Climate Change Effect on Microbial Interactions using Surrogate Modelling of an Individual-Based Model

Lydia Katsini, Jian Wang, Ihab Hashem, Satyajeet S. Bhonsale,   
Jan F.M. Van Impe\*

BioTeC+ - Chemical and Biochemical Process Technology and Control, KU Leuven Ghent, Gebroeders de Smetstraat 1, 9000 Ghent, Belgium

\*Email: jan.vanimpe@kuleuven.be

Abstract

The climate change effects are numerous, affecting all aspects of life on the planet, including microbes. Bottom-up approaches like Individual-Based Models (IBMs) can describe complex microbial interactions that emerge within a microbial community, at the cost of computational time. To tackle this, surrogate models using data-driven modelling methods are developed. The aim of this paper is to develop a surrogate model of an IBM to assess the effect of increasing temperature on a 3-species community dynamics utilizing the maximum growth rate dependence on temperature. The results show a shift in the bacteria interactions with temperature increases. The importance of the initial population levels on the species prevalence, along with the need to further investigate the influence of temperature on microbial interactions are highlighted.

**Keywords**: climate change, microbial interactions, growth rate, individual-based modelling, k-nearest neighbour classification

* 1. Introduction

Climate change is already taking place and is expected to continue in the future. The impact of climate change is wide, from rising sea level to increased incidence of extreme events, like heatwaves, affecting all aspects of life on the planet (Katsini et al., 2022). According to the one health approach, environmental, human, and animal health are interconnected and addressing one means addressing the others as well. Even though microbes are the common denominator between all three pillars of the one health concept, the microbial response to climate change has received limited attention (Cavicchioli et al., 2019).

A major hurdle in tackling this is the misalignment of the scale between the two systems as well as the high complexity that characterizes both. Oftentimes, mathematical modelling facilitates the study of such involved systems. Regarding the microbial system, developed mathematical models have multiple levels of complexity. Although modelling the growth of individual species is straightforward, modelling the spatial growth of multiple species requires more sophisticated tools such as Individual-Based Models (IBMs). In the case of IBMs, the cell is considered as the modelling unit. The cell’s individual functions, i.e., cell traits and behavior are modelled utilizing mechanistic knowledge. Thus, such a bottom-up approach allows IBMs to capture the population characteristics related to the complex socio-spatial interactions that emerge within a microbial community (Hashem and Van Impe, 2022).

As the effect of climate change on microbial communities is the result of both the individual species response and the species interactions, the ability of IBMs to model both makes them ideal candidates to assess the temperature shift impact on multispecies bacterial growth (Katsini et al., 2022). Previous research regarding the effect of temperature changes on microbial communities focused on multi-species in marine environments (Abreu et al., 2023). Åkesson et al. (2021) referred to the capability of IBMs to study the species interactions under climate change, however, by highlighting their computational cost. Surrogate modelling can overcome this issue by replacing the IBM and providing model simulations more efficiently (Trucchia et al., 2019). Thus, the aim of this paper is to study the effect of temperature shifts on microbial interactions, focusing on species prevalence, by developing an IBM and utilizing surrogate modelling to extend the predictions.

* 1. Materials and Methods

The microbial community used as a case study consists of three bacterial species (P, S, V), each with different temperature ranges for growth. The temperature range for P was 10-30 °C, for S 15-35 °C, and for V 20-45 °C. Based on the above, the temperature range, for which all three species can grow is from 20 °C to 30 °C, which is the range for which the IBM and the surrogate model simulations cover. Since the initial population of each species plays a catalytic role in the microbial interactions, several different levels of ratios among the species at the beginning of the simulations are considered.

* + 1. Modelling the temperature effect on bacterial growth

The effect of temperature on the individual response is considered as the effect on the maximum growth rate, , (h-1) for each species.. As the food microbiology field provides extensive quantitative research on the effect of temperature, , on , the model introduced by Zwietering et al. (1991) was used:

|  |  |
| --- | --- |
|  | (1) |

where are the minimum and maximum temperatures for growth (°C), and are Ratkowsky parameters (°C-1), which were the same for all three species. Figure 1 depicts the dependence on temperature based on Eq. (1) for the 3-species system.

A diagram of a temperature

Description automatically generated

Figure 1: The 3-species system in terms of , as a function of temperature.

* + 1. Development of the Individual-Based Model

An IBM simulation consists of two types of entities: the microbial cells and their environment. Each individual cell is defined by its species genre, radius, mass, spatial coordinates, and maximum nutrient uptake rates. For each cell intracellular and diffusion processes, and intercellular and environmental interactions are modeled (Tack et al., 2017, Hashem and Van Impe, 2022). Intracellular processes include cell growth, maintenance, and division, with spatial overlap due to high cell density managed by a cell shoving algorithm (Kreft et al., 2001). The diffusion processes are governed by the second law of Fick modulating nutrient and toxin distribution. IBM simulations are run using MICRODIMS, an in-house IBM modeling platform (Tack et al., 2017), at two time resolutions, one for rapid dynamics, e.g., glucose uptake, and a second one for slower processes, e.g., cell reproduction. The environment is represented as 40 mm × 20 mm section of a 4 mm thick gel, discretized into square units of 10 mm, each characterized by nutrient and toxin concentrations.

* + 1. Development of the surrogate model

The species prevalence obtained at the end of the IBM simulation was used as the label with temperature and initial population ratio as predictors. This means that the modelling exercise was a classification task: predict which of the three species prevails under given conditions. Three different algorithms were trained, i.e., decision tree, support vector machine, and *k*-nearest neighbor, and evaluated in terms of accuracy on unseen data, with the latter achieving the highest performance (87.7 % accuracy). 5-fold cross-validation was used to avoid overfitting and the data were standardized before training to maximize accuracy.

* + - 1. k-Nearest Neighbor (KNN)

The *k*-Nearest Neighbor (KNN) algorithm is a supervised learning method that facilitates the classification of data based on their distance to each other (Zhang and Zhou, 2007). According to this method, the class for each unseen point is selected by evaluating the classification of its *k* nearest neighbors with the most common class among them assigned to the unseen point. Following hyperparameter tuning, the number of nearest neighbors was 10, while the distance metric was the squared inverse of the Euclidean distance.

* 1. Results and Discussion

This section is divided into two main parts: first the results obtained from the IBM, and then the training of the surrogate model along with the predictions for the full set of conditions, i.e., temperature and ratio of initial populations.

* + 1. IBM of the 3-species system

The IBM results show that at the end of each simulation only one species out of the three prevails, therefore the results indicate the prevailing one (Figure 2 and 3). Based on Figure 1, there are two temperature levels where two species have equal , which are higher than the third one. Firstly, at 22.62 °C for P and S, and secondly at 25.17 °C for S and V. Considering the stochastic nature of the developed IBM, this is also reflected in the results of the different iterations at those temperature levels. As shown in Figure 4, at 22.62 °C as the ratio of the initial population of P to V increases, the prevailing species interchanges between P and S.

A red and green pixelated objects

Description automatically generated  
(a)  
A red white and blue flag

Description automatically generated  
(b)

Figure 2: Spatial distribution of the 3-species system (P-green, S-blue, V-red) from the IBM at the middle (a) and the end (b) of the simulation for 25.17 °C.

A group of blue and green objects

Description automatically generated  
(a)  
A group of colorful balls

Description automatically generated  
(b)

Figure 3: Spatial distribution of the 3-species system (P-green, S-blue, V-red) from the IBM at the middle (a) and the end (b) of the simulation for 22.62 °C.

On visually inspecting the prevalence results at 25.17 °C, this distinction is less clear as there is a ratio at which each species prevails at least once through the 10 repetitions. This could be explained by the difference in the values between the two species that have the same and the third one. For 24.13 °C, there is no variation in the prevalent species for the evaluated conditions. This could be attributed to the fact that the tested initial population ratios did not approximate the tipping point for the shift in the prevalence.

A diagram of a temperature

Description automatically generatedA diagram of a temperature

Description automatically generatedA diagram of a temperature

Description automatically generated

Figure 4: Prevalence results from the developed IBM for selected temperature levels.

* + 1. Surrogate model training and application

As mentioned in the previous section, the KNN classification algorithm was the best performing, and thus was selected for the surrogate model. Common practice when building classifiers is to plot the confusion matrix, which provides information on the misclassifications during training. Figure 5 depicts the confusion matrix of the KNN algorithm. The misclassification instances for P and V are lower compared to the misclassification instances for S, which can be explained by inspecting Figure 1. The temperature range where S is prevailing is smaller compared to the other two species and S is the species that has a “conflict” of prevalence with both the other two. Therefore, establishing its clear advantage of prevalence would require additional simulations from the IBM including smaller differences among the initial populations.

A diagram of a confusion matrix

Description automatically generated

Figure 5: Confusion matrix for the training of the k-nearest neighbor (KNN) classifier.

The developed surrogate model was applied for more coarse initial population fractions by covering the relevant temperature range and the species prevalence results are illustrated in Figure 6. It is important to note that for interpreting the results, the fractions correspond to an increase of the initial population of species P, as the ratios increase. For the temperature range 20-22 °C the prevalent species is constantly P, which is expected as the increasing ratios evaluated are to the benefit of species P. Moreover, for this range based on Figure 1, species P has the highest . At 23 °C, even though species P has the highest value (Figure 1), as the initial population of species S increases (low initial P/S and P/V ratios), the prevalent species becomes species S. Surprisingly this pattern is repeated for 24 °C, although for this temperature it is species S with the highest value (Figure 1).

A diagram of a temperature

Description automatically generatedA graph showing different types of species

Description automatically generated

Figure 6: Species prevalence based on the surrogate model for coarser initial population fractions.

One reason for obtaining this result could be the increased misclassification instances of the surrogate model regarding species S. For 25-26 °C, depending on the initial populations, all three species could prevail, which is expected for 25 °C but not for 26 °C. Since at 26 °C, species V has the highest (Figure 1), there are some conditions in which species S prevails, even though the initial population ratios generally correspond to the increase of the initial population of species P. For 27-30 °C, the results are unsurprising, as species V has the highest and the ratios correspond to increase of the initial population of species P.

* 1. Conclusions

Climate change is an urgent and profound issue owed to the amplitude of its effects. This includes, among other aspects of life, microbial life. Climate change driven biodiversity loss is gaining attention, creating the need for studying the effects of shifting temperatures on microbial communities. Nowadays evidence-based policy making is relying more and more on mathematical modelling. In this paper an Individual-Based Model (IBM) of a 3–species bacterial community is developed considering the effect of temperature on the maximum growth rate (. For the IBM it is computationally expensive to simulate a wide range of conditions, thus, a data-driven surrogate model, by applying the k-Nearest Neighbour algorithm, is built to overcome this issue. The results show that even though the is a good indicator for species prevalence, it fails to take into account species interactions, in contrast to the IBM simulations. Thus, such involved models are the adequate tools to evaluate the effect of temperature shifts on bacterial communities. When isolating the species interaction effect, i.e., where values between species are equal, results show that the initial population ratios are the determinant for prevalence. As the more IBM simulations are used, the more reliable the results of the surrogate model are, future work plans include exploring coarser grids of the temperature-initial populations space. Overall, this work contributes to the understanding of the microbial response to climate change, which can be utilized for future-proofing purposes.

Acknowledgement

This work was supported by the PROTECT and E-MUSE projects funded by the European Union’s Horizon 2020 Research and Innovation Programme [MSCA grant 813329 and 956126].

References

Abreu, C.I., Dal Bello M., Bunse C., Pinhassi J., Gore J., 2023.Warmer temperatures favor slower-growing bacteria in natural marine communities. Sci Adv 9(19)

Åkesson, A., Curtsdotter A., Eklöf A., Ebenman B., Norberg, J., Barabás, G., 2021. The importance of species interactions in eco-evolutionary community dynamics under climate change. Nat Commun 12

Cavicchioli, R., Ripple, W.J., Timmis, K.N. et al., 2019. Scientists’ warning to humanity: microorganisms and climate change. Nat Rev Microbiol 17

Hashem, I. & Van Impe, J.F.M., 2022. A Game Theoretic Analysis of the Dual Function of Antibiotics. Front Microbiol 12

Katsini, L. Bhonsale S., Akkermans S., Roufou S. Griffin S., Valdramidis V., Misiou O., Koutsoumanis K., Muñoz López C.A., Polanska M., Van Impe J.F.M., 2022. Quantitative methods to predict the effect of climate change on microbial food safety: A needs analysis. Trends Food Sci 126

Kreft, J.U., Picioreanu, C., Wimpenny, J.W. and van Loosdrecht, M.C., 2001. Individual-based modelling of biofilms. Microbiology, 147(11)

Tack, I.L., Nimmegeers, P., Akkermans, S., Hashem, I. and Van Impe, J.F.M., 2017. Simulation of Escherichia coli dynamics in biofilms and submerged colonies with an individual-based model including metabolic network information. Frontiers in microbiology, 8

Truchhia, A., Mattei M.R., Luongo V., Frunzo L., Rochoux M.C., 2019. Surrogate-based uncertainty and sensitivity analysis for bacterial invasion in multi-species biofilm modelling. Commun Nonlinear Sci Numer Simul 73

Zhang, M., Zhou Z., 2007. ML-KNN: A lazy learning approach to multi-label learning. Pattern Recognit 40(7)

Zwietering MH, de Koos JT, Hasenack BE, de Witt JC, van't Riet K. Modeling of bacterial growth as a function of temperature. Appl Environ Microbiol. 1991 Apr;57(4)