AOB1. This effect is caused by the rapid expansion of the AOB1 population at the start of the simulation. The chances of finding AOB1 agents close to an AMX1 agent are greater than vice versa. Approximately halfway through the simulation the

agent groups are nearly completely separated (Represented by a Manders coefficient of 0). Agents that are strictly not co-located can be an indicator of negative interaction, such as growth suppression through antibiotic production. In this case, however, the separation of the two groups is due to the different niches of the groups. Where AOB require oxygen for their metabolism, AMX growth is inhibited by it.

The interactions between AOB and NOB are of a different nature. The NOB are dependent on AOB for nitrite, while also competing with AOB for oxygen. Both species benefit from being localized at the surface of the biofilm where the oxygen concentration is higher, while NOB also benefits from co-localization with AOB, close to the nitrite supply. This tendency can also be seen in figure 5.4, where NOB1 co-localizes with AOB1.

5.3 Sensitivity analysis: Elementary Effects method

Apart from providing insights on the biofilm and its structure, these quantifications can be used directly in sensitivity analysis. Sensitivity analysis reveals how sensitive the model outputs are to changes in the input parameters. This procedure can also be used to reduce the parameter set to be included for subsequent parameter optimization. Parameter optimization efforts are focused on the parameters the model is most sensitive to. The procedure may further provide insight in the important processes and may guide experimental design.

With the purpose of model calibration in mind it is sensible to perform global sensitivity analysis of the model (as the aim is to uncover the sensitivity of parameters that could be unknown or uncertain). Derivative based global sensitivity measure (DGSM) with Monte Carlo sampling is a popular approach. Unfortunately, Monte Carlo based methods require a very high number of model evaluations (thousands to tens of thousands) and therefore are less suitable for notoriously computational intense models. The Elementary Effects method, also known as Morris screening and Morris method (Morris, 1991), is a good alternative as it may already provide insight after several hundred model evaluations. The method is considered a quasi-global sensitivity analysis method. With modern high-performance computing (HPC) systems, these model evaluations can be performed simultaneously. The main drawback of EE is the risk for false negatives (parameters that are wrongly considered not sensitive), due to elementary effects with opposing signs canceling each other out. This effect can be mitigated by also observing the mean of the distribution of the absolute values of the elementary effects (μ^*) as proposed by Campolongo et al. (2007).

Morris screening has been implemented as a multi-step process that allows for model evaluation on HPC infrastructure (figure 5.5). The first step (staging) has been integrated within the iDynoMiCS 2.0 framework. In this step, a pre-defined parameter space is sampled using Morris sampling. The parameter space is defined in a master file that specifies the parameter ranges for selected parameters. The master file is interpreted and a collection of child protocol files with unique parameter combinations is created. The child protocol files can be used to run concurrent model evaluations. Finally, the resulting model output is analyzed and Elementary Effects can be calculated. Morris screening takes quasi-random samples on a predefined grid distributed within a parameter space of interest. After selecting a point on this grid a perturbation is made in each of the dimensions of the parameter space, these perturbations are in random order and direction.

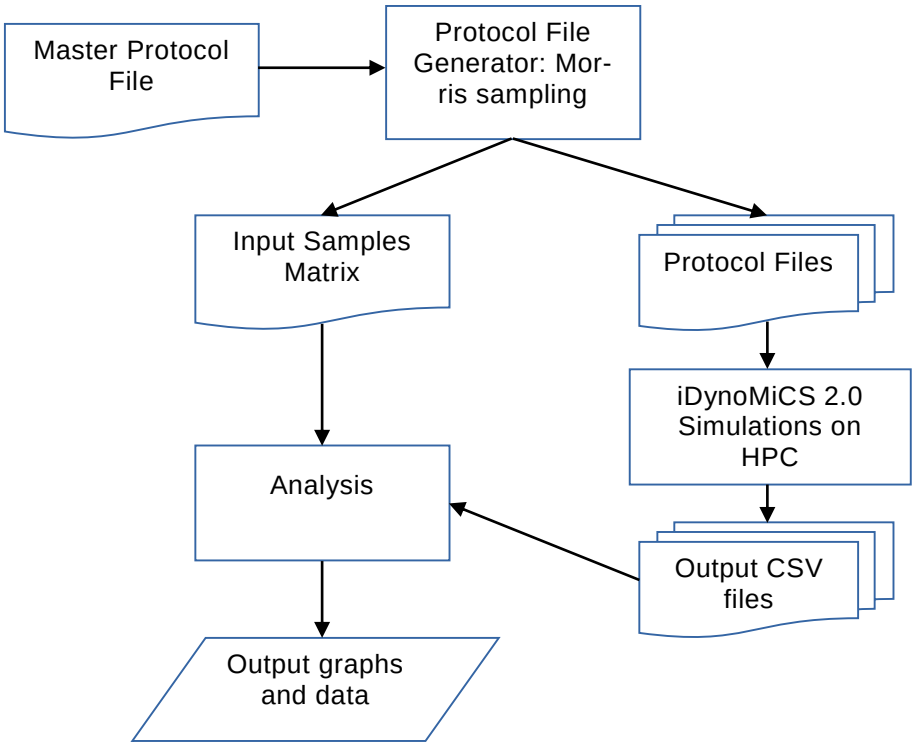


Figure 5.5. The sensitivity analysis is split-up in a few steps. First, the parameters and range of interest for the sensitivity analysis are chosen, these values are used to generate a set of child protocol files (1 per parameter set). Secondly all scenarios are evaluated on a high performance computer. Finally, model results are imported in R and analyzed.

For testing purposes, the model system was simplified to only include the agent species that become dominant (AOB1, NOB1, AMX1). The computational domain was reduced to 64 by 128 µm and the length of the simulation was scaled back to 100 days. This reduced version of the model can be evaluated in a few

minutes up to half an hour. For the 3 species, the affinity and inhibition constants, the maximum growth rates and decay rates are included in Morris screening, for a total of 13 parameters. The lower and upper bounds of the parameter space were chosen by taking -50% and +50% relative to the initial parameters. This parameter space was discretized into 1000 levels. With 50 repetitions, a total of 700 unique parameter combinations were generated. The total microbial agent mass (including inert) for all 700 simulations are shown in figure 5.6. Additional timeseries data are in the appendix, including solute concentrations (figure 9.3), surface roughness and diffusion depth (figure 9.4) and co-localization data (figure 9.5).

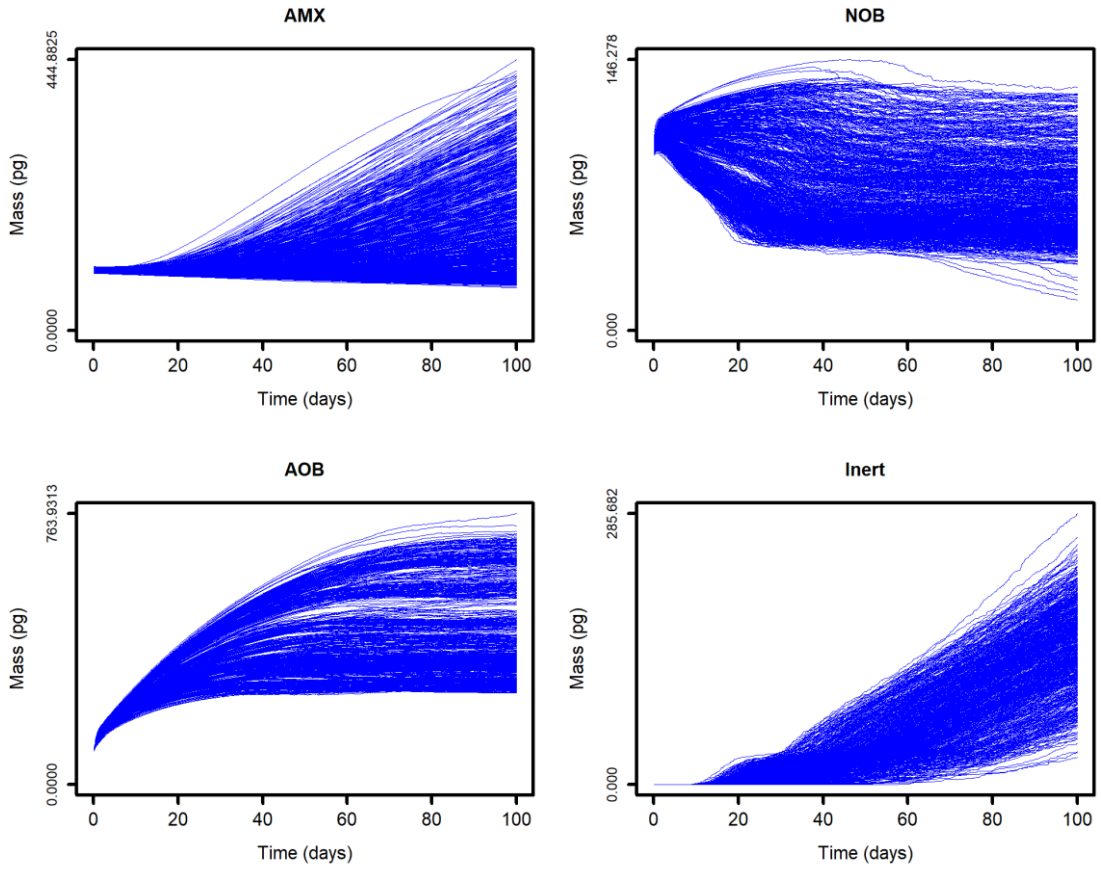


Figure 5.6. AOB (top left), NOB (top right), AMX (bottom left) and inert (bottom right) agent populations over time for all 700 simulations.

Elementary effects can reveal how perturbations in input parameters can affect specific state variables. All elementary effects are evaluated for the final state of the simulation. Plotting the mean (μ) and standard deviation (σ) of the elementary effects can help to get an overview of what input parameters a state variable is sensitive to. This type of plot, further referred to as an EE plot, is included for the total mass of the microbial agents (figure 5.7), for the final

concentration of chemical species in the bulk region (figure 5.8), for biofilm structure parameters (figure 5.9) and for a selection of agent (co-) localization variables (figure 5.10). The plotted lines indicate ±2 times the standard error of the mean (SEM) and indicate which input parameters appear to have a significant effect on the observed state variable (not in between the lines). Additionally, the μ* for all EE are calculated and included in table 9.4, for the observed state variables the μ* did not provide additional insights.

The sensitivity of the total AOB mass to the AOB decay rate (b_{AOB}) is quite evident (figure 5.7). Furthermore, small but significant effects from the oxygen and ammonia affinity can be seen, as well as an effect from the NOB decay rate. A higher NOB decay rate may positively affect the total AOB mass due to reduced competition for oxygen in simulations with less NOB agents.

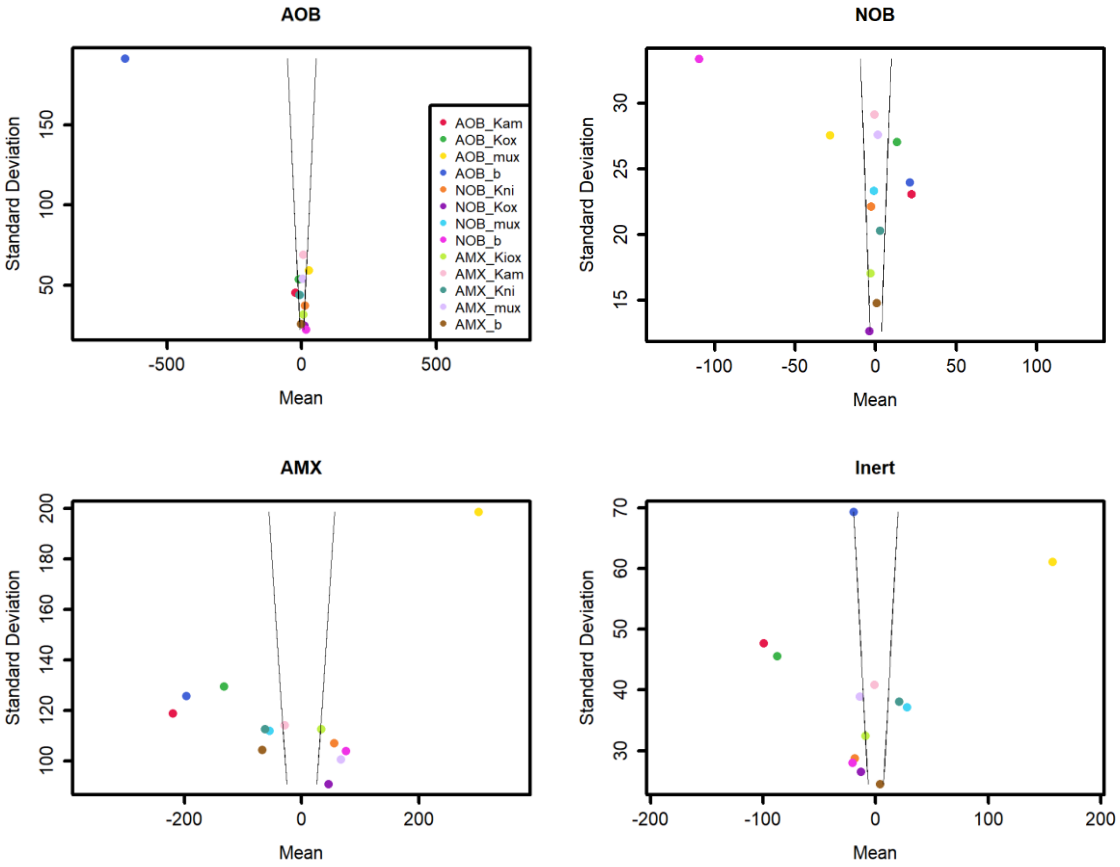


Figure 5.7. EE plots for AOB (top left), NOB (top right), AMX (bottom left) and inert (bottom right) total mass (pg). The lines indicate ±2 times the standard error of the mean (SEM).

The same input parameters appear to have a significant effect on the total NOB mass (although with different estimated effects). The importance of AOB properties on the total NOB mass can be explained by the AOB-NOB interactions, NOB is dependent on nitrite production by AOB, but competes for oxygen. For

AMX, significant effects of nearly all tested parameters are found. Since AMX is inhibited by oxygen, AMX agents benefit from the AOB and NOB oxygen consumption. AMX is also dependent on the nitrite produced by AOB, while it competes with AOB for ammonia and with NOB for nitrite.

AOB properties also play an important role in the bulk solute concentrations (figure 5.8). A large portion of the agents belong to this group while other groups have dependent and competitive interactions with AOB. Although NOB only consume small amounts of ammonium for biomass synthesis, the NOB oxygen affinity, nitrite affinity and decay rate also have a significant effect on the ammonium concentration. Possibly due to competition for oxygen with AOB or the competition for nitrite with AMX.

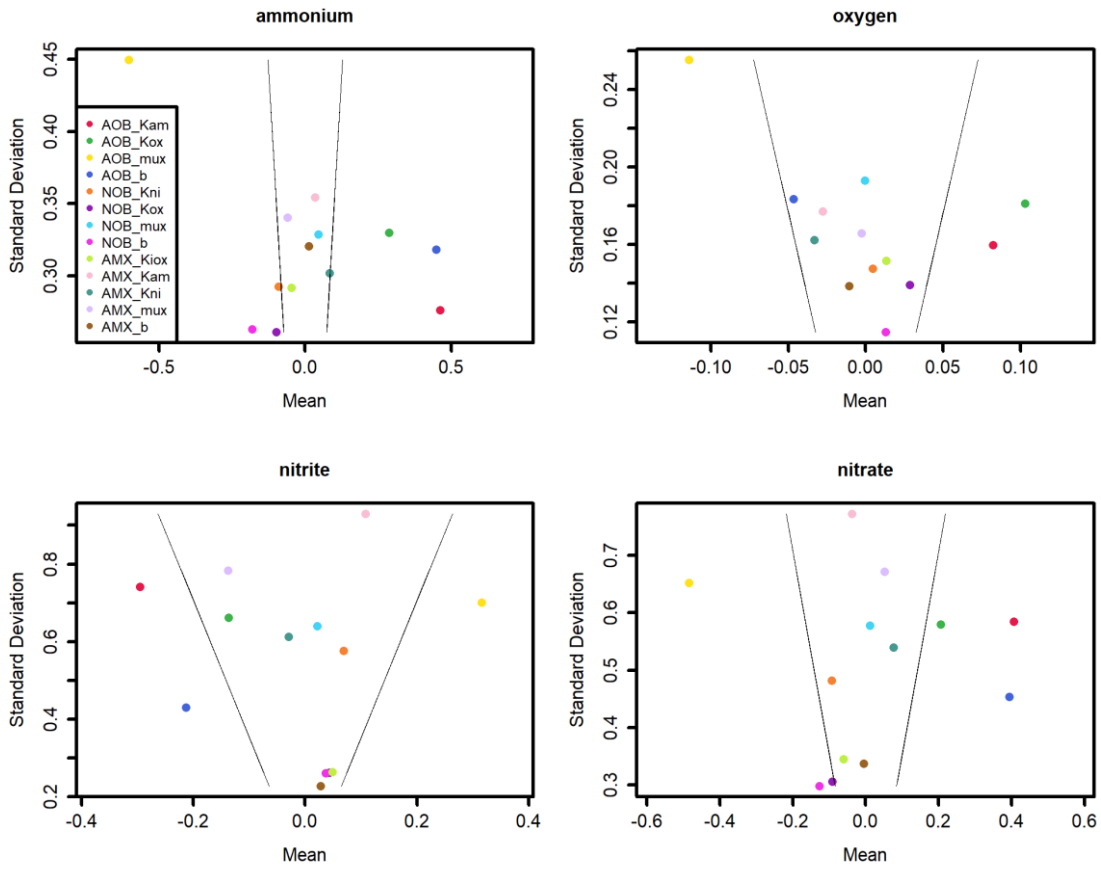


Figure 5.8. EE plots for the final concentration (mg/L) of ammonium (top left), oxygen (top right), nitrite (bottom left) and nitrate (bottom right) in the bulk region. The lines indicate ±2 times the standard error of the mean (SEM).

The biofilm structure parameters included in figure 5.9 also appear mostly sensitive to AOB parameters. Apart from representing a large portion of the biofilm, strongly affecting the average and maximum diffusion depth, AOB is mostly found in the top section of the biofilm. Hence, the surface profile of the

biofilm mostly consists of AOB, explaining the significance of AOB parameters on the biofilm fractal dimension. Further the NOB decay rate also appears to significantly affect the average and maximum diffusion depth. Average and maximum diffusion distance are related model outputs, it should therefore not be surprising the two elementary effect plots show a similar profile of input parameters. Finally, none of the tested parameters appears to have a significant effect on the fraction encapsulated void space.

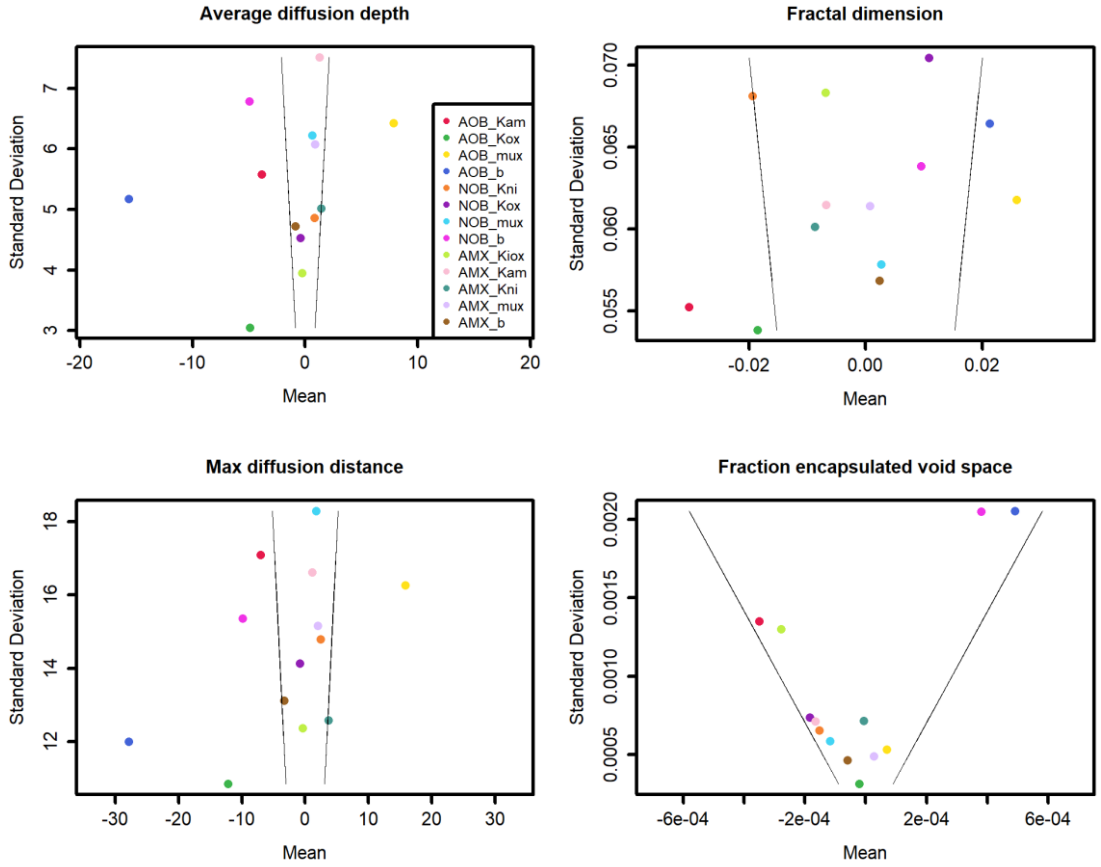


Figure 5.9. EE plots for the average diffusion depth (top left), maximum diffusion depth (top right), fractal dimension (bottom left) and the fraction of encapsulated void space (bottom right). Diffusion depths are in µm, the other output parameters are unitless. The lines indicate ±2 times the standard error of the mean (SEM).

A subset of the quantified localization and co-localization parameters (figure 5.10) again show strong effects from AOB parameters. Since the AMX bacteria are inhibited by oxygen, they are found at the base of the biofilm, where the oxygen concentration is lowest. The EE plot for AMX depth shows a similar profile as that of the average diffusion depth plot, as a thicker biofilm will increase the average diffusion depth and AMX depth in a similar fashion. The

EE plot for NOB depth shows a very different profile. NOB depth appears to be mostly affected by AOB and NOB decay rates.

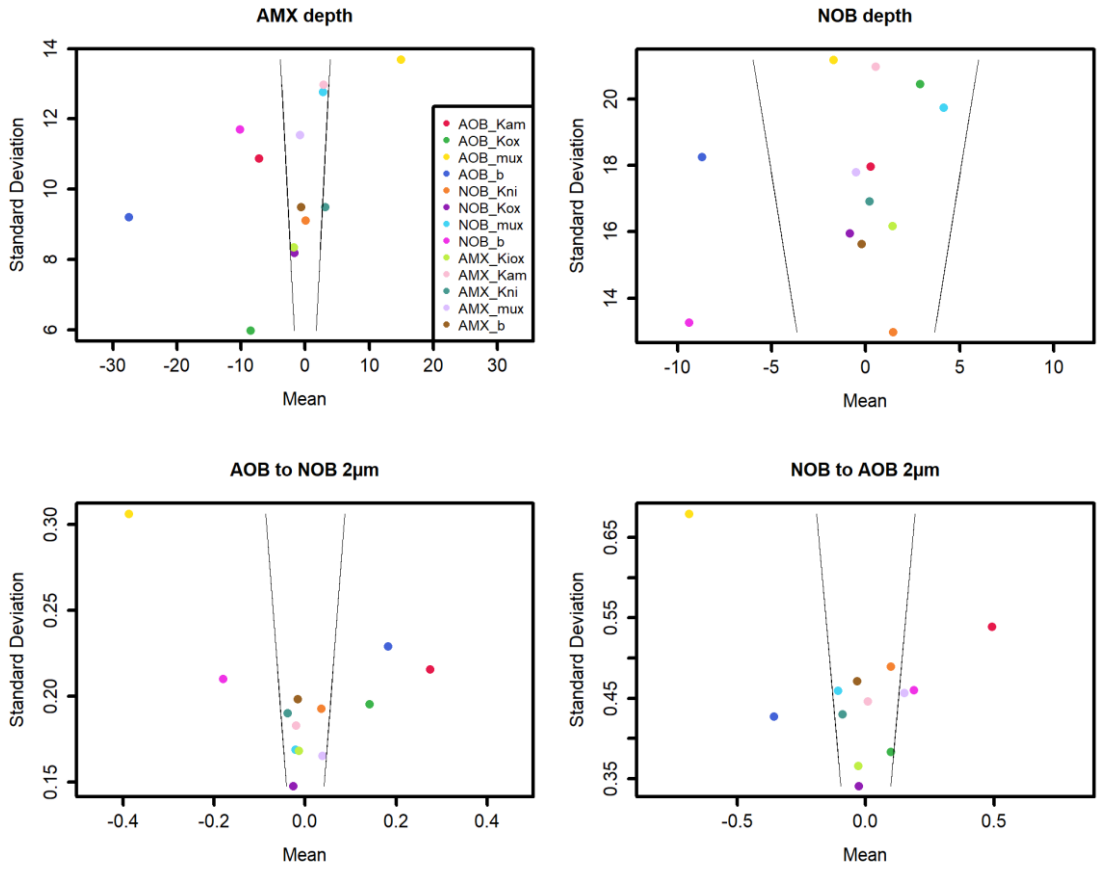


Figure 5.10. EE plots for the average AMX and NOB diffusion depth (top left and right), and the AOB to NOB and NOB to AOB Manders colocalization coefficients (bottom right and left). Diffusion depths are in µm, agents within a 2 voxel range (with a voxel length of 1 µm) are considered co-located. The lines indicate ±2 times the standard error of the mean (SEM).

EE plots for AOB to NOB and NOB to AOB Manders coefficients (2µm) show some overlap between what input parameters are affecting the coefficient. For the AOB and NOB decay rate, the effects are opposite in the two scenarios. A higher NOB decay rate will lead to fewer NOB and potentially fewer NOB microcolonies in the final simulation state, leading to a lower AOB to NOB Manders coefficient. Yet, since there are fewer NOB, there will also be fewer NOB and smaller micro-colonies and the percentage of NOB co-located with AOB will be higher. For the AOB decay rate the effect is inverted. A higher AOB decay rate can lead to the formation of more NOB micro-colonies as there is reduced competition for oxygen and space coming from the AOB.

The Morris screening procedure reveals the dominant position of AOB in the modelled PNA biofilm, and for many state variables, all, or a subset of the AOB parameters are affecting the observed model outcomes. NOB parameters, in particular the NOB decay rate, are sensitive factors for some of the output parameters. AMX parameters appear to have a strong effect mostly on AMX bacteria themselves. Their effects can be seen in the total AMX mass, as well as in co-localization values involving AOB and AMX, small effects on AOB and NOB co-localization, but no strong effects on NOB and AMX co-localization.

5.4 Parameter optimization: Genetic algorithm

For IbMs, parameter optimization poses similar challenges as sensitivity analysis. Due to the typical model complexity, stochasticity, as decision-based agent behavior as well as long model evaluation time, many typical methods are unsuitable. Methods that assume the model to be deterministic can't be used, while methods that rely largely on many sequential evaluations or a very high number of model evaluations may take too long to evaluate. Genetic algorithms (GA) can function well under these circumstances. A GA has been implemented into iDynoMiCS 2.0 (figure 5.11). Each generation of a genetic algorithm can be evaluated in parallel whilst reaching a good estimation with a relatively low number of generations / model evaluations.

The genetic algorithm works by evaluating an initial yet limited sample population from a given parameter space. After a first iteration all sampled parameters are ranked using an objective function. Subsequently a new population is constructed from randomly selecting parameter sets with a high fitness score, these parameter sets undergo an evolution step by randomly interchanging subsets of parameters and by random adjustment of these parameters (mutation). The new generation is then again evaluated and ranked. The process continues until a given fitness threshold is reached or the maximum number of iterations is reached. The entire cycle is implemented in iDynoMiCS and shell script and runs fully automated after initialization.

Ideally, the GA would be utilized to optimize the input parameters based on a rich experimental dataset containing information on the spatial structure of the biofilm, spatial distribution of the microbial species, the microbial community structure as well as (bio-)chemical information such as conversion rates and concentrations. Unfortunately, upon writing this chapter, there is no dataset available that has sufficient detail for this purpose. The parameter optimization could be limited to only optimize for a subset of output parameters discussed

here (such as bio-chemical performance), however this defeats the purpose of looking at the full scope of the biofilm. Therefore, instead a simulated "ground truth" was constructed. The ground truth serves the purpose of replacing data that would have been obtained experimentally otherwise, such that the GA can be tested over the full range of output parameters.

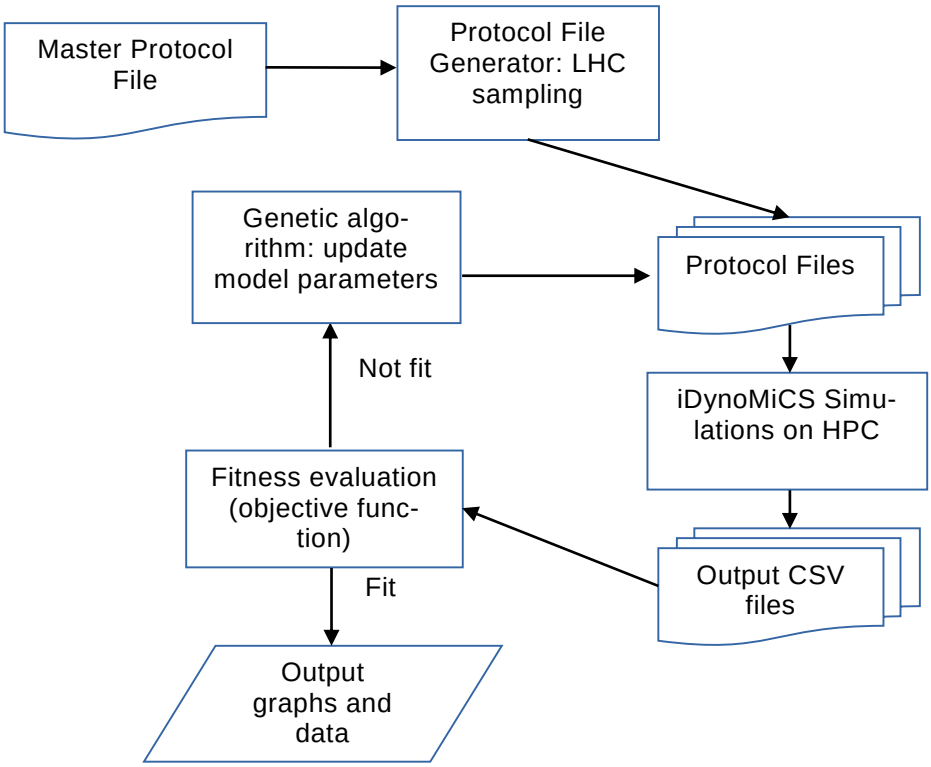


Figure 5.11. The genetic algorithm is an iterative process. A given parameter space is sampled using Latin Hypercube sampling. Child protocols are generated incorporating these parameter sets. After evaluation on a high performance computer the parameter sets are scored. A new generation is created and evaluated until the exit criteria are met. The 'fittest' candidate is the candidate that produces the best representation of the measured data.

The ground truth used here is the same as the simplified PNA biofilm as described in the sensitivity analysis section. Rather than optimizing all model parameters, the input parameter set included in the parameter optimization can be limited to the sensitive parameters as observed in the sensitivity analysis. In this case, the AOB growth and decay parameters are included μ_{AOB}^{max}, $K_{O_2}^{AOB}$, $K_{NH_4}^{AOB}$ and b_{AOB}. The model further showed sensitive to the NOB decay rate b_{NOB}. All other parameters are kept at their default values. There is also a large quantity of outputs the model could be optimized for. There are over 40 different quantified state variables per simulation presented in the first half of this chapter and this number could be extended further, for example by including

chemical gradients and temporal simulation data. Some of these model outcomes may have strong correlations (such as average and maximum diffusion depth, or 1, 2 and 4µm Manders coefficients), while others may not be sensitive to the observed input parameters (such as the fraction of encapsulated void space). Finally, there may be state variables that are generally considered to be of less importance. A subset of all available outputs was selected to be used in the objective function of the GA (table 5.1).

The objective function is used by the GA to rank parameter sets. A model outcome that more closely resembles the target outcome (the observations) will be ranked above outcomes that are less similar. The objective function can be quite simple, such as the mean squared error (MSE). In this case, the MSE is less suitable as the state variables are not directly comparable with each other. It is further important to be aware of the random processes in the model system. Observations such as the number of microcolonies in a section of biofilm, and consequently co-localization coefficients, can have a relatively high variance compared to other observations. Some scenarios may lead to multiple potential outcomes, for example one of the species may die-off or wash out early in the experiment, leading to a distinctly different biofilm after 100 days. By repeatedly simulating the ground truth with different random seeds (thus leading to different initial configurations and random number sequences in the simulations) insight in the stability and variability of the observed biofilm properties is obtained. With sufficient repetitions, scenarios with multiple stable outcomes reveal themselves and the mean and variance of the model outcomes due to stochasticity in the model system itself can be calculated (table 5.1).

By plotting the observations (here state variables of the ground truth instead of experimental observations) for all repetitions, insight is gained on their distribution. Here, several observations appear to be skewed in one or both tails, although the $\pm\sigma$ intervals do appear to follow normal distributions (figure 9.6). Common transformation methods did not resolve the skewness for the observations. However, for the purpose of the objective function this is acceptable. A one-sample Z-test is used for each of the observations to rank each individual simulation in the GA. Due to the skewness of some of the observations, this ranking is not optimal far away from the sample mean μ_0, but it should work well in the $\pm\sigma$ interval of the ground simulations and thus allow the GA to converge to an optimized input parameter set.

Table 5.1. Selected state variables used to calculate the GA objective function and their theoretical range given the model input parameters.

Parameter	**Mean**	**Standard deviation**	**Theoretical range**
Oxygen concentration	1.311 mg/L	0.0360 mg/L	0 to 8.74 mg/L
Ammonium concentration	1.029 mg/L	0.0944 mg/L	0 to 3.0 mg/L
Nitrite concentration	0.102 mg/L	0.0576 mg/L	0 to 3.0 mg/L
Nitrate concentration	1.528 mg/L	0.0947 mg/L	0 to 3.0 mg/L
Total AOB mass	387.16 pg	7.296 pg	0 to 5 mg
Total NOB mass	68.71 pg	4.134 pg	0 to 5 mg
Total AMX mass	169.72 pg	27.60 pg	0 to 5 mg
Fractal dimension	1.058	0.0164	1 to 2
Average diffusion depth	22.98 µm	0.997 µm	0 to 128 µm
AOB NOB co-localization	0.638	0.148	0 to 1
NOB AOB co-localization	0.208	0.057	0 to 1

Fitness scores can be calculated using a probability density function (PDF) for each of the observed model outputs:

$$f_z(z_i) = exp\left\{-\frac{z_i^2}{2}\right\}, \text{ for all } z \in \mathbb{R} \quad 5.1$$

With:

$$z_i = \frac{Y_j - \mu_0}{\sigma} \quad 5.2$$

Where μ_0 is the sample mean and σ the standard deviation of the observation obtained from the ground truth simulations. The objective function is defined as:

$$\frac{1}{n}\sum_{i=1}^{n} 1 - f_z(z_i) \quad 5.3$$

When all output parameters of a GA simulation approach the sample means the objective function will approach 0.

In some cases, one may want to weigh observed output parameters differently from one another, for example if certain parameters are considered more important, if there is an over representation of certain types of parameters, or if some parameters are strongly correlated. In this case there is a similar amount of solute concentration (4), microbial population (3) and spatial structure (4) parameters included. Parameters with obvious strong correlations, for example the maximum diffusion depth which strongly correlates to average diffusion

depth, are not included. Thus, weight adjustments for the objective function are not considered to be necessary.

To populate an initial generation of 100 different input parameter combinations, Latin Hypercube Sampling (LHC) was used to sample the parameter space. Similar to the sensitivity analysis section, the lower and upper bounds of the parameter space were chosen by taking -50% and +50% relative to the initial parameters. The 2 parameter sets of a single generation that scored the best fitness ratings are promoted to the next generation and are thus maintained until replaced by candidates with a better fitness score; these 'elite' parameter sets do not undergo modification before the next model evaluation. The other parameter sets are generated by cloning the fittest candidate out of a randomly selected subset of 8 parameter sets from the previous generation. After selection, the sets are adjusted by parameter cross-over with another randomly selected fit candidate and random mutation (modification of the parameter). The cross-over probability was set to 0.25 (per parameter) and the mutation probability of 0.6 (per parameter). The mutation scale was determined by random sampling from a normal distribution $\mathcal{N}\left(X_i, \frac{b-a}{10}\right)$ with the mean equal to the previous value X_i and the variance scaled to the interval of the input parameter $[a, b]$ as previously defined for the parameter space.

The fittest parameter set resulted in a fitness score of 0.223, after 9 generation a fitness score of 0.022 was reached (table 5.2). The fittest parameter set obtained in the first generation (using Latin Hypercube Sampling) provides a good starting point. The fittest parameter set shows a large deviation in total NOB mass, but most output parameters are already relatively close. Following generations lead to a further optimization of the parameter set, and the GA managed to converge to a parameter set that resembles the ground truth population very closely (table 5.1 and 5.2), after 9 generations all values are within a single standard deviation of the ground truth population mean values.

The average fitness and the most fit parameter set improved over the generations of the genetic algorithm (figure 5.12). Note that the model includes stochastic elements and thus even for the same parameter set the outcome and fitness score can vary over consecutive runs. This effect resulted in a poorer fitness ranking for the fittest candidate in generation 4 and 8 relative to the fittest candidate in generation 3 and 7. Further adjustment of the mutation and cross-over probability parameters used for the GA could improve the convergence of the algorithm.

Table 5.2. The fitness score, input parameters and state variables of the fittest parameter sets found in the first and last (9) generation. Input parameters not included in the table are kept the same as in the base scenario (table 9.1)

	Gen 1	*Gen 9*			*Gen 1*	*Gen 9*	
Fitness scores				**State variables**			
Fitness	*0.223*	*0.022*		*Oxygen concentration*	*1.293*	*1.309*	mg/L
				Ammonium concentration	*1.034*	*1.020*	mg/L
Input parameters				*Nitrite concentration*	*0.112*	*0.080*	mg/L
AOB_Kam	*1.583*	*1.938*	$mg\ N/L$	*Nitrate concentration*	*1.531*	*1.538*	mg/L
AOB_Kox	*0.133*	*0.274*	$mg\ DO/L$	*Total AOB mass*	*391.83*	*387.30*	pg
AOB_mux	*0.452*	*0.817*	d^{-1}	*Total NOB mass*	*96.97*	*69.36*	pg
AOB_b	*0.044*	*0.044*	d^{-1}	*Total AMX mass*	*156.82*	*164.83*	pg
NOB_b	*0.042*	*0.060*	d^{-1}	*fractal dimension*	*1.043*	*1.053*	
				average diffusion depth	*24.20*	*23.42*	μm
				AOB-NOB co-localization	*0.558*	*0.631*	
				NOB-AOB co-localization	*0.214*	*0.205*	

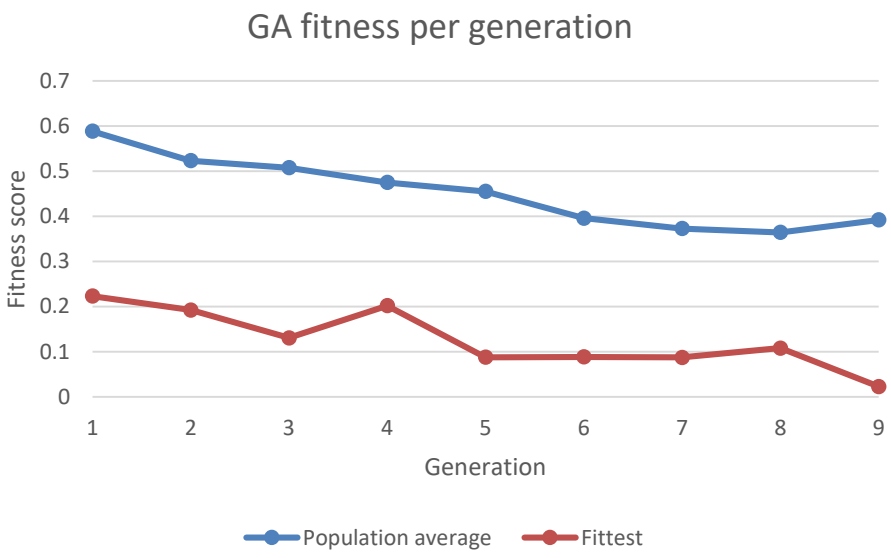


Figure 5.12. The average population fitness and the fitness score of the fittest parameter set in each generation of the genetic algorithm (lower values represent a better fit). The initial population is obtained using Latin Hypercube Sampling, while consecutive populations are obtained from the genetic algorithm.

5.5 Concluding

The procedure described in this chapter shows how methods developed for the analysis of microscopy images can be used to quantify various spatial phenomena observed in IbM outcomes. Quantification of these spatial as well as other model outcomes also allows for sensitivity analysis providing insight into what input parameters affect the model outcomes and in what way (both scale and direction of the effect). The sensitivity analysis can further be used to limit the amount of input parameters considered in a following parameter optimization step, where only the input parameters that have a clear impact on the state variables of interest are optimized, and thereby reducing the amount of required model evaluations. Parameter optimization is essentially a way to adjust the input parameters to have the model better represent the system that is being modelled. Ideally, this procedure would be followed by an uncertainty analysis step and finally model validation on a second independent dataset.

Both sensitivity analysis by Morris screening and parameter estimation by genetic algorithm proved to be effective strategies and met the expectations to effectively reveal the sensitive parameters and estimate parameters with a feasible number of model evaluations. Morris sampling provides qualitative output model parameters can be ranked based on their sensitivity but can't determine the change in input parameter for a certain change in output. Genetic algorithms can be effectively utilized in parameter estimation problems where prior knowledge on these parameter values is limited. The method can be effectively performed in parallel, yet also requires several subsequent iterations. The genetic algorithm was stopped after 9 generations, but already showed strong improvements after a few iterations. This makes the GA very suitable for computationally intensive models like IbMs. The genetic algorithm may benefit from some tweaking to improve its efficiency. As by design the genetic algorithm provides a lot of information around the estimated optimal parameter set. These 'redundant' runs may prove useful to gain initial insights in the model uncertainty, although for a complete picture of the model uncertainties, a better simulation coverage of the parameter space is likely required.

6 Pattern formation and population development in microbial aggregates

iDynoMiCS 2.0 has already been used for multiple computational studies of microbial aggregates. This chapter provides a brief overview of the most notable studies and summarizes the main findings. Much of the content discussed here can be found in more detail in the included manuscripts *I, II and III*.

6.1 Force mediated spatial sorting in aggregate formation

Before a biofilm is formed, several processes and physical interactions play a key role in the initial aggregation of microbial cells. Surface characteristics of suspended cells can change, leading to small groups of cells clustering together. These cells may further mediate the aggregation processes through nanofibers such as pili and flagella extending from the cell surface, and through the production of extracellular polysaccharides (EPS) and other adhesins. The initial reversible attachment phase shifts to what is known as the irreversible phase. In this phase cell-surface interactions become fortified and gene expression of the aggregated microbes shifts; microbial behavior commonly associated with biofilms emerges.

Although this thesis focuses mostly on biofilms in the second phase, the initial stage of biofilm formation can have important implications on the biofilm that is formed. As shown in the paper “biofilms promote altruism” (Kreft, 2004) and reproduced in *Manuscript I*, the initial organization of microbes can have determining implications on the mature biofilm. The initial reversible attachment is often specifically targeted to mitigate biofilm formation on abiotic surfaces. DLVO (Derjaguin, Landau, Verwey and Overbeek) theory (van Oss, 2006) is commonly used to predict microbial attachment success with these types of surfaces. DLVO theory describes (1) Lifshitz-van der Waals interaction, (2) Polar, electron-donor/acceptor, or Lewis acid-base interactions and (3) Electrostatic interactions.

Surface properties of cells can also lead to cell sorting effects, where cell-cell interaction cause to spontaneous segregation of different cell types, altering the spatial structure of the aggregate. Cell sorting due to difference in adhesive properties was first proposed in the 1960s and is also known as the differential adhesion hypothesis (Steinberg, 1962). The principle has been used to explain spontaneous cell sorting in vertebrate embryos (Steinberg, 1963; Duguay et al.,

2003; Foty and Steinberg, 2013; Krieg et al., 2008). The role of differential adhesion mediated cell sorting in microbial biofilms came into focus of researchers much later (Maier, 2021), and is ignored completely in many existing biofilm models.

To investigate the spontaneous aggregation as well as cell sorting of initially planktonic microbial cells, an individual based model using DLVO interaction forces was constructed. iDynoMiCS 2.0 was further extended with a fluid dynamics process. The process implements an open source D2Q9 Lattice Boltzmann algorithm developed by Falcone and Latt (2006). The fluid simulation both introduces mixing within the computational domain, allowing agents near each other to interact and attach, as well as certain degree of shearing, resulting in detachment of weakly adhered cells. This model is described in detail in *Manuscript II*. A short summary of the main results and findings are included in this section.

16 co-aggregation scenarios were simulated (figure 6.1). The scenarios combine strongly adhering cells, as predicted by the large Gibbs free energy of adherence (ΔG_{adh}) calculated using DLVO theory, with 4 different agent types with decreasing ΔG_{adh}, as well as different flow regimes. Scenario A includes two equally strong adhering agent types. In scenarios B, C and D the adhering strength of the second agent type (blue) is reduced stepwise. The surface characteristics of each agent are included in the supplementary information of *Manuscript II*. The model does not include agent growth or decay and thus the observed patterns are exclusively resulting from the modeled physical interactions.

Extended testing revealed four aggregation patterns: (1) both species auto aggregate, (2) only species with strong interactions auto aggregate but then form a scaffold for weak adherents, (3) only strong adhering species aggregate, (4) no species aggregate (typical of high Reynolds number flow regimes). The scenarios presented in figure 6.1 show a transition from pattern 1 (bottom left) to pattern 3 (top right), pattern 4 only occurs with more shearing or a weaker ΔG_{adh}. The size of irregular clusters can be measured as diffusion distance. This revealed trends and differences of aggregate sizes also visible in figure 6.1. In accordance with the differential adhesion hypothesis, when the difference in interaction energy is large, strong adherents form the center of the aggregate whereas weak adherents form the outer shell. In other words, the aggregate becomes stratified. This effect is most clear in Scenarios Re 125 D and Re 250 D, yet the effects are still significant at lower differential adhesion (figure 6.2).

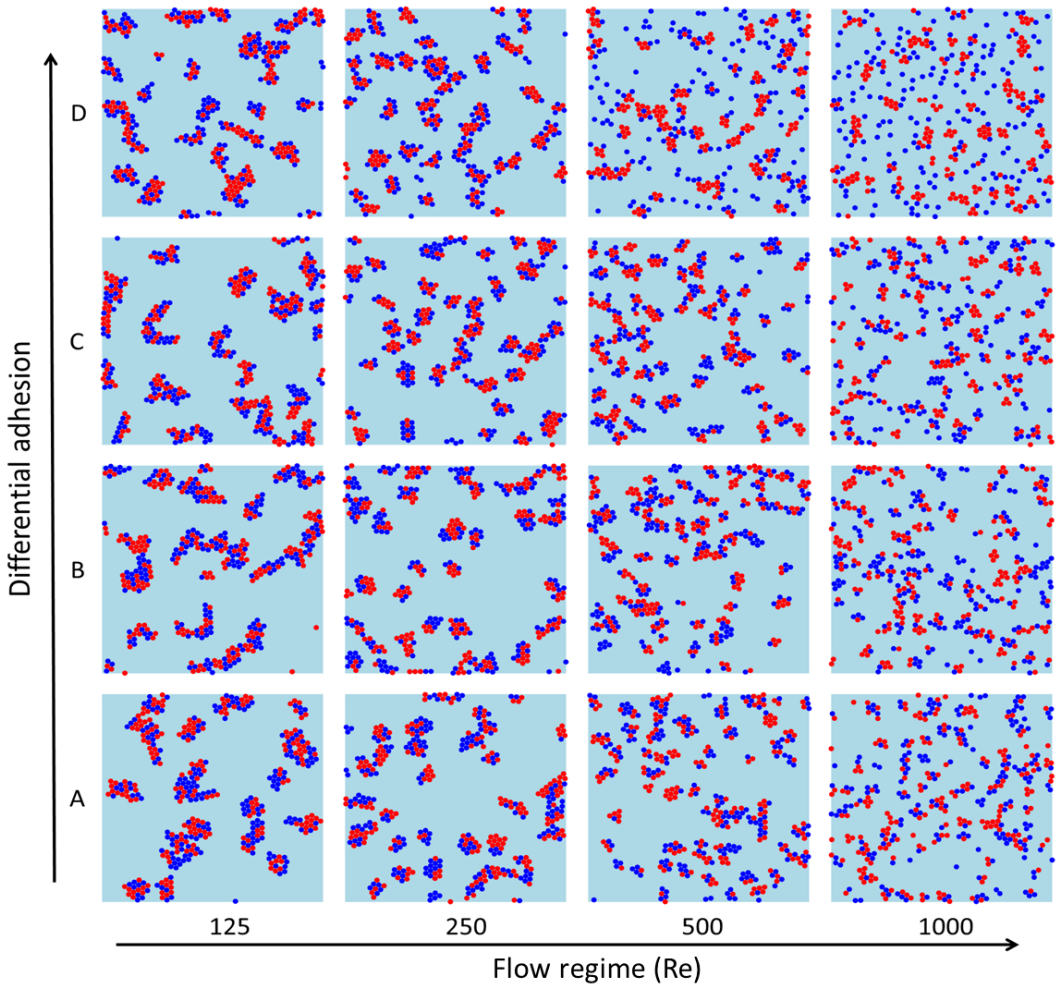


Figure 6.1. Simulation results of 16 aggregation scenarios. Left to right depicts 2 fold increases of turbulence (Reynolds). Bottom to top depicts a stepwise increase of the difference in interaction energies, where red particles remain strong interactors and the blue particles interaction energy decreases. This figure is reproduced from Manuscript II.

The average position within the aggregate relative to the aggregate surface (average diffusion distance from the aggregate surface) of the more adherent cells increases relative to the less adherent cells with increasing difference in adhesiveness (figure 6.2). This trend becomes significant at large enough differences in adhesiveness and holds for all tested flow regimes.

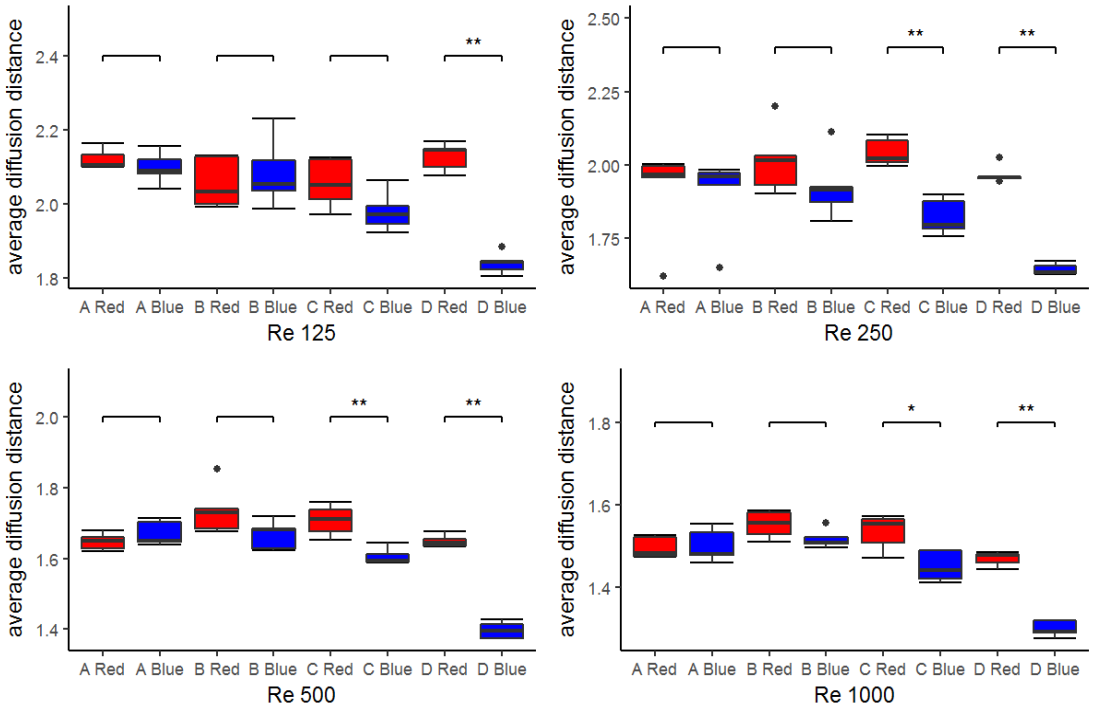


Figure 6.2. Averaged diffusion distances for strong adherent (Red) and weak adherent (Blue). For Re 125, 250, 500 and 1000 scenarios. Scenario A includes two equally strong adhering agent types, while in scenarios B, C and D the adhering strength of the second agent type (blue) is reduced stepwise. Stars indicate a significant difference of the average diffusion depth between the strong and weak adherent in a model scenario. Boxes indicate the interquartile range (IQR) split by a bolt line which represents the median. Bars indicate 1.5 times the IQR over the third quartile and under the second quartile, dots indicate observations outside this range. In the absence of observation outside the 1.5 IQR range, the bar is reduced to indicate the highest or lowest observed value. This figure is reproduced from Manuscript II.

Spatial sorting in microbial aggregates, specifically stratification, is typically explained by the presence of biochemical gradients. This study shows that distinct spatial pattern formation can occur in nascent aggregates due to an interplay between adhesive forces and shear forces occurring in and on the aggregate in absence of biochemical gradients. In the initial stages of microbial aggregates formation, differential adhesion is expected to play a dominant role as biochemical gradients will be weaker in small aggregates. A microbial strategy in which microbes regulate their adhesive properties to adjust their position within an aggregate is conceivable. For example, anammox bacteria, notorious for their strong adhesive properties (Ali et al., 2018), are reversibly inhibited by oxygen and benefit from a position deeper in an aggregate surrounded by aerobic bacteria serving as an oxygen sink.

6.2 Microbial growth strategies for filamentous microbes

Microbial morphologies are often not considered in biofilm models. Although some recent modelling efforts have included non-spherical agents (Smith et al., 2017; Winkle et al., 2017), the majority of IbMs simplify microbial morphology to spheres, despite the importance of microbial shape and size in relation to many selective forces including; nutrient limitation, predation, dispersal, motility, attachment, etc. (Schulz and Jørgensen, 2001; Angert, 2005; Kevin D Young, 2006). iDynoMiCS 2.0 enables the modelling of additional microbial morphologies including rod shapes and filaments. Filament formation is often associated with microbial resource scavenging. The strategy derives from the ability of filaments to focus the growth of biomass into one direction rather than merely producing offspring adding to an existing heap of cells. This way, longer distances can be covered, and new, nutrient rich territories colonized. This strategy is employed by cord-forming fungi who can quickly grow towards new resource hotspots (Aguilar-Trigueros et al., 2022; Boddy, 1993).

A case study was developed extending prior work by Kreft (2004), by introducing filament forming microbes. The original scenario consisted of two competing growth strategies, a Rate Strategist (RS) that grew faster at every substrate concentration (higher μ_{max}, same specific affinity/initial slope of the Monod kinetics) but had a lower growth yield, with a Yield Strategist (YS) that grew slower but had a higher growth yield and therefore converted the substrate diffusing into the biofilm with higher efficiency into biomass. This more economical use of resources is an example of altruistic behavior as it benefits selfish neighbors more than self. It can be reasoned that growing as a filament, which can be considered a cluster of cells in one dimension, gives YS an additional advantage. Here all combinations were tested in a biofilm setting: coccoid RS vs. coccoid YS, filamentous RS vs coccoid YS, coccoid RS vs. filamentous YS and filamentous RS vs. filamentous YS (figure 6.3). As filaments need a larger domain and freedom to bend and spread in all directions, these simulations required a sizable 3D domain, and the 3rd dimension was extended relative to the original study. A full model description is included with *Manuscript I*.

The simulations demonstrate that filamentous growth can provide a strong competitive advantage under nutrient limiting conditions. RS filaments win the competition against YS filaments in each case where the outcome can be inferred from the final biofilm structures (Fig 6.3) and population dynamics over

time (included in *Manuscript I* figure 7). A striking difference between RS filaments and YS filaments is the more open and less dense 'forest' structure produced by RS. This could be due to the lower substrate consumption rate of the YS which allows their filaments to grow better in deeper regions of the biofilms in comparison with the RS filaments. The RS gain a larger advantage when they happen to grow towards the top. Thus, the stronger competition (or self-inhibition) between RS filaments favors expansion over density, leading to a 'fluffier forest' structure.

The advantage of large clusters of YS, who compete less and therefore grow faster than a cluster of RS (figure 6.3 a, m), turns into a disadvantage as RS reach the top of the biofilm faster where substrate flux is highest. The filamentous RS experience stronger positive feedback than the YS (figure 6.3 d, p). This suggests that filamentous growth is a strategy to escape competition between siblings. Given the huge advantage of filamentous growth found here one may wonder why filamentous bacteria are not more abundant. It is certainly common in fungi and streptomycetes, probably because of improved foraging for scattered patches of resources. Also in stream biofilms, filament or chain formation as employed by *Diatoma* spp. enhances nutrient access (Celler et al., 2014). Gradient microbes such as *Beggiatoa* spp. or the intriguing cable bacteria (Pfeffer et al., 2012) form filaments to access electrons from a reduced sediment and electron acceptors from the oxidized water layer above the sediment. Filamentous bacteria are also found in activated sludge flocs in wastewater treatment, where they have the advantage of growing out of the floc into the nutrient richer bulk liquid but are selected against at the settling stage where only fast sinking sludge flocs are recirculated into the activated sludge stage (Martins et al., 2004; Ofiţeru et al., 2014).

Potential disadvantages of filament formation are not represented in the model. These disadvantages include exposure mechanical or physiological stresses and attack by phages or predators and a higher resistance to either active or passive dispersal (Mitchell, 2002; Kevin D. Young, 2006). Cell size is also an important factor for pathogenesis and some microbes may avoid forming filaments to prevent being killed by the host (Yang et al., 2016).

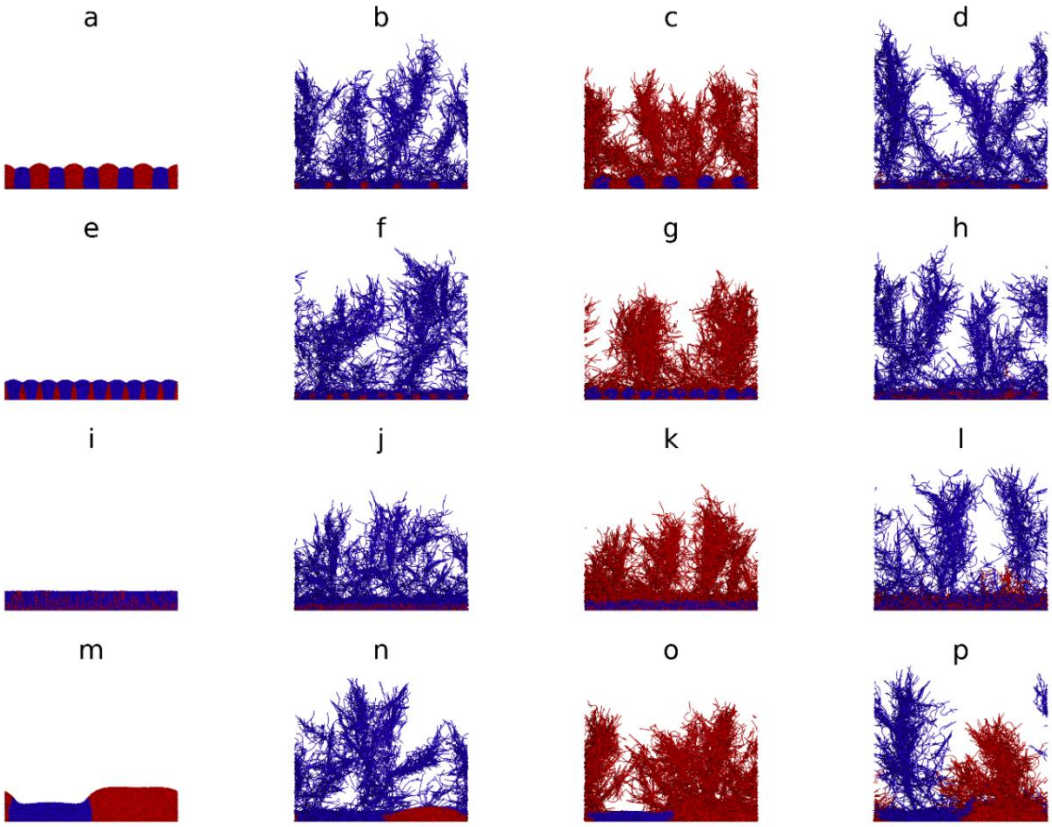


Figure 6.3. Filaments are dominant and give Rate Strategists an advantage. Rate Strategists (RS, blue) and Yield Strategists (YS, red) competed in a 3D biofilm domain (200x200x12.5 µm) for 3 weeks. Simulations in each row were initiated with a different amount of equally spaced agents; 5 each in the first row, 10 each in row 2, 50 each in row 3, row 4 contains 10 each but split between the left and right of the domain. Column 1 corresponds to spherical cell scenarios in Fig 2 of Ref (Kreft, 2004) but were now simulated in 3D. In column 2, RS formed filaments and in column 3, YS formed filaments. In column 4, both formed filaments and RS won or likely won. This figure is reproduced from Manuscript I.

6.3 Individual based gut modeling: fermentation of colonic polysaccharides

Apart from studying relatively abstract aspects of microbial physiology and ecology, IbMs can be used as an effective tool to study more concrete biological systems, including biofilms applied in industrial applications (such as the PNA biofilm presented in chapter 5), but also the biofilms that are part of the human microbiome. The amount of bacterial cell in typical humans is estimated to be over 10^{13} of which the majority is found in the gut (Sender et al., 2016). A growing body of research reveals links between gut microbiota and human health and the list of microbiota associated diseases and disorders includes inflammatory bowel disease, immune disorders, obesity and allergies (Atarashi

et al., 2011; Gophna et al., 2006; Manichanh, 2006; Panzer et al., 2015; Turnbaugh et al., 2006).

A large amount of the gut microbiota research uses animal models such as mice models (Robinson et al., 2019). Animal models allow for a wide range of studies and are relatively easy to work with. One of the drawbacks of most animal models is their difference from human guts in structure, length, cross-sectional area, transit time, and microbial composition. Others have developed *in vitro* gut models using mechanical component to construct an artificial gut system (Vermeiren et al., 2012; Shah et al., 2016). *In silico* models form a third option that, like *in vitro* models, describe a simplified gut system. Examples include gut microbial invasion modeling by Freter et al. (1983) and the colon polysaccharide fermentation model by Muñoz-Tamayo et al. (2010). *In silico* models allow researchers to evaluate a large variety of model scenarios within a small amount of time and few resource requirements.

iDynoMiCS 2.0 was extended to enable individual based gut modeling. This extended software package includes a model of the epithelium, mucosa, and gut lumen. The software, known as eGUT, models epithelial cells explicitly, includes mucus production, models solute transport between the mucosa and blood, and it models the gut lumen as a (well-mixed) bulk compartment. The earlier mentioned gut polysaccharide fermentation model (Muñoz-Tamayo et al., 2010) was extended to include extended to include epithelial dynamics (cellular maintenance, globlet cell mucus production, mucus decay) and implemented in eGUT to test and demonstrate its capabilities (figure 6.4). The model includes four trophic guilds: (1) Glucose-utilizing bacteria, which can hydrolyze polysaccharides into glucose, and ferment glucose, producing lactate, hydrogen, CO_2 and the short-chain fatty acids acetate, propionate and butyrate. (2) Lactate-utilizing bacteria, which ferment lactate, producing acetate, propionate, butyrate, hydrogen and CO_2. (3) Acetogens, which form acetate by reducing CO_2 with hydrogen. (4) Methanogens, which form methane from hydrogen and CO_2. The maintenance of epithelial cells and the metabolic cost of mucus production is modeled through the consumption of butyrate.

The proximal colon lumen receives inflow at a steady rate, containing polysaccharides representing the average-fibre diet proposed in the original model by Muñoz-Tamayo et al. (2010). Digesta (containing agents and solutes) flows from the proximal colon lumen to the transverse colon lumen at an equal rate, and likewise from the transverse colon lumen to the distal colon lumen and out of the distal colon lumen (excretion). Agent and solutes exchange between the

lumen and the corresponding mucosa. A full model description is available in *Manuscript III*.

After five simulated days, the glucose-utilizing bacteria are most abundant in the proximal colon, while the acetogens are most abundant in the transverse and distal colon (figure 6.4). Methanogens are far less abundant than other species. These population differences may be explained by polysaccharide and hydrogen availability in the different compartments.

The proximal colon receives a high influx of polysaccharide, but no hydrogen, while the transverse and distal colon receive hydrogen produced in the prior compartments, but far less polysaccharide. The total biomass in the proximal colon is lower than the total biomass in either of the other colon sections. This is likely due to the products produced by the glucose-utilizing bacteria, including hydrogen, the limiting substrate in acetogenesis, are carried by the flow of digesta from the proximal colon lumen to the other compartments, while the proximal colon receives no supply of hydrogen. Another factor that may drive the differences in species composition between the compartments is the polysaccharide concentration, which differs dramatically between the three compartments.

The case study shows that eGUT can simulate a complex, multi compartmental model of the gut, providing users with comprehensive time courses, solute concentration profiles and detailed image outputs. The patterns that emerge from the model provide opportunities for further research, and data that can be compared to other models and measurements.

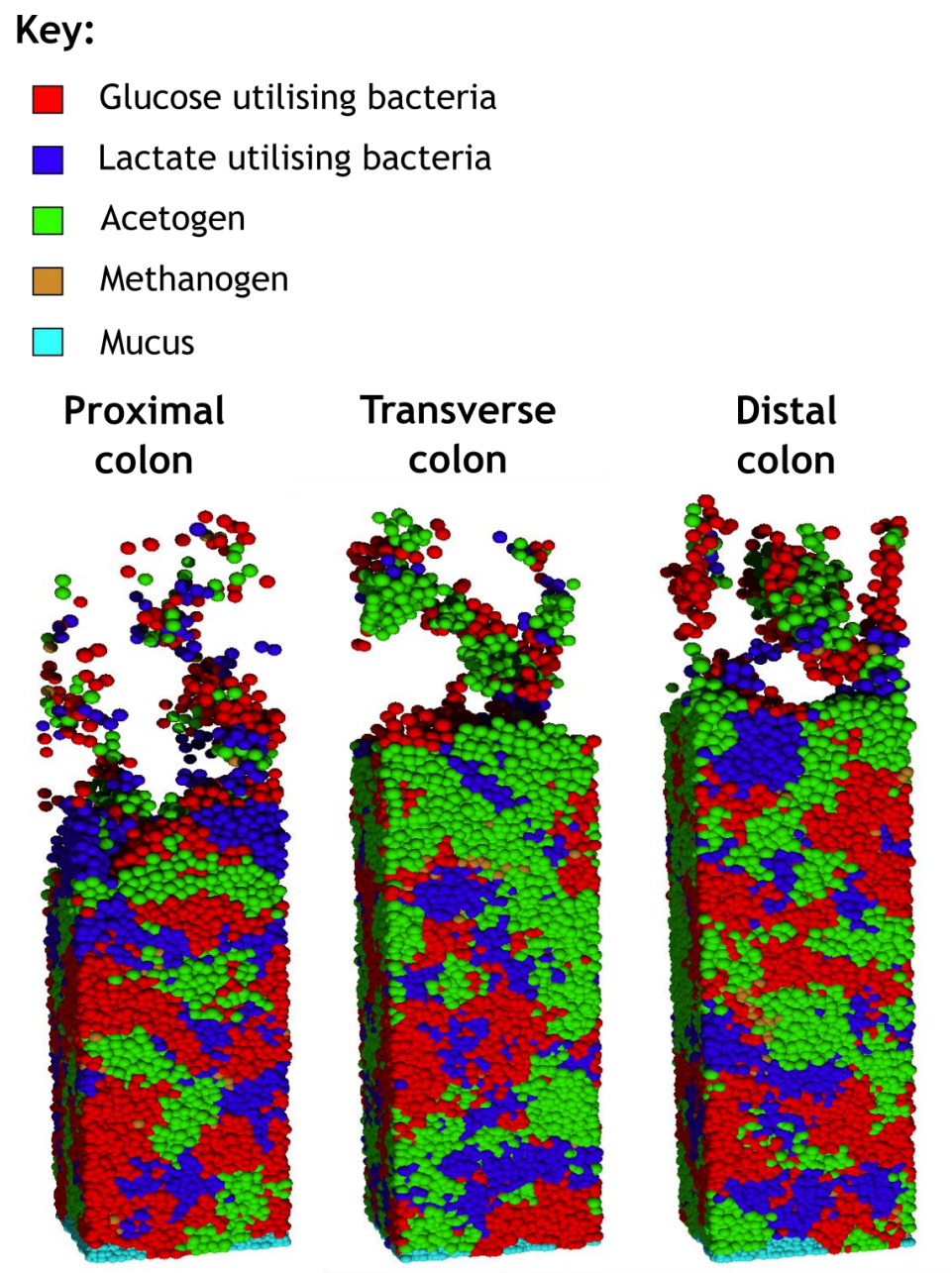


Figure 6.4. Individual based model of the fermentation of colonic polysaccharides after 5 simulated days. Three regions of the colon (proximal, transverse and distal) are modeled using a small representing spatial section of 32x32x256µm representing the mucosa. A lumen compartment is connected to every mucosa section, while also exchanging solutes and agents with the other lumen sections (unidirectional). The lumen, represented by well-mixed compartments, provide the bulk resource of solutes, and allow for the exchange of agents. The base of each mucosa compartment is lined with epithelial (11 colonocytes, 5 globlet cells) agents. Glucose utilizing bacteria are most abundant in the proximal colon, while acetogenic bacteria are most common in the transverse and distal colon. This figure is reproduced from Manuscript III.

7 Conclusions and perspectives

Microbial biofilms and other microbial aggregates play an important role in human and animal health, in various industrial applications, and in the environment. The development of new experimental techniques and increasingly detailed computational models have greatly increased our understanding of microbial aggregates and the dynamic processes occurring within them. New scientific insights and discoveries were made possible by pushing the boundaries of the available research methods.

Computer models that define natural phenomena in mathematical terms and structures, played a key role in obtaining the insight. Computer models help better understand the underlying processes and characteristics which define and influence the microbial aggregate. The processes and phenomena that occur within and around microbial communities that shape the composition, spatial structure and other characteristics of the biofilm are the central interest in this thesis. A better understanding of these processes and phenomena may lead to new control strategies, which in turn may help direct microbial community development in industrial, health and environmental applications. Individual-based microbial community models have helped us understand how individual behavior and local interactions affect the dynamics and spatial structure of microbial aggregates. These models come with several limitations. On one hand, due to the simplified representation of microbes ecologically important properties or behavior may get lost, leading to models that cannot accurately capture essential phenomena. On the other hand, the relative high level of detail and complexity of individual-based models compared to population based modeling approaches can be overwhelming for those not already experienced with IbMs and can lead to a gap between experimental and modeling work. The level of detail can also make the models computationally intense and thereby limit the scale of the model systems. The aim of this thesis has been to improve our capability of accurately modeling these microbial aggregates and the processes and phenomena associated with them, to better predict the dynamic processes and spatial structures found in microbial aggregates and to enable new innovative research that has not been possible before.

Key limitations of existing modeling tools and development needs have been assessed, and the following objectives have been formulated:

1 Enabling a larger range of dynamic individual-based characteristics and behaviors for a better representation of various microbial traits

2 Widening the range of (bio-)chemical sub-models to provide a better representation of microbial metabolisms.
3 Improving the physical representation and interaction models available.
4 Closing the gap between experimental and modeling work facilitating a better integration of experiments and modeling and reducing the required expertise to construct a model.
5 Facilitating model parameterization where individual based experimental observations may be lacking.
6 Improving computational efficiency to allow for models at larger scales.

To achieve this, I have worked with a dedicated group of scientific collaborators to develop the new microbial community modeling framework iDynoMiCS 2.0 and the closely related gut simulation software eGUT. The framework combines many new developments, but also integrates concepts and approaches developed prior.

With iDynoMiCS 2.0 agent properties and behavior can be assembled from orthogonal modules for pick and mix flexibility. Any characteristic of the agent, including the morphology, biochemical and biophysical behavior, can dynamically change in response to external or internal queues. These queues include solute or signal molecule concentrations, the internal availability of storage molecules, or even due to a stochastic process (*Manuscript I*). This gives enormous flexibility in the range of dynamic individual-based characteristics and behaviors that can be modeled with the framework. Possibilities include sporulation, dormancy, signaling, genetic mutation, genetic switching, adaptation, microbial life cycles, ageing, chemotaxis and more.

Prior general purpose individual-based microbial community modeling frameworks are typically restricted to a limited set of kinetic models to formulate the biochemistry of a microbe, while with iDynoMiCS 2.0 the kinetic model can be freely expressed through any kind of arithmetic function (*Manuscript I*). The function may include local solute or signal molecule concentrations as well as properties of the agent, allowing for fully individual behavior.

The first IbMs implement a shoving method in combination with spherical particles to represent agents. The mass-spring approach by Janulevicius et al. (2010) allows for other agent morphologies such as rod-like and filamentous agents, while also greatly improving the physical interaction model of agents. This approach has been refined further and iDynoMiCS 2.0 implementation enables the expression of any type of physical/force-interaction through arithmetic functions. The new implementation is referred to Force-based Mechanics

(FbM). The ability to represent microbes as filament forming agents provides insights in the role of microbial morphology where rate- and yield strategists compete for a limited substrate (*Manuscript I*). While the framework was further used to construct an exDLVO based interaction model. The model implementing CFD based hydrological environment, demonstrates how differential adhesion can lead to different cell sorting patterns under various hydrodynamic conditions (*Manuscript II*).

An IbM framework that enables biologists to formulate their system in their own language rather than computer code is an important factor to bridge the gap between experimentalists and modelers. This concept was taken to hearth during the development of iDynoMiCS 2.0. By implementing numerical solvers that self-inspect convergence and automatically adjust in order produce outcomes within a reasonable error range, by providing default parameters where they are sensible, and by relying on parameters that can be obtained from experimental observations or literature instead of relying on abstract concepts without direct biological parallels. Formulating an individual-based model has been greatly simplified. The gut modelling software eGUT extends iDynoMiCS 2.0 with an epithelium and mucus model and allows for individual-based modeling of the gut lumen and mucosa (*Manuscript III*). The software lowers the barrier for researchers studying gut microbiota to incorporate individual based modeling in their work.

The extensive model analysis, sensitivity analysis and model parameterization tools can further bring experimental and modeling work together. They provide insight into the structural development of microbial aggregates and make it easier to quantitatively compare model and experimental observations. The Morris method implementation makes it easier to link individual traits to emergent properties of the biofilm, while the genetic algorithm is an effective tool for the parameterization of iDynoMiCS 2.0 models. The tools are also an important asset in pattern based modeling and help to identify and replicate the patterns that characterize the microbial aggregate, and thereby mitigating some of the issues caused by the limited availability of individual based observations (*chapter 5, Manuscript III*). In many cases good initial parameter estimates can be obtained with empiric models, such as those for microbial growth characteristics (*chapter 4*).

iDynoMiCS 2.0's orthogonal modules and the possibility to freely formulate mechanical and biochemical models through arithmetic expressions open-up many possibilities for modelers without requiring any modification of source code (*Manuscript I*). The chosen software structure generalizes many standard

procedures through software interfaces, including the management and timing of processes, importing model parameters through protocols and exporting data and figures. This design choice simplifies extending the framework with new functionality for advanced users and iDynoMiCS 2.0 has already received multiple extensions and upgrades. iDynoMiCS 2.0 also comes with a basic graphical user interface allowing any user to run a simulation and analyze the results in just a few mouse clicks, lowering the barrier for new users.

Throughout the development of iDynoMiCS 2.0 the software was subjected to detailed analysis, testing and software profiling to identify computational bottlenecks. The amount of efficiency optimizations are too numerous to summarize in this section, but some of the key improvements are iDynoMiCS 2.0's spatial sorting and searching algorithm, the self-analyzing partial differential equation solver and efficient spatial mapping through Euclidian distance mapping (*Manuscript I*). These optimization efforts made it possible to reach a model scale of 10M+ agents (microbes, inert and EPS) using a single CPU core, an achievement that up until now was only demonstrated with the modeling framework NUFUB using 100 CPU cores.

The rigorous testing process gives me strong confidence about the quality and reliability of the framework. This was done with a focus on numerical solvers, code that inspects the state at each iteration and diagnoses convergence of solvers. The process helped eliminate bugs and software inefficiencies and demonstrated that solutions were numerically correct with deviations of <0.1% from known analytical solutions.

The first individual-based models using iDynoMiCS 2.0 and eGUT already provide interesting insights in societal relevant applications, and they help explain how individual traits can affect the microbial aggregate. The models include the composition and spatial structure development of microbial communities in the human colon and water treatment systems. iDynoMiCS 2.0 also yields interesting insights in more fundamental studies including the effects of microbial morphology on microbial competition and the effects of differential adhesion on the spatial structure of nascent microbial aggregates.

Some of the functionality new to iDynoMiCS is available in other IbM frameworks. For example, agents with rod shaped morphologies are possible through CellModeller (Rudge et al., 2012), BSim 2.0 (Matyjaszkiewicz et al., 2017) and gro (Gutiérrez et al., 2017). Nevertheless, iDynoMiCS 2.0 can provide this functionality at a much larger scale. This makes iDynoMiCS 2.0 more capable of showing the effects of these microscopic differences on the macroscopic

properties of the biofilm. Fluid dynamics is another example of functionality that is available with other frameworks. CellModeller implements a simple implicit advection model which imposes a linear bulk flow in a given direction, while NUFEB (Li et al., 2019) can simulate fluid dynamics explicitly through coupling with the fluid dynamics toolbox OpenFOAM. iDynoMiCS 2.0 also implements fluid dynamics explicitly, but it can do this with agents of various morphologies while also modeling agent interactions based on exDLVO.

The suitability of different individual-based modeling platforms depends on the needs of the researcher. For exploring synthetic bacterial communities where gene regulation and signaling circuits are engineered into cells, CellModeller, gro or BSim 2.0 may be the most suitable platforms. BacArena (Bauer et al., 2017) and ACBM (Karimian and Motamedian, 2020) offer flux-balance models for metabolism, but come with strong limitations on model scales. When it comes to large scale spatially structured biofilm models where mass transport is crucial, including local variability, physical and biochemical interactions between agents and the environment, NUFEB and iDynoMiCS 2.0 are the most suitable modeling platforms. NUFEB outcompetes iDynoMiCS 2.0 if it comes to its maximum model scale (NUFEB can utilize 100+ processors on HPC systems), but iDynoMiCS 2.0 comes with a broader palette of features and functionality and works efficiently on any kind of computer system.

In conclusion, as a central part of this PhD thesis I have worked on the development and have applied the new individual-based modeling framework iDynoMiCS 2.0. The modeling framework combines many new developments within the field, while also innovating and addressing key limitations found in prior work. Dedication and years of hard work have led to a powerful general purpose biofilm modeling framework that is both easy to use and easy to extend. iDynoMiCS 2.0 can be used to predict and analyze the biochemical performance, community composition and spatial structure of biofilms. It can be used in practical applications such as predictive modeling applied for water treatment systems or health applications, as well as for fundamental microbial ecological questions as seen with the fitness shift observed when yield- and rate-strategists form filamentous instead of spherical cells. Extensions of this framework, also covered in this thesis, enable an even broader array of possibilities. This includes eGUT (*Manuscript III)*, which enables individual-based gut modeling, by implementing epithelium and mucus sub-model. It further includes the fluid dynamics extension that enabled the exDLVO based micro-

bial aggregate formation study, revealing differential adhesion mediated spatial pattern formation under various hydrological conditions covered in *Manuscript II*.

7.1 Perspectives

Individual-based biofilm modeling has become increasingly popular in the past decades, possibly aided by the increasing availability of ready to use modeling tools, alongside the excellent ability of IbMs to model and predict emergent properties of the biofilm such as spatial structure and community composition. An increasing number of sub-models and detail is incorporated into IbMs, which allows us to use the methodology in pursuit of an increasing amount of research questions and objectives. One of the difficulties that remains with the approach is the often limited individual-based observations available to parameterize these sub-models. For some parameters, smart methodologies to obtain parameter estimates, such as thermodynamics based approaches to estimate microbial growth and decay kinetics, may help. Observation and quantification of secondary biofilm characteristics, such as used in pattern-oriented modelling (POM) can also play an important role. The modeling tools developed as part of this thesis should make this a lot easier. The next step would be to obtain and utilize experimental observations, rather than relying on synthetic data as done in this thesis, which merely serves as a proof of concept.

As previously asserted (Kreft et al., 2013; Hellweger et al., 2016), a deeper integration of experimental and computational work can be mutually beneficial and plays an important role in deepening our understanding of microbial systems. High quality experimental data is essential for the development of high quality predictive models. In return computational studies can aid experimental work by providing detailed insights in the dynamic development of microbial aggregates or by assisting the optimization of culturing processes and conditions. I believe that the continued development and improvement of advanced modeling tools such as Individual-based models that can adequately capture the biological heterogeneity, local interactions and adaptive behavior occurring within biofilms is an essential part of enabling future biofilm and microbial ecology research. All software developed and covered in this thesis is open source and has been made available online. I hope that these new software tools enable a wider range of researchers to incorporate IbM in their work.

The first iDynoMiCS 2.0 and eGUT models utilize their new functionalities to provide interesting insights in how individual traits affect the formation of spatial patterns and the dynamic nature of the microbial community composition

of microbial aggregates in various applications. Nevertheless, these models only scratch the surface of what is possible. Quite some of the new functionality introduced in this thesis has been tested and included as proof of concept in the software repositories of iDynoMiCS 2.0 and eGUT, but is not used in a fully developed model, or only to a very limited extent. A good example is iDynoMiCS 2.0's dynamic individual-based characteristics and behavior. The functionality was used to give agents a lifecycle in the filamentous rate- and yield strategist model, but the wide range of possibilities including sporulation, dormancy, signaling, genetic mutation, genetic switching, adaptation, ageing, and chemotaxis remains for future model implementations.

The focus of this thesis is on the development of IbM methodologies and procedures, and the application of these methodologies and procedures on theoretical microbial aggregates. I expect this work will aid future research that combines experimental and computational approaches and so assist new insights in microbial aggregates, microbial ecology and so advance our scientific field. I hope this work can facilitate the development of novel treatment strategies for those with gastrointestinal diseases and digestive conditions, and that it may facilitate the development refined bio-based processes such as water treatment, leading to improved water quality and the mitigation of greenhouse gas emissions.

8 References

Ackermann, M., 2013. Microbial individuality in the natural environment. ISME J 7, 465–467. https://doi.org/10.1038/ismej.2012.131

Aguilar-Trigueros, C.A., Boddy, L., Rillig, M.C., Fricker, M.D., 2022. Network traits predict ecological strategies in fungi. ISME COMMUN. 2, 1–11. https://doi.org/10.1038/s43705-021-00085-1

Ali, M., Shaw, D.R., Zhang, L., Haroon, M.F., Narita, Y., Emwas, A.-H., Saikaly, P.E., Okabe, S., 2018. Aggregation ability of three phylogenetically distant anammox bacterial species. Water Research 143, 10–18. https://doi.org/10.1016/j.watres.2018.06.007

Angert, E.R., 2005. Alternatives to binary fission in bacteria. Nat Rev Micro 3, 214–224. https://doi.org/10.1038/nrmicro1096

Atarashi, K., Tanoue, T., Shima, T., Imaoka, A., Kuwahara, T., Momose, Y., Cheng, G., Yamasaki, S., Saito, T., Ohba, Y., Taniguchi, T., Takeda, K., Hori, S., Ivanov, I.I., Umesaki, Y., Itoh, K., Honda, K., 2011. Induction of Colonic Regulatory T Cells by Indigenous *Clostridium* Species. Science 331, 337–341. https://doi.org/10.1126/science.1198469

Bauer, E., Zimmermann, J., Baldini, F., Thiele, I., Kaleta, C., 2017. BacArena: Individual-based metabolic modeling of heterogeneous microbes in complex communities. PLOS Computational Biology 13, e1005544. https://doi.org/10.1371/journal.pcbi.1005544

Becker, P., Hufnagle, W., Peters, G., Herrmann, M., 2001. Detection of Differential Gene Expression in Biofilm-Forming versus Planktonic Populations of Staphylococcus aureus Using Micro-Representational-Difference Analysis. Appl Environ Microbiol 67, 2958–2965. https://doi.org/10.1128/AEM.67.7.2958-2965.2001

Berg, H.C., 1993. Random walks in biology. Princeton University Press, Princeton.

Boddy, L., 1993. Saprotrophic cord-forming fungi: warfare strategies and other ecological aspects. Mycological Research 97, 641–655. https://doi.org/10.1016/S0953-7562(09)80141-X

Bogdanowski, A., Banitz, T., Muhsal, L.K., Kost, C., Frank, K., 2022. McComedy: A user-friendly tool for next-generation individual-based modeling of microbial consumer-resource systems. PLoS Comput Biol 18, e1009777. https://doi.org/10.1371/journal.pcbi.1009777

Breitwieser, L., Hesam, A., De Montigny, J., Vavourakis, V., Iosif, A., Jennings, J., Kaiser, M., Manca, M., Di Meglio, A., Al-Ars, Z., Rademakers, F., Mutlu, O., Bauer, R., 2022. BioDynaMo: a modular platform for high-performance agent-based simulation. Bioinformatics 38, 453–460. https://doi.org/10.1093/bioinformatics/btab649

Campolongo, F., Cariboni, J., Saltelli, A., 2007. An effective screening design for sensitivity analysis of large models. Environmental Modelling & Software 22, 1509–1518. https://doi.org/10.1016/j.envsoft.2006.10.004

Celler, K., Hödl, I., Simone, A., Battin, T.J., Picioreanu, C., 2014. A mass-spring model unveils the morphogenesis of phototrophic *Diatoma* biofilms. Sci Rep 4. https://doi.org/10.1038/srep03649

Clegg, R.J., Dyson, R.J., Kreft, J.-U., 2014. Repair rather than segregation of damage is the optimal unicellular aging strategy. BMC Biol 12, 52. https://doi.org/10.1186/s12915-014-0052-x

Costa, E., Pérez, J., Kreft, J.-U., 2006. Why is metabolic labour divided in nitrification? Trends Microbiol 14, 213–219. https://doi.org/10.1016/j.tim.2006.03.006

Costerton, J.W., Cheng, K.J., Geesey, G.G., Ladd, T.I., Nickel, J.C., Dasgupta, M., Marrie, T.J., 1987. Bacterial biofilms in nature and disease. Annu.Rev.Microbiol. 41, 435–464.

Costerton, J.W., Lewandowski, Z., Caldwell, D.E., Korber, D.R., Lappin-Scott, H.M., 1995. Microbial biofilms. Annu Rev Microbiol 49, 711–745. https://doi.org/10.1146/annurev.mi.49.100195.003431

Daims, H., Lebedeva, E.V., Pjevac, P., Han, P., Herbold, C., Albertsen, M., Jehmlich, N., Palatinszky, M., Vierheilig, J., Bulaev, A., Kirkegaard, R.H., von Bergen, M., Rattei, T., Bendinger, B., Nielsen, P.H., Wagner, M., 2015. Complete nitrification by *Nitrospira* bacteria. Nature 528, 504–509. https://doi.org/10.1038/nature16461

Davey, M.E., O'Toole, G.A., 2000. Microbial biofilms: from ecology to molecular genetics. Microbiol.Mol.Biol.Rev. 64, 847–867.

Dockery, J., Klapper, I., 2001. Finger formation in biofilm layers. SIAM J.Appl.Math. 62, 853–869.

Duguay, D., Foty, R.A., Steinberg, M.S., 2003. Cadherin-mediated cell adhesion and tissue segregation: qualitative and quantitative determinants. Developmental Biology 253, 309–323. https://doi.org/10.1016/S0012-1606(02)00016-7

Dyke, P., 1998. Advanced calculus. Macmillan Press, Ltd., London.

Eberl, H.J., Parker, D.F., van Loosdrecht, M.C.M., 2001. A new deterministic spatio-temporal continuum model for biofilm development. J Theor Med 3, 161–175. https://doi.org/10.1080/10273660108833072

Falcone, J.-L., Latt, J., 2006. Lattice Boltzmann sample, written in Java.

Fenchel, T., Thar, R., 2004. "Candidatus Ovobacter propellens": a large conspicuous prokaryote with an unusual motility behaviour. FEMS Microbiology Ecology 48, 231–238. https://doi.org/10.1016/j.femsec.2004.01.013

Flemming, H.-C., Wuertz, S., 2019. Bacteria and archaea on Earth and their abundance in biofilms. Nature Reviews Microbiology 17, 247–260. https://doi.org/10.1038/s41579-019-0158-9

Foty, R.A., Steinberg, M.S., 2013. Differential adhesion in model systems: Differential adhesion in model systems. WIREs Dev Biol 2, 631–645. https://doi.org/10.1002/wdev.104

Freter, R., Brickner, H., Botney, M., Cleven, D., Aranki, A., 1983. Mechanisms That Control Bacterial Populations in Continuous-Flow Culture Models of Mouse Large Intestinal Flora. Infect Immun 39, 676–685. https://doi.org/10.1128/iai.39.2.676-685.1983

Gogulancea, V., González-Cabaleiro, R., Li, B., Taniguchi, D., Jayathilake, P.G., Chen, J., Wilkinson, D., Swailes, D., McGough, A.S., Zuliani, P., Ofiteru, I.D., Curtis, T.P., 2019. Individual Based Model Links Thermodynamics, Chemical Speciation and Environmental Conditions to Microbial Growth. Front. Microbiol. 10. https://doi.org/10.3389/fmicb.2019.01871

Goñi-Moreno, A., Amos, M., 2015. DiSCUS: A Simulation Platform for Conjugation Computing, in: Calude, C.S., Dinneen, M.J. (Eds.), Unconventional Computation and Natural Computation, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 181–191. https://doi.org/10.1007/978-3-319-21819-9_13

Gonzalez-Gil, G., Sougrat, R., Behzad, A.R., Lens, P.N.L., Saikaly, P.E., 2015. Microbial Community Composition and Ultrastructure of Granules from a Full-Scale Anammox Reactor. Microb Ecol 70, 118–131. https://doi.org/10.1007/s00248-014-0546-7

Gophna, U., Sommerfeld, K., Gophna, S., Doolittle, W.F., Veldhuyzen Van Zanten, S.J.O., 2006. Differences between Tissue-Associated Intestinal Microfloras of Patients with Crohn's Disease and Ulcerative Colitis. J Clin Microbiol 44, 4136–4141. https://doi.org/10.1128/JCM.01004-06

Gorochowski, T.E., Hauert, S., Kreft, J.-U., Marucci, L., Shatil, N., Tang, T.-Y.D., Bandiera, L., Bartoli, V., Dixon, D.O.R., Fedorec, A.J.H., Fellermann, H., Fletcher, A., Foster, T., Giuggioli, L., Matyjaszkiewicz, A., McCormick, S., Montes-Olivas, S.L., Naylor, J., Rubio Denniss, A., Ward, D., 2020. Toward engineering biosystems with emergent collective functions. Front. Bioeng. Biotechnol. 8, Article 705. https://doi.org/10.3389/fbioe.2020.00705

Gorochowski, T.E., Matyjaszkiewicz, A., Todd, T., Oak, N., Kowalska, K., Reid, S., Tsaneva-Atanasova, K.T., Savery, N.J., Grierson, C.S., di Bernardo, M., 2012. BSim: an agent-based tool for modeling bacterial populations in systems and synthetic biology. PLoS ONE 7, e42790. https://doi.org/10.1371/journal.pone.0042790

Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W.M., Railsback, S.F., Thulke, H.H., Weiner, J., Wiegand, T., DeAngelis, D.L., 2005. Pattern-oriented modeling of agent-based complex systems: lessons from ecology. Science 310, 987–991.

Gutiérrez, M., Gregorio-Godoy, P., Pérez del Pulgar, G., Muñoz, L.E., Sáez, S., Rodríguez-Patón, A., 2017. A New Improved and Extended Version of the Multicell Bacterial Simulator gro. ACS Synth. Biol. 6, 1496–1508. https://doi.org/10.1021/acssynbio.7b00003

Heijnen, J.J., Kleerebezem, R., 2010. Bioenergetics of Microbial Growth, in: Encyclopedia of Industrial Biotechnology. Wiley, pp. 1–66. https://doi.org/10.1002/9780470054581.eib084

Heijnen, J.J., Van Dijken, J.P., 1992. In search of a thermodynamic description of biomass yields for the chemotrophic growth of microorganisms. Biotech & Bioengineering 39, 833–858. https://doi.org/10.1002/bit.260390806

Heijnen, J.J., Van Loosdrecht, M.C.M., Tijhuis, L., 1992. A black-box mathematical-model to calculate autotrophic and heterotrophic biomass yields based on gibbs energy-dissipation. Biotechnol.Bioeng. 40, 1139–1154.

Hellweger, F.L., Clegg, R.J., Clark, J.R., Plugge, C.M., Kreft, J.-U., 2016. Advancing microbial sciences by individual-based modelling. Nat Rev Micro 14, 461–471. https://doi.org/10.1038/nrmicro.2016.62

Hoogendoorn, C.J., Meer, T.H. van der, 2005. Fysische transportverschijnselen / II / C.J. Hoogendoorn, T.H. van der Meer, 3e dr. ed. VSSD, Delft.

Horn, H., Lackner, S., 2014. Modeling of Biofilm Systems: A Review, in: Muffler, K., Ulber, R. (Eds.), Productive Biofilms, Advances in Biochemical Engineering/Biotechnology. Springer International Publishing, pp. 53–76.

Hubaux, N., Wells, G., Morgenroth, E., 2015. Impact of coexistence of flocs and biofilm on performance of combined nitritation-anammox granular sludge reactors. Water Research 68, 127–139. https://doi.org/10.1016/j.watres.2014.09.036

Jang, S.S., Oishi, K.T., Egbert, R.G., Klavins, E., 2012. Specification and Simulation of Synthetic Multicelled Behaviors. ACS Synth. Biol. 1, 365–374. https://doi.org/10.1021/sb300034m

Janulevicius, A., Van Loosdrecht, M.C., Simone, A., Picioreanu, C., 2010. Cell Flexibility Affects the Alignment of Model Myxobacteria. Biophys.J 99, 3129–3138.

Jayathilake, P.G., Gupta, P., Li, B., Madsen, C., Oyebamiji, O., González-Cabaleiro, R., Rushton, S., Bridgens, B., Swailes, D., Allen, B., McGough, A.S., Zuliani, P., Ofiteru, I.D., Wilkinson, D., Chen, J., Curtis, T., 2017. A mechanistic Individual-based Model of microbial communities. PLOS ONE 12, e0181965. https://doi.org/10.1371/journal.pone.0181965

Kang, S., Kahan, S., McDermott, J., Flann, N., Shmulevich, I., 2014. *Biocellion* : accelerating computer simulation of multicellular biological system models. Bioinformatics 30, 3101–3108. https://doi.org/10.1093/bioinformatics/btu498

Karimian, E., Motamedian, E., 2020. ACBM: An Integrated Agent and Constraint Based Modeling Framework for Simulation of Microbial Communities. Sci Rep 10, 8695. https://doi.org/10.1038/s41598-020-65659-w

Kissel, J.C., McCarty, P.L., Street, R.L., 1984. Numerical Simulation of Mixed-Culture Biofilm. Journal of Environmental Engineering 110, 393–411. https://doi.org/10.1061/(ASCE)0733-9372(1984)110:2(393)

Kits, K.D., Sedlacek, C.J., Lebedeva, E.V., Han, P., Bulaev, A., Pjevac, P., Daebeler, A., Romano, S., Albertsen, M., Stein, L.Y., Daims, H., Wagner, M., 2017. Kinetic analysis of a complete nitrifier reveals an oligotrophic lifestyle. Nature 549, 269–272.

Kleerebezem, R., Van Loosdrecht, M.C.M., 2010. A Generalized Method for Thermodynamic State Analysis of Environmental Systems. Critical Reviews in Environmental Science and Technology 40, 1–54. https://doi.org/10.1080/10643380802000974

Kreft, J.-U., 2004. Biofilms promote altruism. Microbiology 150, 2751–2760. https://doi.org/10.1099/mic.0.26829-0
Kreft, J.-U., Booth, G., Wimpenny, J.W.T., 1998. BacSim, a simulator for individual-based modelling of bacterial colony growth. Microbiology 144, 3275–3287. https://doi.org/10.1099/00221287-144-12-3275
Kreft, J.-U., Picioreanu, C., Wimpenny, J.W.T., van Loosdrecht, M.C.M., 2001. Individual-based modelling of biofilms. Microbiology 147, 2897–2912. https://doi.org/10.1099/00221287-147-11-2897
Kreft, J.-U., Plugge, C.M., Grimm, V., Prats, C., Leveau, J.H.J., Banitz, T., Baines, S., Clark, J., Ros, A., Klapper, I., Topping, C.J., Field, A.J., Schuler, A., Litchman, E., Hellweger, F.L., 2013. Mighty small: Observing and modeling individual microbes becomes big science. Proc. Natl. Acad. Sci. U.S.A. 110, 18027–18028. https://doi.org/10.1073/pnas.1317472110
Krieg, M., Arboleda-Estudillo, Y., Puech, P.-H., Käfer, J., Graner, F., Müller, D.J., Heisenberg, C.-P., 2008. Tensile forces govern germ-layer organization in zebrafish. Nat Cell Biol 10, 429–436. https://doi.org/10.1038/ncb1705
Lardon, L.A., Merkey, B.V., Martins, S., Dötsch, A., Picioreanu, C., Kreft, J.-U., Smets, B.F., 2011. iDynoMiCS: next-generation individual-based modelling of biofilms. Environ Microbiol 13, 2416–2434. https://doi.org/10.1111/j.1462-2920.2011.02414.x
Lee, E., Jalalizadeh, M., Zhang, Q., 2015. Growth kinetic models for microalgae cultivation: A review. Algal Research 12, 497–512. https://doi.org/10.1016/j.algal.2015.10.004
Leemput, I.A. van de, Veraart, A.J., Dakos, V., Klein, J.J.M. de, Strous, M., Scheffer, M., 2011. Predicting microbial nitrogen pathways from basic principles. Environmental Microbiology 13, 1477–1487. https://doi.org/10.1111/j.1462-2920.2011.02450.x
Li, B., Taniguchi, D., Gedara, J.P., Gogulancea, V., Gonzalez-Cabaleiro, R., Chen, J., McGough, A.S., Ofiteru, I.D., Curtis, T.P., Zuliani, P., 2019. NUFEB: A massively parallel simulator for individual-based modelling of microbial communities. PLoS Comput Biol 15, e1007125. https://doi.org/10.1371/journal.pcbi.1007125
Liu, Y., Lin, Y.-M., Yang, S.-F., 2003. A Thermodynamic Interpretation of the Monod Equation. Current Microbiology 46, 233–234. https://doi.org/10.1007/s00284-002-3934-z
Lupp, C., Urbanowski, M., Greenberg, E.P., Ruby, E.G., 2003. The Vibrio fischeri quorum-sensing systems ain and lux sequentially induce luminescence gene expression and are important for persistence in the squid host: Vibrio fischeri ainS. Molecular Microbiology 50, 319–331. https://doi.org/10.1046/j.1365-2958.2003.t01-1-03585.x
Maier, B., 2021. How Physical Interactions Shape Bacterial Biofilms. Annu. Rev. Biophys. 50, 401–417. https://doi.org/10.1146/annurev-biophys-062920-063646
Manders, E.M.M., Verbeek, F.J., Aten, J.A., 1993. Measurement of co-localization of objects in dual-colour confocal images. Journal of Microscopy 169, 375–382. https://doi.org/10.1111/j.1365-2818.1993.tb03313.x
Manichanh, C., 2006. Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. Gut 55, 205–211. https://doi.org/10.1136/gut.2005.073817
Martin, K.J., Picioreanu, C., Nerenberg, R., 2013. Multidimensional modeling of biofilm development and fluid dynamics in a hydrogen-based, membrane biofilm reactor (MBfR). Water Research 47, 4739–4751. https://doi.org/10.1016/j.watres.2013.04.031
Martins, A.M.P., Picioreanu, C., Heijnen, J.J., van Loosdrecht, M.C.M., 2004. Three-dimensional dual-morphotype species Modeling of activated sludge flocs. Environmental Science & Technology 38, 5632–5641. https://doi.org/10.1021/es0496591
Matyjaszkiewicz, A., Fiore, G., Annunziata, F., Grierson, C.S., Savery, N.J., Marucci, L., di Bernardo, M., 2017. BSim 2.0: An Advanced Agent-Based Cell Simulator. ACS Synth. Biol. 6, 1969–1972. https://doi.org/10.1021/acssynbio.7b00121
McCarty, P.L., 1965. Thermodynamics of biological synthesis and growth. Air Water Pollut 9, 621–639.

Michalewicz, Z., Schoenauer, M., 1996. Evolutionary Algorithms for Constrained Parameter Optimization Problems. Evolutionary Computation 4, 1–32. https://doi.org/10.1162/evco.1996.4.1.1
Mitchell, J.G., 2002. The Energetics and Scaling of Search Strategies in Bacteria. The American Naturalist 160, 727–740. https://doi.org/10.1086/343874
Morgenroth, E., Wilderer, P.A., 2000. Influence of detachment mechanisms on competition in biofilms. Water Research 34, 417–426. https://doi.org/10.1016/S0043-1354(99)00157-8
Morris, M.D., 1991. Factorial Sampling Plans for Preliminary Computational Experiments. Technometrics 33, 161–174. https://doi.org/10.1080/00401706.1991.10484804
Muñoz-Tamayo, R., Laroche, B., Walter, É., Doré, J., Leclerc, M., 2010. Mathematical modelling of carbohydrate degradation by human colonic microbiota. Journal of Theoretical Biology 266, 189–201. https://doi.org/10.1016/j.jtbi.2010.05.040
Naylor, J., Fellermann, H., Ding, Y., Mohammed, W.K., Jakubovics, N.S., Mukherjee, J., Biggs, C.A., Wright, P.C., Krasnogor, N., 2017. Simbiotics: A Multiscale Integrative Platform for 3D Modeling of Bacterial Populations. ACS Synth Biol 6, 1194–1210. https://doi.org/10.1021/acssynbio.6b00315
Noguera, D.R., Pizarro, G.E., Regan, J.M., 2004. Modeling Biofilms, in: Microbial Biofilms. John Wiley & Sons, Ltd, pp. 222–249. https://doi.org/10.1128/9781555817718.ch13
Ofiţeru, I.D., Bellucci, M., Picioreanu, C., Lavric, V., Curtis, T.P., 2014. Multi-scale modelling of bioreactor-separator system for wastewater treatment with two-dimensional activated sludge floc dynamics. Water Res 50, 382–395. https://doi.org/10.1016/j.watres.2013.10.053
Palomo, A., Jane Fowler, S., Gülay, A., Rasmussen, S., Sicheritz-Ponten, T., Smets, B.F., 2016. Metagenomic analysis of rapid gravity sand filter microbial communities suggests novel physiology of Nitrospira spp. ISME J 10, 2569–2581. https://doi.org/10.1038/ismej.2016.63
Panzer, J.A., Anand, R., Dalmau, J., Lynch, D.R., 2015. Antibodies to dendritic neuronal surface antigens in opsoclonus myoclonus ataxia syndrome. Journal of Neuroimmunology 286, 86–92. https://doi.org/10.1016/j.jneuroim.2015.07.007
Parsek, M.R., Fuqua, C., 2004. Biofilms 2003: Emerging Themes and Challenges in Studies of Surface-Associated Microbial Life. J Bacteriol 186, 4427–4440. https://doi.org/10.1128/JB.186.14.4427-4440.2004
Peces, M., Dottorini, G., Nierychlo, M., Andersen, K.S., Dueholm, M.K.D., Nielsen, P.H., 2022. Microbial communities across activated sludge plants show recurring species-level seasonal patterns. ISME COMMUN. 2, 18. https://doi.org/10.1038/s43705-022-00098-4
Pfeffer, C., Larsen, S., Song, J., Dong, M., Besenbacher, F., Meyer, R.L., Kjeldsen, K.U., Schreiber, L., Gorby, Y.A., El-Naggar, M.Y., Leung, K.M., Schramm, A., Risgaard-Petersen, N., Nielsen, L.P., 2012. Filamentous bacteria transport electrons over centimetre distances. Nature 491, 218–221. https://doi.org/10.1038/nature11586
Pfeiffer, T., Bonhoeffer, S., 2004. Evolution of cross-feeding in microbial populations. Am.Nat. 163, E126–E135.
Picioreanu, C., Kreft, J.-U., van Loosdrecht, M.C.M., 2004. Particle-based multidimensional multispecies biofilm model. Appl Environ Microbiol 70, 3024–3040. https://doi.org/10.1128/AEM.70.5.3024-3040.2004
Picioreanu, C., Pérez, J., van Loosdrecht, M.C.M., 2016. Impact of cell cluster size on apparent half-saturation coefficients for oxygen in nitrifying sludge and biofilms. Water Research 106, 371–382. https://doi.org/10.1016/j.watres.2016.10.017
Picioreanu, C., van Loosdrecht, M.C.M., Heijnen, J.J., 2000a. A theoretical study on the effect of surface roughness on mass transport and transformation in biofilms. Biotechnol.Bioeng. 68, 355–369.

Picioreanu, C., van Loosdrecht, M.C.M., Heijnen, J.J., 2000b. Effect of diffusive and convective substrate transport on biofilm structure formation: A two-dimensional modeling study. Biotechnol.Bioeng. 69, 504–515.

Picioreanu, C., van Loosdrecht, M.C.M., Heijnen, J.J., 1998a. Mathematical modeling of biofilm structure with a hybrid differential-discrete cellular automaton approach. Biotechnology and Bioengineering 58, 101–116. https://doi.org/10.1002/(SICI)1097-0290(19980405)58:1<101::AID-BIT11>3.0.CO;2-M

Picioreanu, C., van Loosdrecht, M.C.M., Heijnen, J.J., 1998b. A new combined differential-discrete cellular automaton approach for biofilm modeling: application for growth in gel beads. Biotechnol Bioeng 57, 718–731.

Pinto, A.J., Marcus, D.N., Ijaz, U.Z., Bautista-de Lose Santos, Q.M., Dick, G.J., Raskin, L., 2016. Metagenomic Evidence for the Presence of Comammox *Nitrospira* -Like Bacteria in a Drinking Water System. mSphere 1, e00054-15. https://doi.org/10.1128/mSphere.00054-15

Purcell, E.M., 1977. Life at low Reynolds number. American Journal of Physics 45, 3–11.

Reichert, P., 1994. Aquasim: A tool for simulation and data analysis of aquatic systems. Wat.Sci.Technol. 30, 21–30.

Reichert, P., Wanner, O., 1997. Movement of solids in biofilms: significance of liquid phase transport. Water Science and Technology 36, 321–328.

Rhodes, M.J., 2008. Introduction to particle technology, 2nd ed. ed. Wiley, Chichester, England ; Hoboken, NJ.

Rittmann, B.E., Manem, J.A., 1992. Development and experimental evaluation of a steady state, multispecies biofilm model. Biotechnol Bioeng 39, 914–922.

Rittmann, B.E., McCarty, P.L., 2018. Environmental Biotechnology: Principles and Applications. McGraw-Hill Education, New York, N.Y.

Robinson, N.B., Krieger, Katherine, Khan, F.M., Huffman, W., Chang, M., Naik, A., Yongle, R., Hameed, I., Krieger, Karl, Girardi, L.N., Gaudino, M., 2019. The current state of animal models in research: A review. International Journal of Surgery 72, 9–13. https://doi.org/10.1016/j.ijsu.2019.10.015

Rudge, T.J., Steiner, P.J., Phillips, A., Haseloff, J., 2012. Computational modeling of synthetic microbial biofilms. ACS Synth. Biol. 1, 345–352. https://doi.org/10.1021/sb300031n

Ruocco, G., 2018. Introduction to transport phenomena modeling: a multiphysics, general equation-based approach. Springer.

Schindler, J., Rataj, T., 1992. Fractal geometry and growth models of a *Bacillus* colony. Binary 4, 66–72.

Schluter, J., Nadell, C.D., Bassler, B.L., Foster, K.R., 2015. Adhesion as a weapon in microbial competition. ISME J 9, 139–149. https://doi.org/10.1038/ismej.2014.174

Schulz, H.N., Jørgensen, B.B., 2001. Big bacteria. Annual Review of Microbiology 55, 105–137. https://doi.org/10.1146/annurev.micro.55.1.105

Sender, R., Fuchs, S., Milo, R., 2016. Revised Estimates for the Number of Human and Bacteria Cells in the Body. PLoS Biol 14, e1002533. https://doi.org/10.1371/journal.pbio.1002533

Seoane, J., Yankelevich, T., Dechesne, A., Merkey, B., Sternberg, C., Smets, B.F., 2010. An individual-based approach to explain plasmid invasion in bacterial populations. FEMS Microbiol.Ecol.

Shah, P., Fritz, J.V., Glaab, E., Desai, M.S., Greenhalgh, K., Frachet, A., Niegowska, M., Estes, M., Jäger, C., Seguin-Devaux, C., Zenhausern, F., Wilmes, P., 2016. A microfluidics-based in vitro model of the gastrointestinal human–microbe interface. Nat Commun 7, 11535. https://doi.org/10.1038/ncomms11535

Shaw, A., Takacs, I., Pagilla, K., Riffat, R., DeClippeleir, H., Wilson, C., Murthy, S., 2015. Toward Universal Half-Saturation Coefficients: Describing Extant K_S as a Function of Diffusion. water environ res 87, 387–391. https://doi.org/10.2175/106143015X14212658614072

Shaw, A., Takács, I., Pagilla, K.R., Murthy, S., 2013. A new approach to assess the dependency of extant half-saturation coefficients on maximum process rates and estimate intrinsic coefficients. Water Research 47, 5986–5994. https://doi.org/10.1016/j.watres.2013.07.003

Smith, W.P.J., Davit, Y., Osborne, J.M., Kim, W., Foster, K.R., Pitt-Francis, J.M., 2017. Cell morphology drives spatial patterning in microbial communities. Proceedings of the National Academy of Sciences 114, E280–E286. https://doi.org/10.1073/pnas.1613007114

Steinberg, M.S., 1963. Reconstruction of Tissues by Dissociated Cells: Some morphogenetic tissue movements and the sorting out of embryonic cells may have a common explanation. Science 141, 401–408. https://doi.org/10.1126/science.141.3579.401

Steinberg, M.S., 1962. ON THE MECHANISM OF TISSUE RECONSTRUCTION BY DISSOCIATED CELLS, I. POPULATION KINETICS, DIFFERENTIAL ADHESIVENESS, AND THE ABSENCE OF DIRECTED MIGRATION. Proc. Natl. Acad. Sci. U.S.A. 48, 1577–1582. https://doi.org/10.1073/pnas.48.9.1577

Stevens, A.B., Hrenya, C.M., 2005. Comparison of soft-sphere models to measurements of collision properties during normal impacts. Powder Technology 154, 99–109. https://doi.org/10.1016/j.powtec.2005.04.033

Stewart, P.S., William Costerton, J., 2001. Antibiotic resistance of bacteria in biofilms. The Lancet 358, 135–138. https://doi.org/10.1016/S0140-6736(01)05321-1

Storck, T., Picioreanu, C., Virdis, B., Batstone, D.J., 2014. Variable cell morphology approach for Individual-based Modeling of microbial communities. Biophys J 106, 2037–2048. https://doi.org/10.1016/j.bpj.2014.03.015

Sweeney, E.G., Nishida, A., Weston, A., Bañuelos, M.S., Potter, K., Conery, J., Guillemin, K., 2019. Agent-Based Modeling Demonstrates How Local Chemotactic Behavior Can Shape Biofilm Architecture. mSphere 4, e00285-19. https://doi.org/10.1128/mSphere.00285-19

Turnbaugh, P.J., Ley, R.E., Mahowald, M.A., Magrini, V., Mardis, E.R., Gordon, J.I., 2006. An obesity-associated gut microbiome with increased capacity for energy harvest. Nature 444, 1027–1031. https://doi.org/10.1038/nature05414

van Kessel, M.A.H.J., Speth, D.R., Albertsen, M., Nielsen, P.H., Op den Camp, H.J.M., Kartal, B., Jetten, M.S.M., Lücker, S., 2015. Complete nitrification by a single microorganism. Nature 528, 555–559. https://doi.org/10.1038/nature16459

van Loosdrecht, M.C.M., Heijnen, J.J., Eberl, H.J., Kreft, J.U., Picioreanu, C., 2002. Mathematical modelling of biofilm structures. Antonie van Leeuwenhoek 81, 245–256. https://doi.org/10.1023/A:1020527020464

van Oss, C.J., 2006. Interfacial forces in aqueous media, 2. ed. ed. CRC, Taylor & Francis, Boca Raton, Fla.

van Waveren, H., 1999. Good modelling practice handbook. STOWA ; Rijkswaterstaat-RIZA ; SDU, afd. SEO/RIZA [etc. distr.], [Utrecht], Lelystad, Den Haag.

Vázquez-Padín, J., Mosquera-Corral, A., Campos, J.L., Méndez, R., Revsbech, N.P., 2010. Microbial community distribution and activity dynamics of granular biomass in a CANON reactor. Water Research 44, 4359–4370. https://doi.org/10.1016/j.watres.2010.05.041

Vermeiren, J., Den Abbeele, P., Laukens, D., Vigsnaes, L.K., Vos, M., Boon, N., Wiele, T., 2012. Decreased colonization of fecal Clostridium coccoides/Eubacterium rectale species from ulcerative colitis patients in an in vitro dynamic gut model with mucin environment. FEMS Microbiol Ecol 79, 685–696. https://doi.org/10.1111/j.1574-6941.2011.01252.x

Vlaeminck, S.E., Terada, A., Smets, B.F., De Clippeleir, H., Schaubroeck, T., Bolca, S., Demeestere, L., Mast, J., Boon, N., Carballa, M., Verstraete, W., 2010. Aggregate Size and Architecture Determine Microbial Activity Balance for One-Stage Partial Nitritation and Anammox. Appl.Environ.Microbiol. 76, 900–909.

von Stockar, U., 2013. Biothermodynamics : the role of thermodynamics in biochemical engineering. EFPL Press, Lausanne, Switzerland.

Wang, L.-L., Wang, L.-F., Ren, X.-M., Ye, X.-D., Li, W.-W., Yuan, S.-J., Sun, M., Sheng, G.-P., Yu, H.-Q., Wang, X.-K., 2012. pH Dependence of Structure and Surface Properties of Microbial EPS. Environ. Sci. Technol. 46, 737–744. https://doi.org/10.1021/es203540w

Wanner, O., Eberl, H.J., Morgenroth, E., Noguera, D.R., Picioreanu, C., Rittmann, B.E., Van Loosdrecht, M.C.M., 2006. Mathematical modeling of biofilms. IWA Publishing, London.

Wanner, O., Gujer, W., 1986. A multispecies biofilm model. Biotechnology and Bioengineering 28, 314–328. https://doi.org/10.1002/bit.260280304

Wanner, O., Reichert, P., 1996. Mathematical modeling of mixed-culture biofilm. Biotechnol Bioeng 49, 172–184.

Widder, S., Allen, R.J., Pfeiffer, T., Curtis, T.P., Wiuf, C., Sloan, W.T., Cordero, O.X., Brown, S.P., Momeni, B., Shou, W., Kettle, H., Flint, H.J., Haas, A.F., Laroche, B., Kreft, J.-U., Rainey, P.B., Freilich, S., Schuster, S., Milferstedt, K., van der Meer, J.R., Großkopf, T., Huisman, J., Free, A., Picioreanu, C., Quince, C., Klapper, I., Labarthe, S., Smets, B.F., Wang, H., Isaac Newton Institute Fellows, Soyer, O.S., 2016. Challenges in microbial ecology: building predictive understanding of community function and dynamics. ISME J 10, 2557–2568. https://doi.org/10.1038/ismej.2016.45

Wimpenny, J.W.T., Colasanti, R., 1997. A unifying hypothesis for the structure of microbial biofilms based on cellular automaton models. FEMS Microbiol.Ecol. 22, 1–16.

Winkle, J.J., Igoshin, O.A., Bennett, M.R., Josić, K., Ott, W., 2017. Modeling mechanical interactions in growing populations of rod-shaped bacteria. Phys. Biol. 14, 055001. https://doi.org/10.1088/1478-3975/aa7bae

Wright, R.J., Clegg, R.J., Coker, T.L.R., Kreft, J.-U., 2020. Damage Repair versus Aging in an Individual-Based Model of Biofilms. mSystems 5, 27. https://doi.org/10.1128/mSystems.00018-20

Xavier, J.B., Picioreanu, C., van Loosdrecht, M.C.M., 2005. A general description of detachment for multidimensional modelling of biofilms. Biotechnology and Bioengineering 91, 651–669. https://doi.org/10.1002/bit.20544

Yang, D.C., Blair, K.M., Salama, N.R., 2016. Staying in Shape: the Impact of Cell Shape on Bacterial Survival in Diverse Environments. Microbiol Mol Biol Rev 80, 187–203. https://doi.org/10.1128/MMBR.00031-15

Yang, X., Beyenal, H., Harkin, G., Lewandowski, Z., 2000. Quantifying biofilm structure using image analysis. J.Microbiol.Methods 39, 109–119.

Young, Kevin D., 2006. The Selective Value of Bacterial Shape. Microbiol Mol Biol Rev 70, 660–703. https://doi.org/10.1128/MMBR.00001-06

Young, Kevin D, 2006. The selective value of bacterial shape. Microbiol Mol Biol R 70, 660–703. https://doi.org/10.1128/MMBR.00001-06

Zimmermann, M., Escrig, S., Hübschmann, T., Kirf, M.K., Brand, A., Inglis, R.F., Musat, N., Müller, S., Meibom, A., Ackermann, M., Schreiber, F., 2015. Phenotypic heterogeneity in metabolic traits among single cells of a rare bacterial species in its natural environment quantified with a combination of flow cell sorting and NanoSIMS. Front. Microbiol. 06. https://doi.org/10.3389/fmicb.2015.00243

9 Appendix

9.1 Substrate utilization and diffusion in a catalyst layer over an inert surface

For a catalyst utilizing a substrate with concentration S with first-order kinetic the following kinetic expression can be used:

$$\frac{dS}{dt} = -kS \tag{9.1}$$

Here k is the first-order kinetic parameter and t is time. This can be combined with Fick's second law of diffusion resulting in:

$$\frac{dS}{dt} = D\nabla^2 S - kS \tag{9.2}$$

Here D is the solute diffusivity and ∇^2 the Laplacian operator. For an inert particle with a constant solute concentration at the surface a steady state concentration profile can be reached:

$$0 = D\nabla^2 S - kS \tag{9.3}$$

Within a Cartesian coordinate system this can be rewritten as:

$$0 = D\left(\frac{d^2S}{dx^2} + \frac{d^2S}{dy^2} + \frac{d^2S}{dz^2}\right) - kS \tag{9.4}$$

The catalyst is assumed to be perfectly homogeneously distributed over a flat inert surface such that a solute gradient only forms perpendicular to the surface ($\frac{dS}{dy} = 0$, $\frac{dS}{dz} = 0$), and thus the equation can be simplified to:

$$\frac{d^2S}{dx^2} = {}^{k}/_{D}\, S \tag{9.5}$$

The general solution to this problem is:

$$S^*(x) = Ae^{-x\sqrt{{}^{k}/_{D}}} + Be^{x\sqrt{{}^{k}/_{D}}} \tag{9.6}$$

A and B can be found using boundary conditions. At the inert surface ($x = 0$), there is no solute gradient ($\frac{dS}{dx} = 0$). The solute concentration at the top of the

catalyst layer is assumed to be known $S(x_b) = S_b$. Using these boundary conditions, the steady state solute concentration in the catalyst layer can be found:

$$S^*(x) = S_b e^{-(x-x_b)\sqrt{k/D}} \frac{1 + e^{2x\sqrt{k/D}}}{1 + e^{2x_b\sqrt{k/D}}} \quad 9.6$$

9.2 Substrate utilization and diffusion on spherical particle

Like substrate diffusion and consumption of a catalyst on a flat inert surface, for a spherical catalyst particle, or a catalyst layer around a spherical particle the problem can be solved analytically. This section elaborates on an analytical solution by Hoogendoorn and Meer (2005).

For a catalyst utilizing a substrate with concentration S with first-order kinetic the following kinetic expression can be used:

$$\frac{dS}{dt} = -kS \quad 9.7$$

Here k is the first-order kinetic parameter and t is time. This can be combined with Fick's second law of diffusion resulting in:

$$\frac{dS}{dt} = D\nabla^2 S - kS \quad 9.8$$

Here D is the solute diffusivity and ∇^2 the Laplacian operator. For an inert particle with a constant solute concentration at the surface a steady state concentration profile can be reached:

$$0 = D\nabla^2 S - kS \quad 9.9$$

When the Laplacian operator for spherical coordinates is written out (the Laplacian can be found in any standard calculus handbook) this equation becomes:

$$0 = \frac{d^2S}{dr^2} + \frac{1}{r^2}\frac{d^2S}{d\varphi^2} + \frac{1}{r^2 sin^2\varphi}\frac{d^2S}{d\theta^2} + \frac{2}{r}\frac{dS}{dr} + \frac{cos\varphi}{r^2}\frac{dS}{d\varphi} - \frac{k}{D}S \quad 9.10$$

With the center of the inert particle on the origin of the coordinate system at r=0. The assumption is made that all concentric layers surrounding the particle

have a perfectly homogenous composition and structure, and thus a concentration gradient will only exist in the r dimension ($\frac{dS}{d\theta} = 0, \frac{dS}{d\varphi} = 0$). This allows the simplification of previous equation:

$$\frac{d^2S}{dr^2} + \frac{2}{r}\frac{dS}{dr} - \frac{k}{D}S = 0 \qquad 9.11$$

A general solution for this differential exists:

$$S = \frac{A}{r} sin\left(\sqrt{-\frac{k}{D}}r\right) + \frac{B}{r} cos\left(\sqrt{-\frac{k}{D}}r\right) \qquad 9.12$$

The validity of this solution can be checked by finding the first and second derivative (S' and S'') of this equation and checking whether $S'' + \frac{2}{r}S' - \frac{k}{D}S$ equals zero.

Since both k and D are greater than zero, $\sqrt{-1} = i, \sin(ix) = i * \sinh(x), \cos(ix) = \cosh(x)$ the equation can be rewritten as:

$$S = \frac{A^*}{r} sinh\left(\sqrt{\frac{k}{D}}r\right) + \frac{B}{r} cosh\left(\sqrt{\frac{k}{D}}r\right) \qquad 9.13$$

Here $A^* = iA$. A^* and B can be found using boundary conditions. The solute gradient in the center of the catalyst particle (or at the inert particle surface) r_0 is zero ($\frac{dS}{dr} = 0$), this boundary condition is only met with B = 0 and thus:

$$S = \frac{A^*}{r} sinh\left(\sqrt{\frac{k}{D}}r\right) + \frac{0}{r} cosh\left(\sqrt{\frac{k}{D}}r\right) = \frac{A^*}{r} sinh\left(\sqrt{\frac{k}{D}}r\right) \qquad 9.14$$

A^* can be found by saying that on the outermost layer of the catalyst region $r = R$ the concentration will be equal to the 'boundary concentration' $S = S_b$:

$$S_b = \frac{A^*}{R} sinh\left(\sqrt{\frac{k}{D}}R\right) \qquad 9.15$$

Thus:

$$A^* = S_b R \frac{1}{sinh\left(\sqrt{\frac{k}{D}} R\right)} \qquad 9.16$$

This results in the following analytical solution for steady state solute concentration S^* at a distance r from the center of the catalyzing particle, or from the inert particle surface:

$$S^*(r) = S_b \frac{R}{r} \frac{sinh\left(\sqrt{\frac{k}{D}} r\right)}{sinh\left(\sqrt{\frac{k}{D}} R\right)} \qquad 9.17$$

This reaction diffusion model was used in an assignment for the master course '12104 Modelling of Environmental Processes and Technologies', where the students had to use the model to predict solute penetration depth.

9.3 Partial nitritation anammox

Table 9.1. Microbial growth parameters partial nitritation anammox model.

Parameter	Symbol	Unicode	Value	source
AOB1				
Species/name			*Nitrosomonas sp. ML1*	Park2007
Maximum growth rate	μ_{AOB}^{max}	AOB_mux	0.66 d^{-1}	
Oxygen affinity	$K_{O_2}^{AOB}$	AOB_Kox	0.24 $mg\ DO/L$	
Ammonia affinity	$K_{NH_4}^{AOB}$	AOB_Kam	1.62 $mg\ N/L$	
Decay rate	b_{AOB}	AOB_b	0.044 d^{-1}	
Growth yield	Y_{AOB}		0.15 $gCOD/gN$	Wiesmann1994
Nitrogen content	i_{NXB}		0.083 $gN/gCOD$	Hubaux2014
AOB2				
Species/name			*Nitrosomonas sp. NL7*	Park2007
Maximum growth rate	μ_{NOB}^{max}		0.77 d^{-1}	
Oxygen affinity	$K_{O_2}^{AOB}$		1.22 $mg\ DO/L$	
Ammonia affinity	$K_{NH_4}^{NOB}$		0.48 $mg\ N/L$	
Decay rate	b_{AOB}		0.83 d^{-1}	
Growth yield	Y_{AOB}		0.15 $gCOD/gN$	Wiesmann1994
Nitrogen content	i_{NXB}		0.083 $gN/gCOD$	Hubaux2014
NOB1				
Species/name			*Nitrospira sp. ND1*	Ushiki2017
Maximum growth rate	μ_{NOB}^{max}	NOB_mux	1.71 d^{-1}	
Oxygen affinity	$K_{O_2}^{NOB}$	NOB_Kox	0.13 $mg\ DO/L$	
Nitrite affinity	$K_{NO_2}^{NOB}$	NOB_Kni	0.084 $mg\ N/L$	
Growth yield	Y_{NOB}		0.041 $gCOD/gN$	Wiesmann1994
Decay rate	b_{NOB}	NOB_b	0.06 d^{-1}	Koch2000
Nitrogen content	i_{NXB}		0.083 $gN/gCOD$	Hubaux2014
NOB2				
Species/name			*N. japonica*	Ushiki2017
Maximum growth rate	μ_{NOB}^{max}		1.26 d^{-1}	
Oxygen affinity	$K_{O_2}^{NOB}$		0.083 $mg\ DO/L$	
Nitrite affinity	$K_{NO_2}^{NOB}$		0.14 $mg\ N/L$	
Growth yield	Y_{NOB}		0.041 $gCOD/gN$	Wiesmann1994
Decay rate	b_{NOB}		0.06 d^{-1}	Koch2000
Nitrogen content	i_{NXB}		0.083 $gN/gCOD$	Hubaux2014
AMX1				
Species/name			*B. Anammoxidans*	Kartal2012
Maximum growth rate	μ_{AMX}^{max}	AMX_mux	0.065 d^{-1}	
Oxygen inhibition constant	K_{i,O_2}^{AMX}	AMX_Kiox	0.032 $mg\ DO/L$	
Ammonia affinity	$K_{NH_4}^{AMX}$	AMX_Kam	0.07 $mg\ N/L$	
Nitrite affinity	$K_{NO_2}^{AMX}$	AMX_Kni	0.07 $mg\ N/L$	
Growth yield	Y_{AMX}		0.159 $gCOD/gN$	Strous1998
Decay rate	b_{AMX}	AMX_b	0.003 d^{-1}	Hao2002
Nitrogen content	i_{NXB}		0.058 $gN/gCOD$	Hubaux2014
AMX2				
Species/name			*B.Sinica*	Kartal2012
Maximum growth rate	μ_{AMX}^{max}		0.098 d^{-1}	
Oxygen inhibition constant	K_{i,O_2}^{AMX}		2.016 $mg\ DO/L$	
Ammonia affinity	$K_{NH_4}^{AMX}$		0.392 $mg\ N/L$	
Nitrite affinity	$K_{NO_2}^{AMX}$		1.204 $mg\ N/L$	
Growth yield	Y_{AMX}		0.159 $gCOD/gN$	Strous1998
Decay rate	b_{AMX}		0.003 d^{-1}	Hao2002
Nitrogen content	i_{NXB}		0.058 $gN/gCOD$	Hubaux2014

Table 9.2. Petersen (stoichiometric) matrix for microbial growth and decay in the partial nitritation anammox model.

		Growth	AOB	NOB	AMX	Decay	AOB	NOB	AMX
S_{NH_4}	$gN\ m^{-3}$		$-i_{NXB} - \frac{1}{Y_{AOB}}$	$-i_{NXB}$	$-i_{NXB} - \frac{1}{Y_{AMX}}$				
S_{O_2}	$gCOD\ m^{-3}$		$-\frac{3.43 - Y_{AOB}}{Y_{AOB}}$	$-\frac{1.14 - Y_{NOB}}{Y_{NOB}}$					
S_{NO_2}	$gN\ m^{-3}$		$\frac{1}{Y_{AOB}}$	$-\frac{1}{Y_{NOB}}$	$-\frac{1}{Y_{AMX}} - \frac{1}{1.14}$				
S_{NO_3}	$gN\ m^{-3}$			$\frac{1}{Y_{NOB}}$	$\frac{1}{1.14}$				
X_{AOB}	$gCOD\ m^{-3}$		1				-1		
X_{NOB}	$gCOD\ m^{-3}$			1				-1	
X_{AMX}	$gCOD\ m^{-3}$				1				-1
EPS	$gEPS\ m^{-3}$		$\frac{3}{20}$	$\frac{3}{20}$	$\frac{3}{20}$				
Kinetic expression			$\mu_{AOB}^{max} * X_{AOB} * \frac{S_{NH_4}}{K_{NH_4}^{AOB} + S_{NH_4}} * \frac{S_{O_2}}{K_{O_2}^{AOB} + S_{O_2}}$	$\mu_{NOB}^{max} * X_{NOB} * \frac{S_{NO_2}}{K_{NO_2}^{NOB} + S_{NO_2}} * \frac{S_{O_2}}{K_{O_2}^{NOB} + S_{O_2}}$	$\mu_{AMX}^{max} * X_{AMX} * \frac{S_{NO_2}}{K_{NO_2}^{AMX} + S_{NO_2}} * \frac{S_{NH_4}}{K_{O_2}^{AMX} + S_{O_2}} * \frac{S_{O_2}}{K_{i,O_2}^{AMX} + S_{O_2}}$		$b_{AOB} * X_{AOB}$	$b_{NOB} * X_{NOB}$	$b_{AMX} * X_{AMX}$

Table 9.3. Operation conditions and computational domain properties of the partial nitritation anammox model.

	X	**Y**	**DBL**	**Bulk/surface ratio**	**Dilution rate**
Base model size	128 µm	256 µm	20 µm	3.13 m^3/m^2	0.015 h^{-1}
reduced model size	64 µm	128 µm	20 µm	3.13 m^3/m^2	0.015 h^{-1}
	ammonium	**oxygen**	**nitrite**	**nitrate**	
Initial/influent concentration	3 mg/l	8.74 mg/l	0.05 mg/l	0 mg/l	

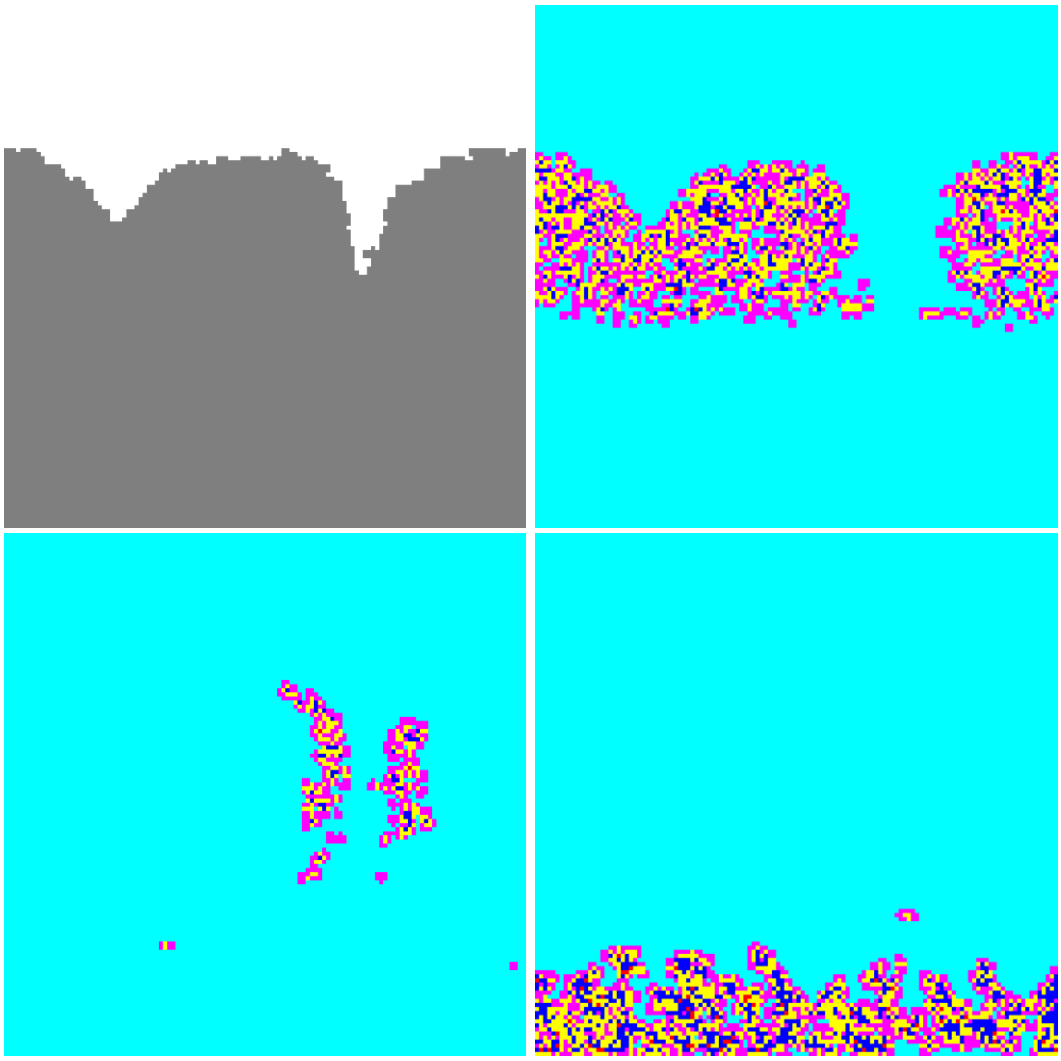


Figure 9.1. Discretization of the computational domain allows for detailed spatial analysis. The computational domain is discretized with a voxel size of 1 µm, a visualization of several regions at the final timestep is shown. The biofilm region is detected (top left), gray indicates the region with biomass. The algorithm can also detect regions with specific traits or species, here the regions with AOB1 (top right), NOB1 (bottom left) and AMX1 (bottom right) are visualized. Colors indicate the number of agents that meet the filter criteria (fx. Species = AOB1) colliding with a voxel: cyan = 0, magenta = 1, yellow = 2, blue = 3, red = 4.

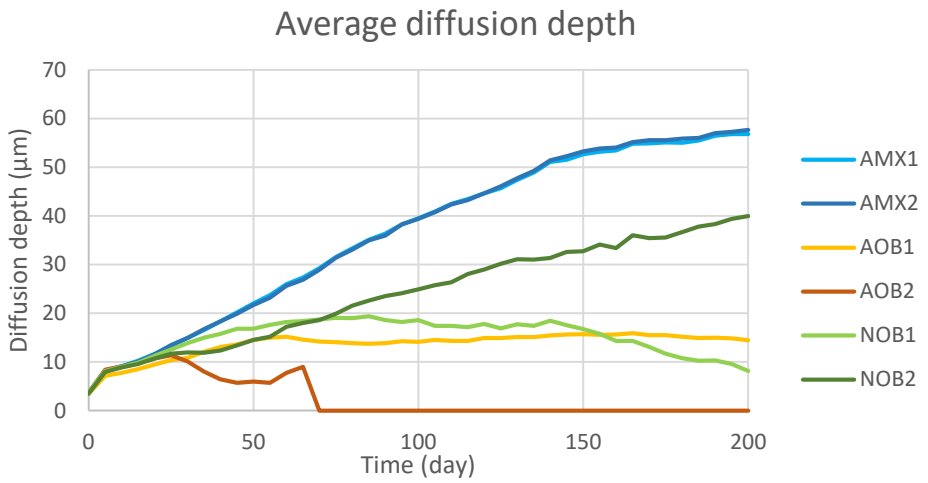


Figure 9.2. average species specific diffusion depth over the course of the simulation. Anammox bacteria are inhibited by oxygen and are thus found deeper in the biofilm, where the oxygen concentration is lower. AOB and NOB1 are located where their substrates are most abundant. After 70 days AOB2 disappeared from, due to its competition for ammonium with AOB1 and oxygen with AOB1 and the NOB. NOB2 is located relatively deeper in the biofilm than NOB1.

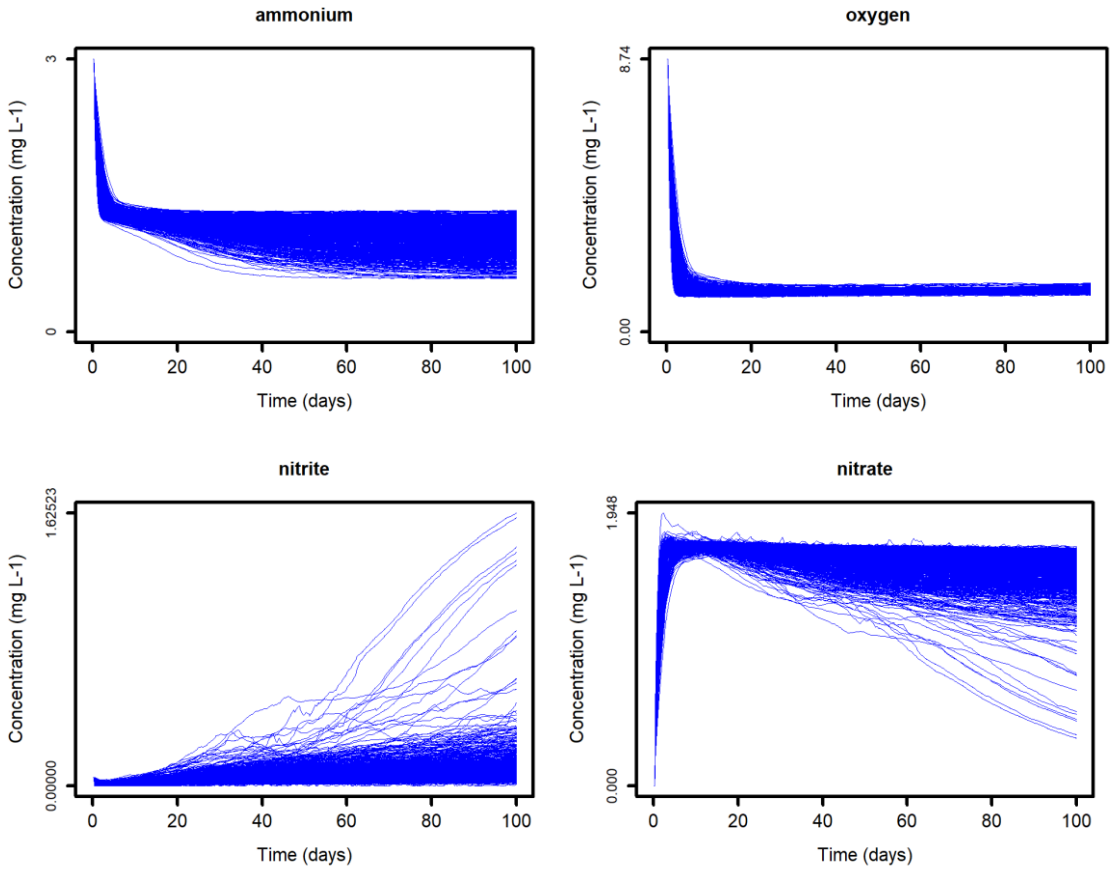


Figure 9.3. Solute concentration over time for all 700 simulations. The panels show ammonium (top left), oxygen (top right), nitrite (bottom left) and nitrate (bottom right) concentrations.

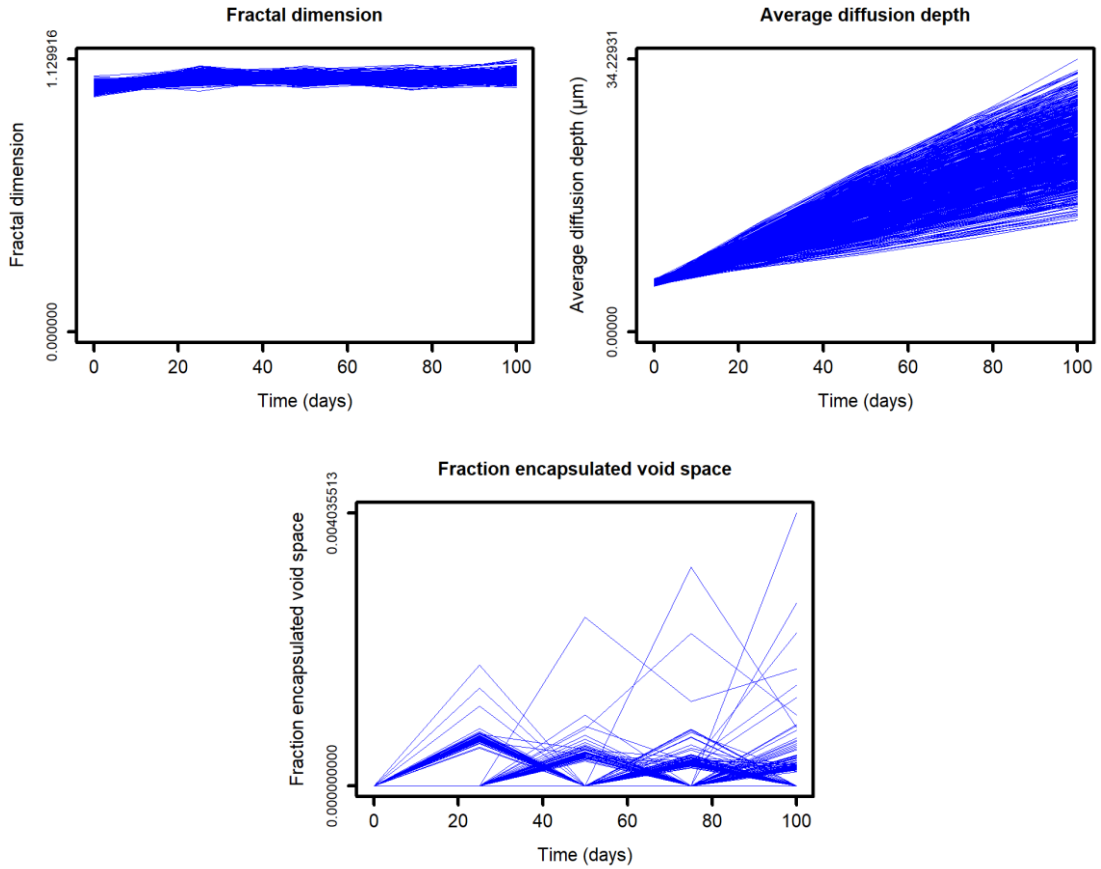


Figure 9.4. A subset of structural parameters over time for all 700 simulations, quantified every 25 days (simulation time). The panels show the fractal dimension (top left), average diffusion depth (top right) and the fraction of encapsulated void space (bottom).

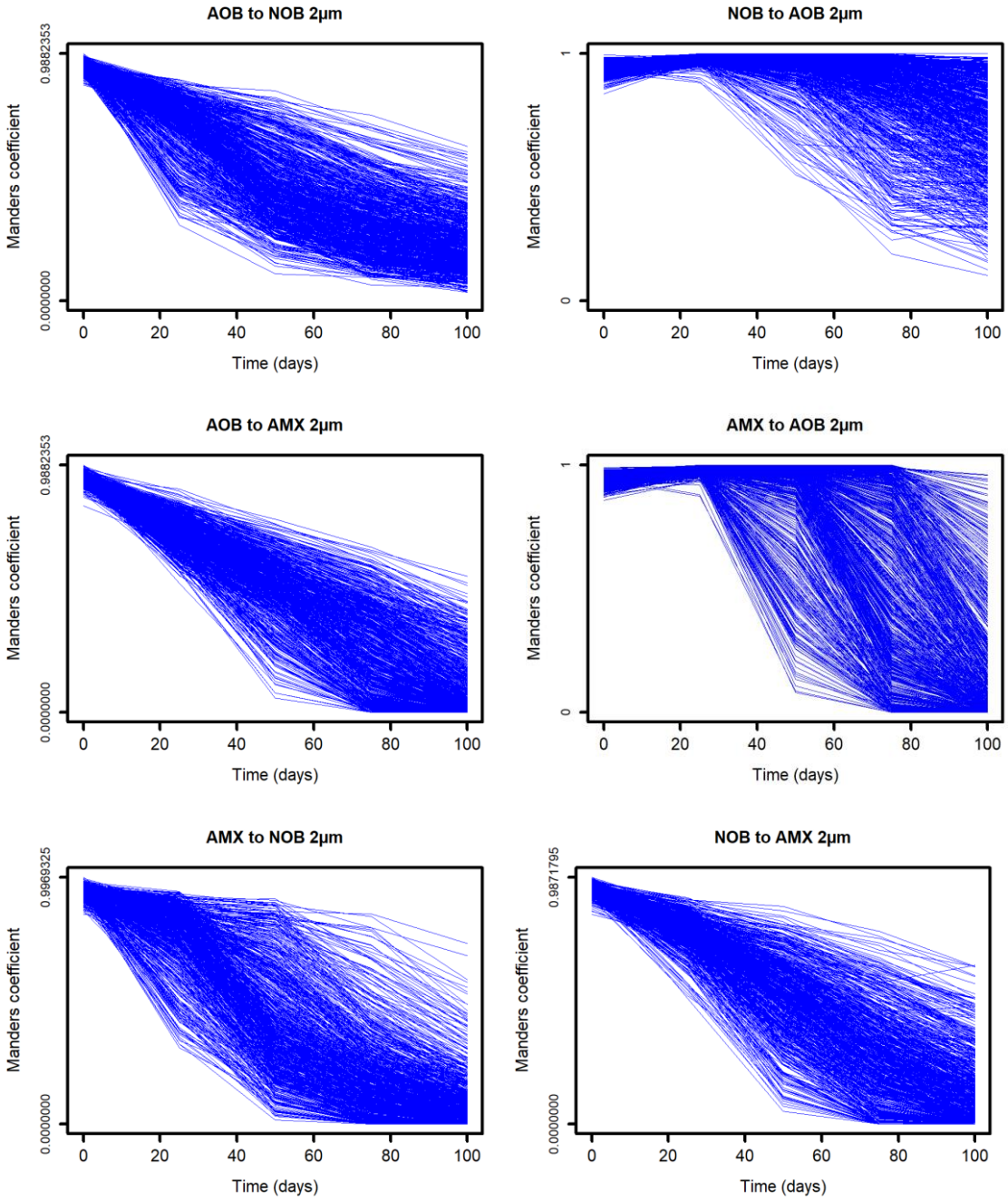


Figure 9.5. A subset of species to species co-localizationl parameters over time for all 700 simulations, quantified every 25 days (simulation time). The panels show AOB to NOB (top left), NOB to AOB (top right) AOB to AMX (mid let), AMX to AOB (mid right), AMX to NOB (bottom left) and NOB to AMX (bottom right).

Table 9.4. The mean of the distribution of the absolute values of the elementary effects (μ*) for the sampled input parameters and observed state variables.

	$K_{NH_4}^{AOB}$	$K_{O_2}^{AOB}$	μ_{AOB}^{max}	b_{AOB}	$K_{NO_2}^{NOB}$	$K_{O_2}^{NOB}$	μ_{NOB}^{max}	b_{NOB}	K_{i,O_2}^{AMX}	$K_{NH_4}^{AMX}$	$K_{NO_2}^{AMX}$	μ_{AMX}^{max}	b_{AMX}
Bulk solute conc.													
oxygen	0.150	0.177	0.230	0.137	0.115	0.102	0.147	0.086	0.098	0.128	0.113	0.132	0.099
ammonium	0.465	0.357	0.636	0.469	0.239	0.190	0.242	0.234	0.172	0.251	0.230	0.236	0.222
nitrite	0.322	0.296	0.377	0.247	0.269	0.179	0.283	0.173	0.136	0.362	0.242	0.321	0.157
nitrate	0.421	0.347	0.554	0.419	0.309	0.226	0.344	0.236	0.186	0.389	0.301	0.336	0.231
Total agent mass													
AOB	30.203	30.221	41.606	656.495	24.227	17.477	26.901	20.203	17.716	32.379	22.526	28.842	18.320
NOB	22.943	19.883	30.070	22.421	13.353	9.432	14.390	109.651	9.434	15.403	11.812	13.923	10.646
AMX	219.921	147.468	308.371	199.035	83.610	72.006	87.038	92.301	76.932	83.383	91.520	87.210	93.000
EPS	53.075	39.336	88.361	11.183	8.265	6.898	9.236	8.142	6.229	6.956	6.630	8.248	7.925
inert	102.145	89.293	156.742	62.823	26.111	23.040	35.212	27.957	22.200	25.349	24.911	27.388	17.265
Biofilm properties													
fractal dimension	0.050	0.044	0.049	0.057	0.054	0.056	0.046	0.050	0.051	0.051	0.048	0.052	0.047
max diffusion distance	14.685	13.726	19.902	27.797	11.389	9.351	11.932	12.886	8.332	11.209	9.590	10.490	9.652
average diffusion depth	5.958	5.274	8.974	15.677	3.461	3.070	3.997	5.887	2.799	4.556	3.451	4.207	3.141
fraction encapsulated void space (x1000)	0.495	0.098	0.154	0.711	0.264	0.219	0.229	0.691	0.454	0.268	0.259	0.229	0.176
Agent localization													
AMX depth	11.093	9.336	17.291	27.525	6.968	6.078	9.368	11.154	6.016	8.879	6.939	8.043	6.581
AOB depth	4.447	3.044	5.376	20.229	3.280	3.491	3.433	3.869	2.364	3.701	3.097	3.005	2.886
NOB depth	11.184	12.075	15.395	13.534	9.309	11.319	13.792	13.314	9.100	12.896	11.439	10.329	10.011
AOB to AMX ADD	14.959	9.228	22.831	8.186	5.676	6.449	8.231	6.768	6.237	7.355	5.392	7.750	5.420
AMX to AOB ADD	12.545	8.226	19.545	11.417	3.001	3.014	4.761	4.597	2.392	3.041	3.703	3.550	2.657
AOB to AMX 1μm	0.108	0.080	0.156	0.082	0.038	0.047	0.050	0.061	0.038	0.032	0.033	0.047	0.039
AMX to AOB 1μm	0.247	0.188	0.323	0.188	0.075	0.088	0.090	0.122	0.077	0.064	0.058	0.097	0.073
AOB to AMX 2μm	0.250	0.165	0.359	0.133	0.067	0.084	0.098	0.105	0.071	0.072	0.064	0.096	0.078
AMX to AOB 2μm	0.450	0.331	0.646	0.351	0.113	0.141	0.166	0.182	0.114	0.125	0.096	0.169	0.124
AOB to AMX 4μm	0.384	0.238	0.609	0.198	0.121	0.141	0.151	0.179	0.126	0.149	0.097	0.157	0.127
AMX to AOB 4μm	0.577	0.418	0.892	0.452	0.139	0.152	0.191	0.233	0.127	0.148	0.116	0.209	0.140
NOB to AMX ADD	17.090	14.714	24.546	18.526	11.560	11.910	17.088	10.632	12.055	14.837	12.931	15.725	11.221
AMX to NOB ADD	12.439	15.289	20.365	17.161	9.634	6.800	11.497	10.670	7.764	10.537	9.020	9.009	7.890
NOB to AMX 1μm	0.154	0.158	0.192	0.139	0.091	0.107	0.088	0.102	0.114	0.120	0.111	0.078	0.121
AMX to NOB 1μm	0.085	0.096	0.117	0.095	0.041	0.044	0.039	0.084	0.058	0.051	0.055	0.035	0.057
NOB to AMX 2μm	0.324	0.323	0.429	0.337	0.189	0.214	0.225	0.194	0.223	0.272	0.230	0.196	0.234
AMX to NOB 2μm	0.202	0.236	0.305	0.249	0.109	0.099	0.109	0.185	0.128	0.118	0.115	0.098	0.132
NOB to AMX 4μm	0.491	0.507	0.686	0.527	0.312	0.358	0.438	0.349	0.339	0.400	0.378	0.350	0.348
AMX to NOB 4μm	0.379	0.444	0.576	0.472	0.207	0.203	0.267	0.309	0.204	0.239	0.235	0.189	0.262
NOB to AOB ADD	2.475	1.423	3.414	1.997	1.454	1.334	1.678	1.845	1.193	1.327	1.293	1.634	1.522
AOB to NOB ADD	8.545	5.101	11.430	8.306	4.075	3.982	4.286	6.019	4.217	5.203	4.760	4.226	4.804
NOB to AOB 1μm	0.364	0.214	0.496	0.329	0.284	0.180	0.226	0.213	0.220	0.230	0.214	0.278	0.258
AOB to NOB 1μm	0.136	0.092	0.180	0.099	0.088	0.067	0.073	0.127	0.074	0.072	0.081	0.075	0.085
NOB to AOB 2μm	0.559	0.310	0.782	0.434	0.379	0.282	0.348	0.374	0.258	0.324	0.313	0.365	0.336
AOB to NOB 2μm	0.289	0.189	0.404	0.232	0.156	0.120	0.127	0.223	0.133	0.139	0.144	0.130	0.148
NOB to AOB 4μm	0.378	0.206	0.492	0.278	0.214	0.192	0.259	0.286	0.182	0.199	0.197	0.244	0.225
AOB to NOB 4μm	0.454	0.302	0.651	0.423	0.228	0.195	0.220	0.324	0.231	0.233	0.211	0.187	0.239

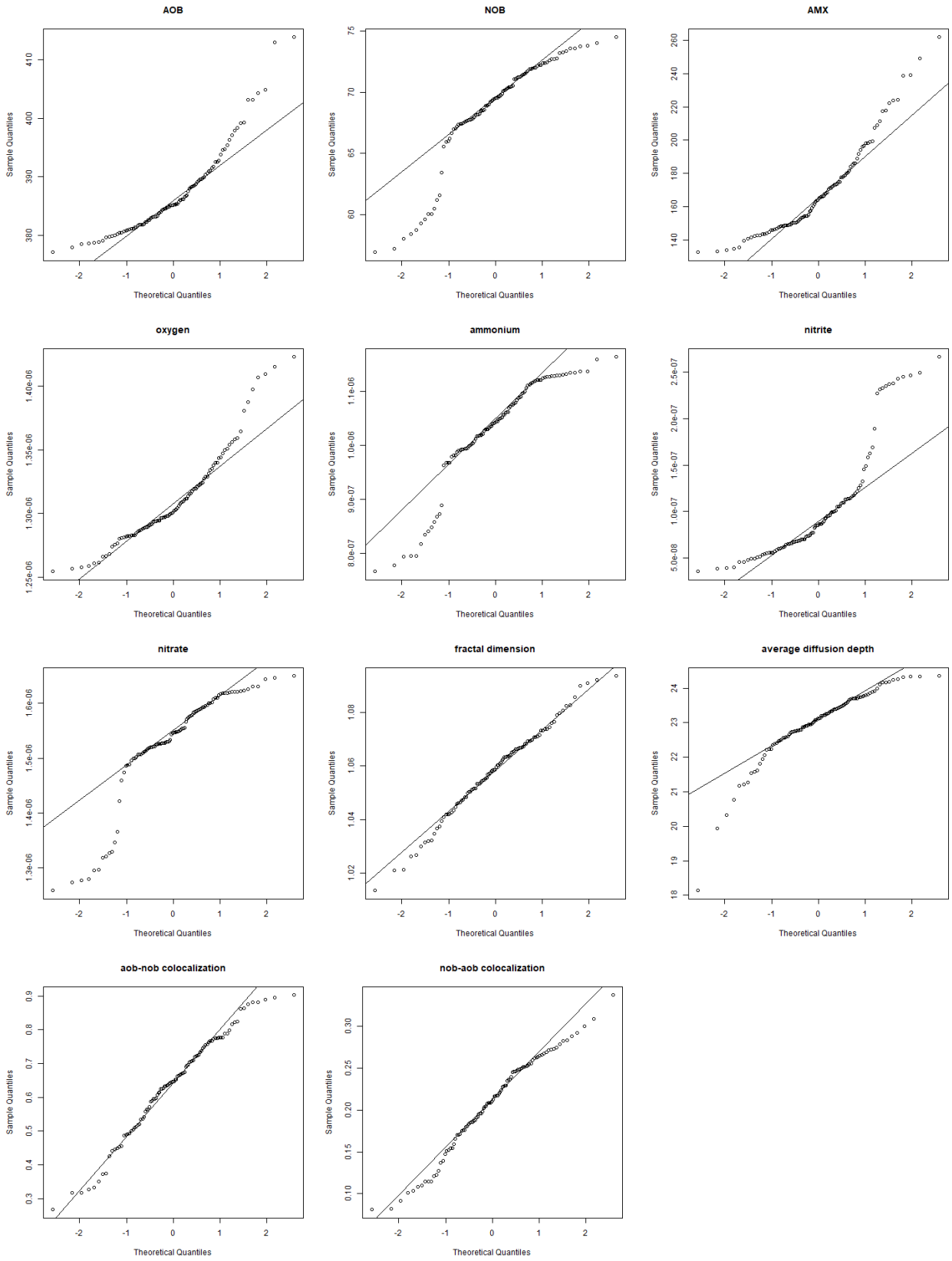
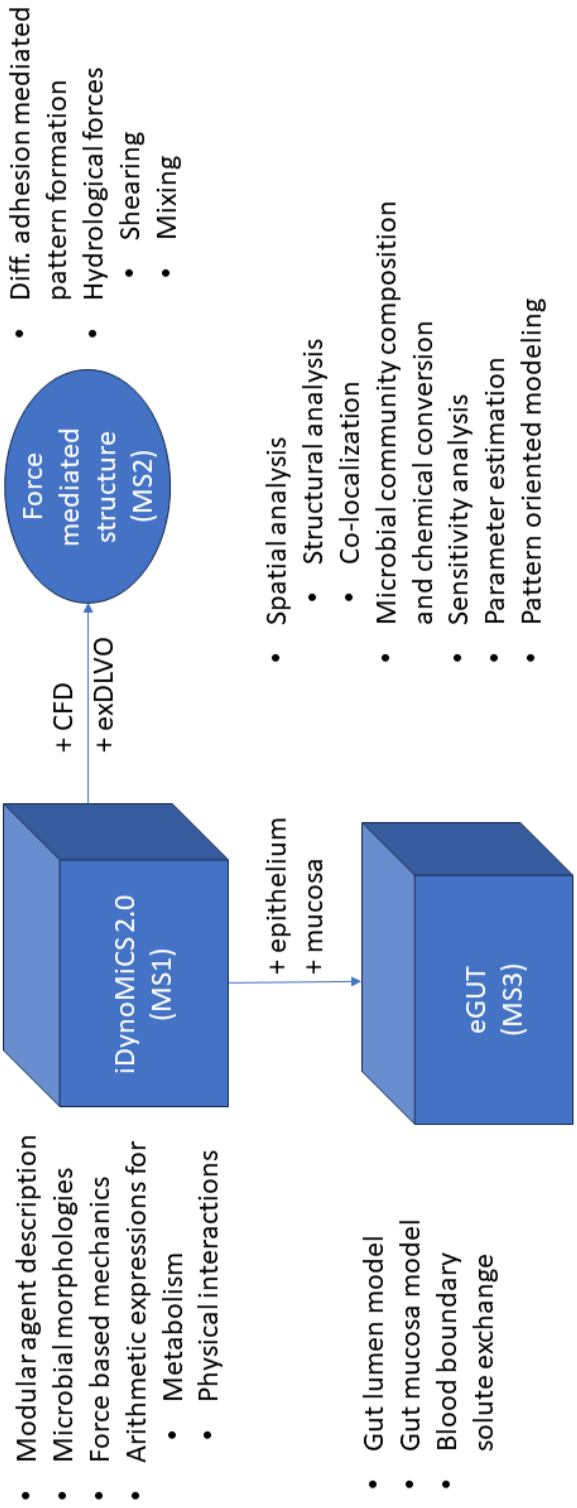


Figure 9.6. QQ-plots reveal a normal distribution of the observed output parameters.

9.4 Graphical thesis overview



9.5 Publications

Bastiaan J R Cockx, Tim Foster, Robert J Clegg, Kieran Alden, Sankalp Arya, Dov J Stekel, Barth F Smets, Jan-Ulrich Kreft, 2023. *Is it selfish to be filamentous in biofilms? Individual-based modeling links microbial growth strategies with morphology using the new and modular iDynoMiCS 2.0.*

Bastiaan J R Cockx, Jan-Ulrich Kreft, Barth F Smets. *Force mediated spatial structure in granular biofilm assembly.*

Tim Foster, Bastiaan J R Cockx, Robert J Clegg, Shirin Moossavi, Kieran Alden, Tariq H Iqbal, Barth F Smets, Jan-Ulrich Kreft. *eGUT, a platform enabling agent-based modelling of the gut microbiome and its spatial interactions with the host mucosa.*