Universidade do Minho
Escola de Engenharia

Juliana Marques Pereira

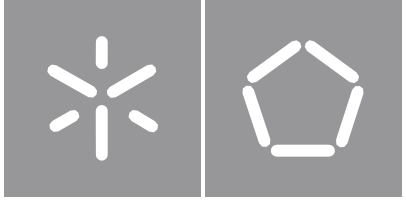
Ecological modelling to describe the role of light on microbial interactions in *Ulva spp*. with implications in aquaculture

October of 2022

Ecological modelling to describe the role of light on microbial interactions in *Ulva spp*. with implications in aquaculture

Juliana Pereira

UMinho | 2022



Universidade do Minho
Escola de Engenharia

Juliana Marques Pereira

Ecological modelling to describe the role of light on microbial interactions in *Ulva spp*. with implications in aquaculture

Master's Dissertation
Master in Bioinformatics

Work carried out under the guidance of
Doctor Eva Balsa-Canto
Doctor Óscar Dias

October of 2022

Acknowledgments

Firstly I would like to thank my family for the support they gave me in this journey, as well as the advice and backup to make this dissertation the best it could be.

Secondly, I would like to thank Dr. David Henriques from the Institute of Marine Research (IIM-CSIC) in Vigo, Spain, for suggesting and challenging me to do this dissertation. I would also like to thank Dr. Eva Balsa-Canto from the Biosystems and Bioprocess Engineering Group of the Institute of Marine Research (IIM-CSIC) for all the support, help, good disposition, availability, even in the tightest hours, and patience she gave me in the development of this dissertation. I would also like to thank my colleagues in this research group Dr. David Henriques, PhD student Artai Moimenta, PhD student Diego Troitiño Jordedo, PhD student Geoffrey Roudaut and Dr. Benjamin Kuchen for all the help either in the integration in Spain, in the institute and in the development of this dissertation, in those hours when nothing worked and despair was installed they were always there for advice.

My thanks to Dr. José Pintado from the Ecology and Marine Resources group as well as PhD student Gonzalo Del Olmo Berenguer (member of the same group) for all the help either in the understanding of the biological component of this dissertation and in the help provided for the realization of the necessary laboratory experiment to obtain new data more adjusted to this model. Also a word of thanks to all the people at the institute that I have not mentioned otherwise the list would be enormous, but who interacted with me and made me feel at home in Spain. It was an incredible experience to meet you all and I hope to come back one day to catch up.

Finally, a thanks to Dr. Óscar Dias from the BIOSYSTEMS group of the University of Minho for facilitating my going to Spain and for the knowledge he gave me about bioinformatics modelling.

Thank you very much to all of you.

Funding

This dissertation has been funded by a Research Introduction Grant JAE Intro ICU financed by the Spanish Ministry of Science and Innovation and the Consejo Superior de Investigaciones Científicas.

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

Resumo

Com a população mundial a aumentar e as quotas de pesca a estagnar, são necessárias novas formas de produzir alimentos de origem marinha sem comprometer o ambiente. Uma destas formas é através da Aquacultura Multi-Trófica Integrada num Sistema de Recirculação de Água (IMTA-RAS). Um dos maiores problemas das produções intensivas é o dos agentes patogénicos ou oportunistas que podem causar uma taxa de mortalidade dos peixes da ordem dos 75% ou mesmo superior. Num sistema do tipo IMTA-RAS, são necessárias pelo menos duas espécies, uma alimentada (como os peixes) e outra extrativa (como, por exemplo, as algas) capaz de remover os nutrientes orgânicos e inorgânicos da água. Neste trabalho, consideramos *Ulva ohnoi*, uma espécie que tem atraído grande atenção devido à sua facilidade de cultivo, produtividade, elevado teor proteico, e outros nutrientes essenciais. As comunidades bacterianas associadas a *Ulva spp.* desempenham um papel funcional importante tanto na morfogénese como na reprodução de algas. Uma espécie bacteriana específica encontrada na superfície de *Ulva*, *Phaeobacter gallaeciensis*, pode produzir um antibiótico natural atuando contra agentes patogénicos oportunistas como o agente patogénico dos peixes, *Vibrio anguillarum*. No entanto, a retenção de *Phaeobacter gallaeciensis* na superfície das algas é afetada pelas condições de funcionamento do sistema IMTA. Mais especificamente, na intensidade da luz. Aqui propusemos a formulação de um modelo ecológico para ter em conta esses efeitos e descrever como a intensidade da luz afeta as interações entre espécies: alga - microbioma bacteriano - *Phaeobacter*. Para este propósito, primeiro foi realizada uma experiência para obter dados de crescimento das espécies com diferentes intensidades de luz. Os dados foram então utilizados para identificar iterativamente um modelo Lotka-Volterra. Foi realizada uma análise de identificabilidade, o que levou a um modelo reduzido. Posteriormente, utilizando uma abordagem de estimativa de parâmetros multi-experimental, foram estimados os tipos de interações e a sua dependência da intensidade luminosa. Os resultados finais revelam que a taxa de crescimento das algas depende da intensidade da luz. No entanto, intensidades mais elevadas podem ser prejudiciais. Além disso, as melhores condições de crescimento das algas parecem ser as piores para a retenção de *Phaeobacter*. Embora estes resultados necessitem de mais validação experimental, concluímos que a intensidade da luz deve ser seleccionada para obter bons compromissos entre o crescimento das algas e a produção de antibióticos naturais.

Palavras-chave: *Ulva ohnoi*, *Phaeobacter gallaeciensis*, modelação ecológica, Lotka-Volterra, luz, aquacultura.

Abstract

With the world's population increasing and fishing quotas stagnating, new ways to produce marine food without compromising the environment are needed. One of these ways is through Integrated Multi-Trophic Aquaculture in a water Recirculation System (IMTA-RAS). One of the greatest problems of intensive productions is the opportunistic pathogens that can cause a mortality rate of fish on the order of 75% or even higher. In an IMTA-RAS type of system, at least two species are needed, one fed (like fish) and other extractive (like, for example, algae) capable of removing organic and inorganic nutrients from the water. In this work, we considered *Ulva ohnoi*, a species that has attracted important attention due to its ease of cultivation, productivity, high protein content, and other essential nutrients. The bacterial communities associated with *Ulva spp.* play an important functional role in both morphogenesis and algae reproduction. A particular bacterial species found on the surface of *Ulva*, *Phaeobacter gallaeciensis*, can produce a natural antibiotic acting against opportunistic pathogens as fish pathogen, *Vibrio anguillarum*. However, the retention of *Phaeobacter gallaeciensis* on the surface of the algae is affected by the operating conditions of the IMTA system. More specifically, on light intensity. Here we proposed the formulation of an ecological model to account for those effects and describe how light intensity affects the interactions between species: algae - bacterial microbiome - *Phaeobacter*. For this purpose, an experiment was first performed to obtain species growth data at different light intensities. The data were then used to iteratively identify a Lotka-Volterra model. Identifiability analysis was performed, which led to a reduced model. Afterwards, using a multi-experiment parameter estimation approach, the types of interactions and their dependence on the light intensity were estimated. The final results reveal that the algae growth rate depends on the light intensity. However, higher intensities can be detrimental. In addition, the best algae growth conditions appear to be the worst for retention of *Phaeobacter*. Although these results need further experimental validation, we concluded that light intensity must be selected to obtain good compromises between algae growth and the production of natural antibiotics.

Keywords: *Ulva ohnoi*, *Phaeobacter gallaeciensis*, ecological modeling, Lotka-Volterra, light, aquaculture.

Índice

Índice de Figuras

Índice de Tabelas

1 Introduction

Fish is an important dietary source of proteins of high biological value, vitamin D, vitamin E, iodine, and long chain omega-3 polyunsaturated fatty acids ($\omega - 3$ PUFA), and is low in saturated fatty acids. In particular, fish and seafood are the primary dietary source of eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA), which provide various health benefits, such as a decrease in the risk of cardiovascular diseases. With its valuable properties, fish can play an important role in correcting unbalanced diets and countering obesity.

Due to the growth of the population, fish consumption has increased worldwide. Fisheries and aquaculture are important sources of nutritious food. The total global capture fishery production had increased over many decades during the past century. However, the sector is under stress from pollution, habitat degradation, overfishing and harmful practices, climate variability, and climate change (thermal structure, storm frequency, acidification, etc.). These stresses contribute to modify species composition, productions and yield, distribution and seasonality, or the presence of pathogens which affect safety and security, as well as efficiency and costs.

Since capture fishery production has been rather static since the 1990s, aquaculture has been responsible for the growth of the supply of fish for human consumption. In 2020, aquaculture production represented 49.2% (122.6 million tonnes) of global production. The production of aquatic animals (87.5 billion tonnes) corresponds to US$ 264.8 billion. In terms of algae, 35.1 million tons are produced in aquaculture, representing US$ 16.5 billion worldwide. 89% of the global production of aquatic animals through fisheries and aquaculture is consumed by humans. Over the years, we have seen an increase in aquaculture production. This increase helps to address the existing fish shortage on the market (Chopin et al., 2010; Lavaud, Guyondet, Filgueira, Tremblay, & Comeau, 2020; FAO, 2022a).

Intensive fish production generates a large amount of organic waste that remains in the water. Therefore, to reduce the ecological footprint of the production processes, integrated multi-trophic aquaculture (IMTA) has emerged (Chopin et al., 2010; Oca, Cremades, Jiménez, Pintado, & Masaló, 2019; Fraga-Corral et al., 2021). IMTA aims to produce together two types of species, as fish that are fed (usually with a specific diet); and those extractive species, which use excess inorganic and organic nutrients released by the fish to grow (such as algae or suspension and deposit feeding invertebrates) (Chopin et al., 2010).

In nature, green algae of the genus Ulva grow in regions where there are large amounts of nutrients, such as coastal zones. These algae adapt easily to environmental conditions and have a great capacity for nutrient uptake (such as nitrogen) and growth (Oca et al., 2019; Lavaud et al., 2020). Ulva species have attracted important attention in aquaculture due to their ease of cultivation, productivity, high protein content, and other essential nutrients.

Integrated multitrophic aquaculture with recirculation systems (IMTA-RAS) reduces water use and allows the installation in locations away from the coast. This method can combine more than two species of different multitrophic levels to reuse nutrients in water. For example, the production of fish and seaweed can recycle fish residues and increase the productivity of both species in a closed system. Furthermore, cultivation in IMTA-RAS systems could improve the control of disease outbreaks.

The bacterial communities associated with *Ulva spp.* play an important functional role in both morphogenesis and algae reproduction. In addition, *Ulva spp.* hosts antibiotic-producing bacteria (APB) with known antagonism to fish pathogens such as *Vibrio anguillarum*. Therefore, it is possible to use *Ulva spp.* colonized with APB as a pathogen control strategy in multitrophic fish-algal cultures in water recirculation systems, improving fish health.

However, the optimal conditions for the cultivation of *Ulva* have a determining influence on the maintenance of these APBs. High light intensities that lead to high algae growth rates correlate with a lower persistence of APBs. Therefore, the question is how to grow algae to optimize the permanence of APB.

In this project, this question is addressed using a model-based approach. Modeling the interactions between *Ulva ohnoi*, its bacterial microbiome, and *Phaeobacter gallaeciensis* will be considered taking into account the role of light intensity. The model will then be used to identify the best compromise between algae growth and the permanence of APB in the system.

1.1 Objectives

The objective of this dissertation is, through a Lotka-Volterra model, to understand the optimal light intensity for algae *Ulva ohnoi* growth-retaining bacteria *Phaeobacter gallaeciensis* on its surface in an IMTA-RAS system.

To accomplish this general objective, the following sub-objectives are to be achieved:

- Analyze available experimental data and perform additional experiments if required.
- Formulate a three-species generalized Lotka-Volterra model
- Perform the model structural identifiability analysis
- Calibrate the model to decipher the types of interactions between species: algae-bacterial microbiome; algae-*Phaeobacter*; bacterial microbiome-*Phaeobacter*.
- Characterize the interaction coefficients as functions of the light intensity.
- Identify a range of light conditions to improve the permanence of *Phaeobacter* in the IMTA system.

2 State of the art

At this moment we are 7.6 billion people inhabiting this planet. However, according to world estimates, it is expected that in 2030 there will be 8.6 billion people, in 2050 there will be 9.8 billion and in 2100 there will be 11.2 billion people (Nations, n.d.). One of the biggest concerns today is: Will we be able to feed so many people? Will we reach the carrying capacity of the planet? How can we sustain population growth?

In agriculture, due to the scarcity of water, governments are currently limiting the use of water for irrigation and consumption (e.g., Portugal and Spain) due to the lack of rain and extreme drought, with temperatures rising increasingly caused by climate change. Also in regions such as Galicia (Spain), the use of nonessential water is prohibited (e.g., closing footwash and shower taps on beaches, limiting the filling of swimming pools, or washing cars).

In terms of fishing, the situation is also dramatic. In the 1990s there was a peak in catches (Pauly & Zeller, 2016). With an increase in population, stagnation of the fishing quota and the average annual consumption of marine animals that increases to around 3% per year (9.0 kg per capita in 1961 to 20.5 kg per capita in 2019), it is necessary to find alternative ways to produce fish (FAO, 2022a). Here, aquaculture plays an important role.

2.1 Aquaculture

2.1.1 Production

Currently, aquaculture represents a way to bring more fish to human consumption.

It is estimated that 178 million tonnes of aquatic organisms were produced worldwide in 2020. Of this, 49% (88 million tonnes) comes from aquaculture production. Of this global production, 112 million tonnes are related to marine organisms and 30% of this figure comes from aquaculture production. The remaining 66 million tonnes represent production in inland waters. Of this, 83% comes from aquaculture production. In monetary terms, aquaculture production of aquatic organisms is valued at US$ 264.8 billion. In addition to aquatic animals, 35.1 million tons of algae (wet weight) have been produced worldwide, of which 97% comes from aquaculture (mainly marine) and are valued at US$ 16.5 billion. Of the annual production of aquatic animals, approximately 89% (157 million tonnes) is destined for human consumption. The

remaining 11% are destined for non-food uses, such as the production of fish meal and fish oil (FAO, 2022a). But will this production be enough to feed the human population in 8 years? And 28 years from now?

The data we have say that it is not enough. In 2020, between 720 and 811 million people around the world are believed to be facing hunger. About half of the 2.37 billion people who face moderate and severe food insecurity are in Asia and a third are in Africa (UNICEF et al., 2021). Therefore, it is necessary to find new ways to increase food production without causing so much damage to ecosystems, without relying so heavily on fresh water, and which at the same time is safe for human consumption.

2.1.2 Marine aquaculture

Although most aquaculture is monoculture in extensive and semi-intensive regimes, especially in marine aquaculture, intensive marine-fed aquacultures have begun to appear (Chopin et al., 2010). The problem with these intensive productions is that they are usually concentrated in a certain geographical location and are often found in suboptimal regions where the assimilation capacity is poorly understood and may be excessive (Chopin et al., 2010). Unfortunately, monocultures cause the independence of production of the two types of aquaculture (fed versus extractive) in geographically different regions (Chopin et al., 2010). In an aquaculture with fixed spatial limits, where the aim is to increase profitability, the environment is often overloaded with individuals. This can cause the natural system to become destabilized (Chopin et al., 2010; Fraga-Corral et al., 2021). It is necessary to find a way to compensate for excess nutrients in the system in water to avoid overloading the system.

Intensively farmed fish species retain only 20-40% of the nitrogen consumed, demonstrating the large amount of nutrients such as nitrogen left in the water (Teles, Couto, Enes, & Peres, 2020). If extractive aquaculture (such as seaweed production) is implemented in the system, it is possible to convert excess nutrients and energy into commercially valuable products (Chopin et al., 2010; Oca et al., 2019).

Therefore, in order for aquaculture to be more environmentally, economically and socially sustainable, multitrophic integrated aquaculture (IMTA) has emerged (Chopin et al., 2010; Correia et al., 2020; Fraga-Corral et al., 2021).

2.1.3 Integrated Multi-trophic Aquaculture in Recirculation Systems (IMTA-RAS)

Integrated Multi-trophic Aquaculture (IMTA) allows the intensive production of species fed (usually with sustainable commercial diets) with extractive species that use excess organic and inorganic nutrients in water to grow (Chopin et al., 2010; Oca et al., 2019). IMTA allows mixing and producing organisms from more than two trophic levels. Examples include fish such as sole (*Solea senegalensis)* and sea bass (*Dicentrarchus labrax*), seaweed species, worms, crustaceans, echinoderms and shellfish, among others (Chopin et al., 2010; Oca et al., 2019; Correia et al., 2020).

IMTA systems can be implemented within a water recirculation system (recirculating aquaculture system, RAS). RAS is based on the use of mechanical and biological filters that allow the reuse of the water used in the production of fish and other organisms in aquaculture (Figures 1 and 2) (Bregnballe et al., 2010). Depending on the amount of water that can be recirculated or reused in aquaculture, the intensity of the RAS can be adjusted. The formula used to calculate the degree of recirculation is as follows:

$$\frac{Internal recirculation flow}{Internal recirculation flow + New water intake} * 100 \tag{2.1}$$

The amount of water that is reused in an IMTA-RAS system allows the removal of nutrients left over from the species being produced, since there is recycling of waste from one species to feed other species at other trophic levels. This system not only allows for more intensive production without the species being dependent on existing environmental conditions in nature, but also allows the development of new products, thus increasing the economic return of aquaculturists (Bregnballe et al., 2010; Qiu, Carter, Hilder, & Hadley, 2022).

Qiu et al., 2022 have developed a model with three trophic levels: Atlantic salmon, *Ulva lactuca* and an invertebrate grazing species. The RAS system in this model had incorporated biological and mechanical filtration. In this model, different waste treatment methods were simulated using filtration of the RAS system with the incorporation of algae. Simulations of this model with the addition of algae and biological and mechanical filtration mechanisms were able to decrease the level of dissolved inorganic nitrogen by 66% and the level of dissolved inorganic phosphorus by 31% compared to the model without algae. When

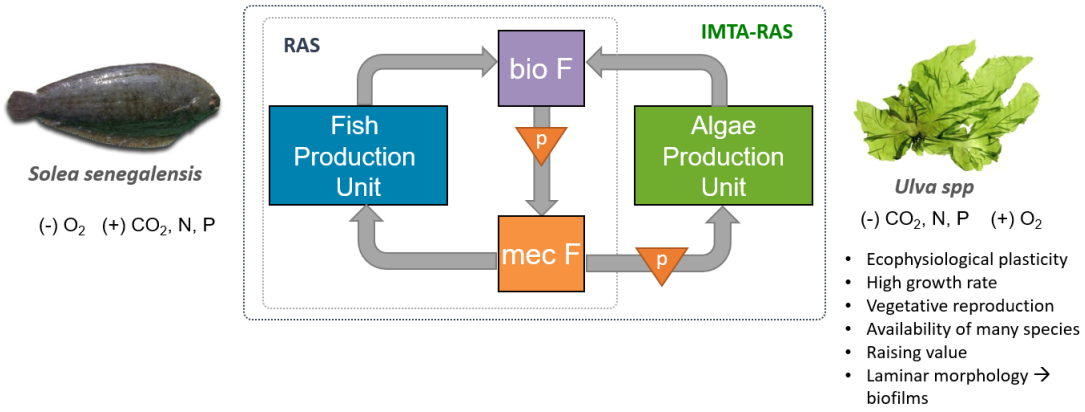


Figure 1: Scheme with an example of an integrated multitrophic aquaculture with a recirculating system (IMTA-RAS) with two trophic levels (*Solea senegalensis* and *Ulva spp.*). In RAS is possible to see a fish production unit, a biofilter (bio F), a pump (P) and a mecanic filter (mec F). The IMTA-RAS is adding another tank with algae production. This tank will oxigenate the water and remove carbon dioxide, nitrogen and phosphorus.

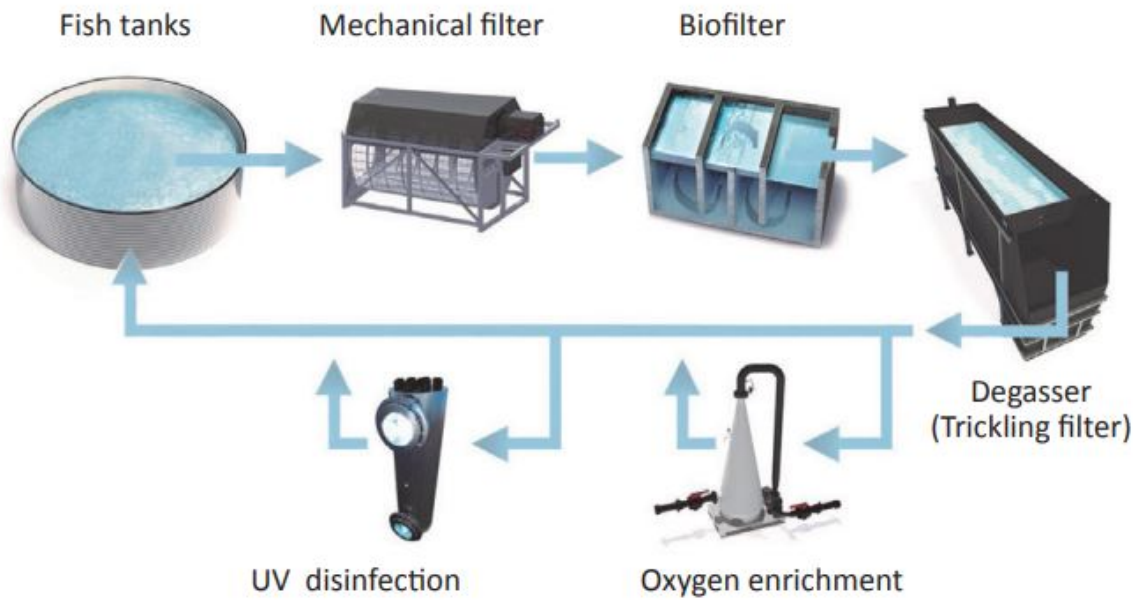


Figure 2: Scheme of a water recirculation system (RAS) where there is a basic water treatment through mechanical filtration, biofiltration and aeration/stripping. An oxigen enricher or a UV disinfection system can also be included in this system (Bregnballe et al., 2010).

extraction was optimized, the removal of these residues increased to 94% for dissolved inorganic nitrogen and 45% for dissolved inorganic phosphorus, and the biomass of the by-product increased by 41%. This demonstrates how reuse of waste/nutrients is important not only for increasing feed and profits but also for more sustainable aquaculture.

When it comes to feeding fish, the ideal is to feed them several times a day with dry food, as this type of food is safer and contains the most appropriate energy and nutritional content to meet the biological needs of fish at different stages of growth. This type of feed also allows for maximum protein absorption and a significant reduction in certain excretion products, such as ammonia (Bregnballe et al., 2010).

With regard to production tanks, these must be suitable for the species being produced. For example, demersal species (species that, despite their swimming ability, live most of the time at the bottom), such as sole and turbot, require shallow depth and water current velocity tanks due to their biological and ecological characteristics. However, pelagic species, such as salmonids, require larger and deeper tanks with a higher current velocity of the water (Bregnballe et al., 2010).

One of the most important advantages of RAS is disease control. In traditional fish farming, the water introduced comes from the environment, such as rivers, lakes, or the sea. This water can carry certain pathogens that can ruin production. In a recirculating water system, there is more control over what is in the water, and it is much less likely that a pathogen will affect production, so no chemicals are needed to treat the diseased organisms. However, diseases can be introduced into the system by taking infested eggs or fish. To prevent this, it is best to choose eggs that can be disinfected against diseases and avoid stocking fish from outside (Bregnballe et al., 2010).

2.2 Why is there a growing interest in algae?

There are three main reasons for the growing interest in algae. First, algae are extremely versatile and can be used in a variety of industries such as pharmaceuticals, cosmetics, textiles (thickening / emulsifying agents), agriculture and bioremediation, as well as in human food and animal feed (Mantri, Kazi, Balar, Gupta, & Gajaria, 2020; FAO, 2022b). Second, algae are primary producers and are at the base of the aquatic food chain, producing the food resources that fish are adapted to consume. Thus, algae are good candidates for multitrophic aquaculture. Third, algae have good nutritional value. Due to population growth and resource scarcity, there is a demand for new, more nutritious, and sustainable

food sources. The main nutritional characteristics of algae are (FAO, 2022b):

- Algae have various minerals such as iron, calcium, iodine, potassium and selenium, and vitamins such as A, C and B-12;
- They are one of the few sources of natural omega-3;
- They have soluble dietary fiber and some species can be good sources of protein;
- They can be beneficial for health, as certain species have anti-inflammatory, prebiotic, and antioxidant components;
- The addition of seaweed to animal feed has shown several advantages. In cattle, this addition to feed has been shown to decrease methane production by up to 98%, increase weight by 42% and without side effects (Kinley et al., 2020).

Many different algae already play a vital role in aquaculture. Algae are photosynthetic organisms (containing chlorophyll) that obtain energy from the sun and carbon from carbon dioxide. Their size ranges from one micrometer to many meters. All organisms that use carbon dioxide to meet their carbon requirement are called autotrophs. Algae are generally beneficial in aquaculture, providing oxygen and a natural food base for cultured fish, but also acting as biofilters that are able to remove nitrogenous waste from recirculating marine aquaculture.

2.3 Ulva spp.

2.3.1 Ecology

In this genus, there are marine and freshwater species. The largest number of recorded species is found in Asia, but these algae can be found on all seven continents (Mantri et al., 2020). Tubular species are known to be more tolerant to low salinity and can therefore be seen in fresh water. On the other hand, leafy stalk species are found in marine environments. *Ulva* species have been found in waters with salinity between <0.5 and 49 (Rybak, 2018).

The geographical distribution of the different species is related to their tolerance and adaptation to different temperatures and salinities. Other factors that are also important for growth are light and the concentration of nutrients in water (Fortes & Lüning, 1980; Oca et al., 2019). The success of the wide

geographical distribution of algae of this genus is due to the physiological, biochemical, and molecular characteristics that allow algae to adapt to different environmental conditions (Mantri et al., 2020). It is this plasticity that allows extreme events to occur, such as algae blooms (also known as green tides) due to eutrophication, which can cause hypoxia or even anoxia events in these areas (Lavaud et al., 2020).

The Ecology and Marine Resources group at IIM-CSIC is interested in implementing an IMTA-RAS system for aquaculture of *Solea senegalensis*, due to its commercial value. The aim is to define a productive system in combination with *Ulva ohnoi* that can act as a biofilter (Oca et al., 2019).

2.3.2 Ulva ohnoi: Effects of light on growth and productivity

Oca et al., 2019, developped a model of algal growth was developed to analyze the influence of the density of the light and biomass stock on the production of effluent fed *Ulva ohnoi* from tanks with *Solea senegalensis* in an IMTA-RAS system. In this study, growth and productivity rates were determined in three flat-bottom algae tanks with different incident photon irradiances (163, 280 and 886 $\mu mol\ m^{-2}\ s^{-1}$) in a photoperiod of 12:12h. The results of this study showed that algae exposed to the highest intensity of light at a stock density below $170 gdw\ m^{-2}$ suffered morphological damage and a strong decrease in growth rate. This demonstrates that algal growth is not linearly related to the light intensity directed at the algal surface but that there is a specific optimum intensity for algal growth.

2.3.3 Microbial interactions in Ulva spp.

Due to the production and release of nutrients, macroalgae are good candidates for bacterial colonization and proliferation (Roth-Schulze et al., 2018). Different algae species may have a unique bacterial community that may be related to the external morphology of the algae and the exuded substances (Longford et al., 2007; Lachnit, Fischer, Künzel, Baines, & Harder, 2013). In the genus *Ulva*, existing bacterial communities are recruited from surrounding seawater (Burke, Thomas, Lewis, Steinberg, & Kjelleberg, 2011b). However, the bacteria found in different algae of the same species differ substantially. This shows that the bacteria found in *Ulva spp.* are not related by phylogeny or taxonomy of community members, but are related through a core set of functions involved in algal-bacterial interactions (Burke, Steinberg, Rusch, Kjelleberg, & Thomas, 2011a; Burke et al., 2011b; Roth-Schulze et al., 2018).

Within the various bacteria found on the surface of algae, bacteria of the genus *Phaeobacter* have already been found. These bacteria have the ability to decrease growth or even kill certain pathogens that can be found in fish by producing the antibiotic TDA (tropodithietic acid) (Brinkhoff et al., 2004).

2.4 Phaeobacter gallaeciensis

2.4.1 Importance of Phaeobacter gallaeciensis in an IMTA-RAS system

In intensive fish production, the main cause of fish death in farmed fish is the presence of pathogen or opportunistic bacteria such as those belonging to the *Vibrionaceae* family (Prol-García, Gómez, Sánchez, & Pintado, 2014). Two of the bacteria (*Vibrio anguillarum* and *Vibrio splendidus*) belonging to this family are found in the normal microbiome of these farms and can enter these facilities through seawater or already infected fish, leading to generalized infections and mortality (Prol-García, Planas, & Pintado, 2010; Prol-García & Pintado, 2013; Prol-García et al., 2014). Antibiotics and disinfectants can be used to treat this infection and clean the system (Prol-García et al., 2014). However, excessive use of antibiotics can cause the emergence of antibiotic-resistant bacteria. Disinfectants, on the other hand, affect the stable microbial population, which can lead to disequilibrium in the rearing system, allowing opportunistic bacteria to colonize (Prol-García & Pintado, 2013; Prol-García et al., 2014).

Phaeobacter gallaeciensis is a Gram-negative $\alpha-$proteobacterium that belongs to the Rhodobacteraceae family (Martens et al., 2006). This bacterium produces tropodithietic acid (TDA), which is an antibacterial compound that has the ability to efficiently inhibit bacteria *Vibrio anguillarum* and *Vibrio splendidus* (Brinkhoff et al., 2004; D'alvise et al., 2012; Prol-García & Pintado, 2013; Prol-García et al., 2014). According to Prol-García & Pintado, 2013, in a culture of turbot (*Psetta maxima*) larvae infected with bacteria of the genus *Vibrio*, the mortality rate was around 76%. Including a *Phaeobacter* biofilter, the mortality rate of the larvae was much lower (35-40%) and at the end of 144 h no *Vibrio* could be identified in the water by real-time PCR analysis. The presence of *Phaeobacter* lasted only 11 days. Prol-García et al., 2014 also analyzed, with biofilters made of different materials, whether *Phaeobacter* could control *Vibrio* infection. This study demonstrated for the first time the benefits of using biofilters with *Phaeobacter*. The study also showed that the presence of probiotics in the biofilter can make the system less susceptible to colonization by opportunistic pathogenic bacteria by stabilizing

the bacterial microbiota. The next step is to understand whether it is possible to introduce *Phaeobacter* so that it stays longer in the system.

Because *Phaeobacter* can be found on the surface of algae belonging to the genus *Ulva*, it is possible to design an IMTA system in which the combination of *Ulva* and *Phaeobacter* can prevent opportunistic bacteria such as *Vibrio* from colonizing and causing infection/mortality in fish species [Figure 3].

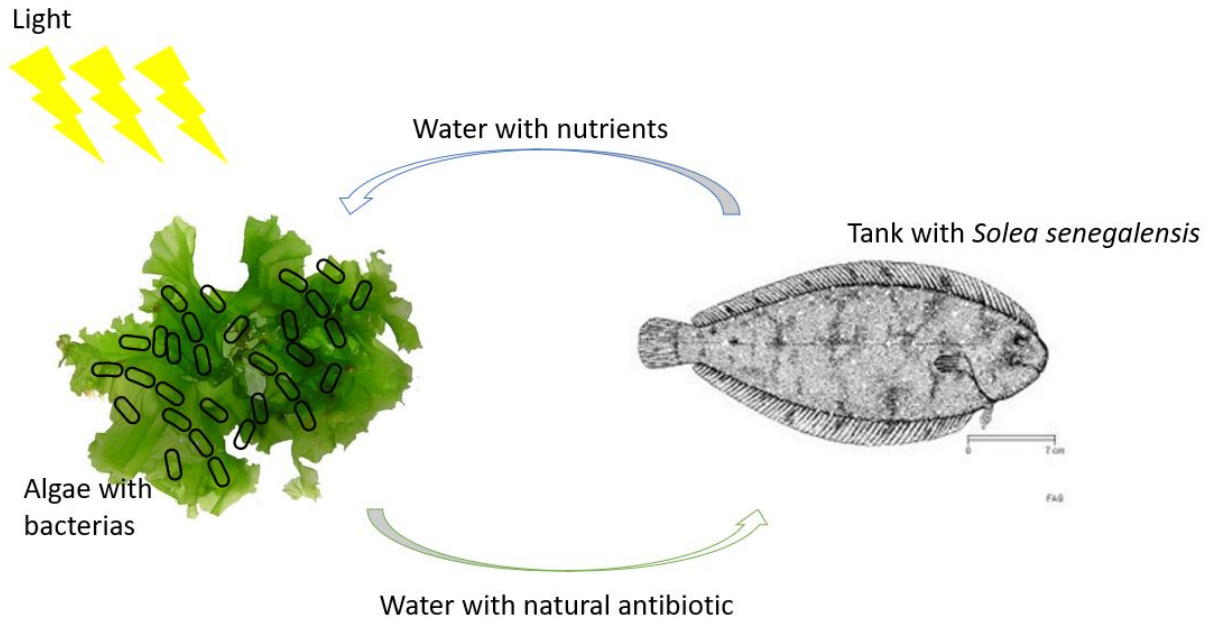


Figure 3: Schematic representation of the IMTA-RAS system considered in the present work. The system includes a tank for fish (specifically *Solea senegalensis*) and a tank with *Ulva ohnoi*—bacterial microbiome (including *Phaeobacter gallaeciensis*).

2.4.2 Effects of light on Phaeobacter gallaeciensis in Ulva ohnoi

Various factors, such as oxygen, defense and secondary metabolites, and surface competition, can influence the maintenance of bacterial communities on the outer surface of algae (Egan et al., 2013). Although there are studies that demonstrate the advantages of having *Phaeobacter* in a fish rearing system (Prol-García & Pintado, 2013), there is still no clear view on how environmental conditions in *Ulva* culture, such as agitation and high light intensities, can affect the maintenance of *Phaeobacter* on the surface of algae.

In an attempt to find an answer, Pintado, Ruiz, Cremades, & Wichard, 2022 started to study, in small-scale cultures, how different light intensities (high, medium, low, and no light) affected *Phaeobacter* (unpublished data). A decrease in *Phaeobacter* was recorded in cultures when light was present, indicating that light negatively influenced the maintenance of *Phaeobacter* on the surface of *Ulva ohnoi*. This did not occur on an inert surface (glass cover slip) (Unpublished data). Furthermore, there was no decrease in *Phaeobacter* in small-scale cultures in darkness, and these factors may demonstrate that the intensity of light can affect the maintenance of *Phaeobacter* on the surface of *Ulva*.

Therefore, it is of critical importance to design environmental conditions, particularly light intensity, to ensure that *Phaeobacter* can be maintained in the system and play its role as a probiotic. In this context, ecological models can help identify optimal operation conditions.

2.5 Ecological modeling

Mathematics models have three important functions: first, they help better understand the biological phenomena studied; second, they facilitate the design of experiments to predict certain features of the biological system, and then to be experimentally verified; and finally, they allow us to summarize and communicate the current knowledge about the system in a simple way. The development of such a model involves a series of steps, starting with the definition of its purpose and ending with the preliminary working model.

2.5.1 Model building loop

The questions to be addressed by the model will help to select the modeling framework and the information to be included in the model. Only elements that could have an impact on the questions to be addressed by the model should be included. A conceptual model may be a starting point for this.

Conceptual models are usually box- and arrow-diagrams, providing a compact visual representation of the system, allowing one to determine the key parts of the system to be considered. Boxes often represent state variables that describe the state of ecosystem components. The arrows show the relationships among state variables, such as the transfer of materials or ecological interactions.

The next step to refine the conceptual model could be based on available data: species quantities, environmental factors, role of time or space, etc. The model should strike a balance between incorporating

enough detail to capture the necessary ecological structure and processes and being simple enough to be useful in generating hypotheses and organizing available information.

The quantitative model is a set of mathematical equations with coefficients and data attached to the boxes and arrows in the conceptual model. Solving these equations allows for the prediction of state variables under particular conditions.

Ecologists use quantitative models for various purposes, including explaining existing data, formulating predictions, and guiding research. Some predictive models are empirical, meaning they represent relationships determined strictly by data. Since empirical models are not based on knowledge of the underlying mechanisms, they are more useful within the limits of the data used for their formulation. Other predictive models are more mechanistic, based on hypotheses about the particular ecological processes that cause observed behavior.

Once an initial mathematical model (or a battery of models) is formulated, the experimental data must then be used to verify the hypothesis and discriminate, if possible, among different model alternatives. The candidates will often depend on a number of unknown nonmeasurable parameters (e.g., growth rates, carrying capacities, interaction coefficients) that can be computed by means of experimental data fitting (parameter estimation). This crucial step provides the mathematical structure with the ability to reproduce a given data set, make predictions, and discriminate between different model candidates.

2.5.2 Lotka-Volterra model

When species interact, the population dynamics of each species is affected. There are three main types of interactions (Murray, 2002): (i) Predation: requires one species, the predator, to kill and eat another, the prey; (ii) Competition: the interaction of individuals competing for a common resource that is in limited supply, so the growth rate of each species is decreased; and (iii) Symbiosis: two or more species live purposefully in direct contact with each other.

Volterra, 1926, first proposed a simple model for the predation of one species by another to explain the oscillatory levels of certain fish catches in the Adriatic. If x denotes the prey population and y the predator population, the simplest set of equations that describe the system is:

$$\frac{dx}{dt} = ax - bxy \tag{2.2a}$$

$$\frac{dy}{dt} = -cy + dxy \tag{2.2b}$$

where the parameter a represents the exponential growth of prey in the absence of predators; the parameter c represents the mortality rate of predators in the absence of prey; the term xy represents the probability that a prey meets a predator when these species are uniformly distributed in their habitat and move randomly, and the ratio of the parameters b/d is analogous to the efficiency of predation, that is, the efficiency of converting a unit of prey into a unit of predator mass (Edelstein-Keshet, 2005).

The model, known as the Lotka-Volterra model, assumes: (i) the prey grows exponentially in the absence of any predation; this is the ax term; (ii) the effect of predation is to reduce the growth rate of prey per capita by a term proportional to the prey and predator populations; this is the $-bxy$ term; (iii) in the absence of any prey for food, the predator death rate results in exponential decay, that is, the $-cy$ term; (iv) the prey contribution to the predator growth rate is dx; that is, it is proportional to the available prey and the size of the predator population. This model has some drawbacks. The model considers that the two species compete for resources and sums up the role of intraspecific -within the same species- and interspecific - between different species- competitive effects. Both effects are considered to depend linearly on the population of both species. As a consequence, in the absence of a predator, the prey can grow exponentially.

Alternatives to this classical Lotka-Volterra (LV) model have become increasingly abundant in the literature (see, for example, the recent work by Gavina et al., 2018 and the works cited therein). Generalized versions of the LV model allow for more flexibility by including non-linear intraspecific and interspecific competition terms. The underlying idea is to account for higher-order interactions. Gilpin & Ayala, 1973 proposed a non-linear intraspecific competition, and others added non-linear decay terms (see the recent work by Gavina et al., 2018 for various examples).

For a general system of m species, generalized Lotka-Volterra models would read:

$$\frac{dx_i}{dt} = \mu_i x_i \left(1 - \left(\frac{x_i}{K_i}\right)^{\theta_i} - (\sum_j a_{i,j} f_c(x_j))\right) \quad i,j = 1,\ldots,m \quad i \neq j \tag{2.3}$$

where μ_i corresponds to the specific growth rate of the species i; K_i corresponds to the so-called carrying capacity of the species i in the specific medium; θ_i controls the degree of nonlinearity in intraspecific growth. The coefficients $a_{i,j}$ measure the competitive strength of species j over i. If

$a_{i,j} > 0$ species are said to be competing. $f_c(x_j)$ describe the interspecific competition, which depends linearly on x_j for the case of the Lotka-Voterra (LV) and Gilpin-Ayala (GA) models, while non-linear versions may account for crowding effects, saturation, etc.

It should be noted that the model parameters may vary with environmental factors. Therefore, secondary models must be considered (Balsa-Canto, Alonso-del Real, & Querol, 2020). In this respect, secondary models will be formulated taking into account the available experimental data and using multi-experimental parameter estimation.

2.5.3 Initial Value Problem

The generalized Lotka-Volterra models are defined as a set of ordinary differential equations, that is, an initial value problem. Therefore, its solution will depend on the initial conditions and the values of the parameters (μ_i, K_i, $a_{i,j}$). There is a huge variety of numerical methods to solve this system of equations: implicit Euler, Runge-Kutta, Adams method, Backward-differentiation-formulae (BDF), etc. The most widely used methods are implemented in the SUNDIALS suite (Hindmarsh et al., 2005a). In this work we will use CVODES, a C code implementation of a BDF method available in the SUNDIALS suite. CVODES solves the ordinary differential set and allows for the simultaneous computation of parametric sensitivities.

2.5.4 Parameter estimation

Given a set of data, the objective of parameter estimation is to compute some or all parameters $\theta = [\mu_i, K_i, a_{i,j}]$; $i, j = 1 \dots m$ in order to minimize the distance between the data and the predictions of the model. The maximum likelihood principle yields an appropriate cost function to quantify such a distance, which, for the case of Gaussian noise with known or constant variance, reads as the widely used weighted least-squares function:

$$J(\theta) = \sum_{e=1}^{n_e} \sum_{o=1}^{n_o^e} \sum_{s=1}^{n_s^{e,o}} \left[\frac{y_s^{e,o}(\theta) - ym_s^{e,o}}{\sigma_s^{e,o}} \right]^2 \tag{2.4}$$

where $\sigma_s^{e,o}$ collects the information related to a given experimental noise of the data.

Parameter estimation is then formulated as a non-linear optimization problem, where the decision variables are the parameters and the objective is to minimize $J(\theta)$ subject to the generalized Lotka-

Volterra model and, possibly, linear constraints that define the box of feasible values for the parameters.

Since the problem does not have an analytical solution, nonlinear programming (NLP) methods must be used for its solution.

2.5.5 Nonlinear programming methods

Nonlinear programming methods are designed to generate, from one or several initial guesses, a sequence of solutions that eventually converges to the minimum of the cost function. The way this sequence is generated gives rise to hundreds of different NLP solvers. A first classification of the methods would be in those able to handle nonlinear convex problems, local methods, and those able to handle nonlinear nonconvex or multimodal problems, global methods.

Direct-search methods make use of the value of the cost function in several points in the vicinity of the current iterate to generate new iterates. Probably the most popular within this category is the Nelder-Mead simplex method (Nelder & R, 1965) which uses the concept of adaptive simplex to explore the search space. An alternative, dynamic hill climbing, was proposed by (De La Maza & Yuret, 1994). This algorithm is divided into two loops: an inner loop that will find the local optimums and an outer loop that directs the inner loop to different parts of the search space in order to ensure that all the search space has been explored.

Global methods have emerged as an alternative to search for the global optimum. One of the simplest global methods is a multistart method. Here, a large amount of initial guesses are drawn from a distribution and subjected to a parameter estimation algorithm based on a local optimization approach. The smallest minimum is then regarded as the global optimum. In practice, however, there is no guarantee of arriving at the global solution and the computational effort can be quite large. These difficulties arise because it is *a priori* not clear how many random initial guesses are necessary.

Within global optimizers, stochastic and hybrid methods are possibly the most widely used in the context of parameter estimation of dynamic systems (Balsa-Canto, Alonso, & Banga, 2010). Stochastic global optimization algorithms make use of pseudorandom sequences to determine search directions toward the global optimum. This leads to an increasing probability of finding the global optimum during the runtime of the algorithm. The main advantage of these methods is that they rapidly arrive at the proximity of the solution. The number of stochastic methods has rapidly increased in the last decades. The most

successful approaches lie in one (or more) of the following groups: pure random search and adaptive sequential methods, clustering methods, population based methods or nature inspired methods (Dréo J & P, 2006).

Despite the fact that many stochastic methods can locate the vicinity of global solutions very quickly, the computational cost associated with the refinement of the solution is usually very large. In order to surmount this difficulty, hybrid methods and metaheuristics have been recently presented for the solution of parameter estimation problems that speed up these methodologies while retaining their robustness. In particular, the enhanced Scatter Search metaheuristic (eSS) (Egea, Balsa-Canto, Garcia, & Banga, 2009) showed speeds between one and two orders of magnitude with respect to the use of stochastic global methods.

eSS is a population-based approach that combines global exploration with calls to a local method. As implemented in AMIGO2 tooldbox (Balsa-Canto, Henriques, Gabor, & Banga, 2016), eSS allows to use several local optimizers. In this work, we have used DHC and Nelder-Mead methods. Note that Nelder-Mead is available in MATLAB as *fminsearch*.

2.6 Software

In this work, we have used several software tools: FIJI (Schindelin et al., 2012), GenSSI2 (Ligon et al., 2018) and AMIGO2 (Balsa-Canto et al., 2016). In this section, the tools are briefly introduced.

2.6.1 FIJI

Fiji is the distributor of the open source software ImageJ. This software is focused on biological image analysis and uses modern software engineering to update the underlying architecture of ImageJ, making it more attractive to biologists, bioinformaticians and computer scientists in the process of developing science (Schindelin et al., 2012).

- **Biology researchers.** They can input their data and interact with multidimensional image data.
- **Bioinformaticians.** Using the Script Editor plugin, bioinformaticians can develop pipelines for image processing suitable for what is to be analyzed.

- **Software engineers.** Can use capable software libraries to transform the mathematical formulations of computer science algorithms into functional programs, thus allowing the management of existing code and the implementation of new algorithms.

2.6.2 GenSSI2

GenSSI2 (Generating Series for testing Structural Identifiability) is a software toolboox that allows performing structural identifiability analysis on biological models. In other words, the tool is devoted to analyzing whether the parameter estimation problem will have a unique solution for a given model. Obviously, the structural identifiability of any candidate model should be checked at the beginning of the model-building loop. However, checking this property for arbitrary nonlinear dynamic models is not an easy task. GenSSI2 is an update of GENSSI (Chiş, Banga, & Balsa-Canto, 2011) and enables nonexpert users to perform such an analysis. The toolbox runs under the popular Matlab environment and is accompanied by detailed documentation and relevant examples.

2.6.3 AMIGO2 toolbox

AMIGO2 (Balsa-Canto et al., 2016) is a MATLAB based toolbox devoted to the identification and optimization of general dynamic systems.

The software has the following capabilities:

- **Model.** AMIGO supports general nonlinear ODE models using a simple syntax, C++ or MATLAB. Allows to import sbml and black-box user-defined models.
- **Experimental scheme.** The tool allows flexible experimental schemes -one or more experiments, input profiles, initial conditions, experiment durations, and sampling times- that are to be performed *in silico*.
- **Experimental data.** The tool allows introducing or loading real experimental data with different types of experimental noise, homoscedastic or heteroscedastic. In addition, pseudo-experimental data can be generated for numerical tests.
- **Parameter estimation.** The tool allows multi-experiment fitting with local (experiment dependent) and global unknowns (parameters and initial conditions). Several types of cost

function, weighted least squares or log-likelihood, may be used depending on the available information about the experimental noise.

- **Practical identifiability analysis.** Computes local sensitivities, the correlation matrix from the Fisher information matrix depending on the experimental noise conditions, cost function contour plots for pairs of unknowns, and the robust Monte Carlo-based approach.

- **Optimal experimental design.** Solves the D-, E-, Modified E or A- optimal experimental design problem as a general open-loop optimal control problem allowing sequential and parallel designs. It is possible to optimize the sampling times, input conditions, duration of the experiment, and initial conditions for one or more simultaneous experiments. Several Fisher matrix formulations are available depending on the experimental noise.

Numerical methods. It incorporates several state–of–the–art initial value problem (IVP) and nonlinear optimization (NLP) methods to deal with both parameter estimation and experimental design problems. Regarding IVP solvers, explicit and implicit Runge-Kutta, Adams, and BDF methods have been incorporated together with methods to compute sensitivities. Concerning NLP solvers, several direct and indirect local methods, multistart of local methods, global stochastic, and hybrid optimization methods are available. Computational demanding tasks are interfaced with C++ compiled code.

Reporting. Generates reports and plots according to user specifications for different tasks. The complete working session is saved in a Matlab structure and may be reloaded at any time.

3 Materials and Methods

3.1 Experiment

3.1.1 Ulva ohnoi culture

Ulva ohnoi samples were obtained from a single clone culture from the laboratory of Professor Javier Cremades at the University of A Coruña (Spain), and were processed into discs each with 2 cm in diameter. Individually, the discs were placed in six multiwell plates (BioLite 6 well multidish, THERMO SCIENTIFIC, United States) filled with 10 mL *Ulva* culture medium (UCM) per well.

The UCM contains 33 g/L sea salt (Instant Ocean, AQUARIUM SYSTEMS, United Kingdom), 0.1 g/L Guillardś F/2 medium (Cell-Hi F2P, VARICON AQUA, United Kingdom) supplemented with NO_3Na to a final level of 20 mgN/L (MERK, Germany) and pH adjusted to 8. The UCM medium was sterilized using 0.22 μm membrane filtration.

3.1.2 Phaeobacter sp. 4UAC3 culturing and quantification

The probiotic *Phaeobacter sp.* strain 4UAC3 was previously isolated from *Ulva rigida* in Galicia, Spain, and was used as the probiotic candidate due to previous antagonism activity against *V. anguillarum* (unpublished data). *Phaeobacter sp.* was stored at -80°C in Marine Broth (MB, Marine Broth 2216, BD-DIFCO, Spain) and 15% glycerol (Glicerina, VORQUÍMICA S.L., Spain). The bacteria were reactivated following the protocol described by Prol, Bruhn, Pintado, & Gram, 2009. This protocol follows the following steps:

- First, we start reactivating *Phaeobacter* in 5 ml of Marine Broth for 72 hours in darkness at 20°C;
- Then we inoculate 1 mL of preculture in 100 mL MB (Marine Broth, CONDALAB, Spain);
- After this, we leave the bacteria incubating for 48 hours in darkness at 20°C under stagnant conditions.

The purity and concentration of the respective cultures was checked with Marine Agar (MB with 15 g/L industrial agar, CONDALAB, Spain) in 90 mm Petri dishes (90 mm Petri dish triple vent, THERMO SCIENTIFIC, United States). The medium cultures were previously sterilized in an autoclave at 121°C for 20 min.

The colonies of *Phaeobacter sp.* clearly differ morphologically from other marine bacteria. This happens because nutrient-rich iron-containing medium causes precipitation of the TDA-iron complex, making these bacteria distinguishable by their brown coloration on the plates (D'Alvise, Phippen, Nielsen, & Gram, 2016).

3.1.3 Experimental design

Ulva ohnoi was cultured in multiwell plates to determine the optimal light intensity for *Phaeobacter* to remain on the surface of the algae, also taking into account the growth of the algae itself. All multi-well plates containing *U. ohnoi* discs were placed in a climate controlled culture chamber (Medline, LIEBHERR, Switzerland) at 20°C under a daylight LED panel (44 W – 6000K) with a light intensity of approximately 45 $\mu m\ m^{-2}s^{-1}$ in a dark: light photoperiod of 12:12h and orbital agitation (Orbital Shaker DOS-20L, ELMI, United States) of 80rpm for a four-day acclimatization period (Figure 5). After acclimatization, the plates were randomly distributed and placed under four different light intensities: high light (HL, 224 $\mu m\ m^{-2}s^{-1}$), medium light (ML, 120 $\mu m\ m^{-2}s^{-1}$), low light (LL, 60 $\mu m\ m^{-2}s^{-1}$) and darkness (DL, 0 $\mu m\ m^{-2}s^{-1}$). On day 6 (T6), the UCM of each well was renewed with 10 mL fresh UCM and inoculated with 100 μL of a 48-hour culture of *Phaeobacter sp.* (explained the procedure in the previous subsection) (concentration: $6.9 \times 10^8 CFU/mL$), giving a concentration of $6.9 \times 10^7 CFU/mL$ in the UCM. The discs were kept under different light conditions for a total of 14 days (Figure 4). On T0, T4, T6, T8, T11, T13, T15 and T18, four *U. ohnoi* discs of each light treatment were randomly sampled for microbial analysis of the *U. ohnoi* surface and for growth measurements. The UCM of the wells sampled was taken for physicochemical analysis.

3.1.4 Microbial analysis

Algae surface biofilm sampling was performed using the swab-rinse technique (Favero, McDade, Robertsen, Hoffman, & Edwards, 1968). The algae discs were rubbed with a sterile cotton swab with an applicator stick for 1 min in a laminar flow hood (Heraeus LaminAir TL 2448, Heraeus SA, Germany). The swab was then immersed in a microcentrifuge tube containing 1 mL sterile seawater, the applicator stick was cut off, and 2 min of vortex step was applied. Subsequently, the cotton swab was removed and serial dilutions were performed to count the colony forming units of the sampled biofilm. 100 μL were taken and diluted in 900 μL sterile seawater. This step was repeated as many times as necessary so that 3 days after plating the dilution on Marine Agar culture medium plate, a number of colonies between

30 and 300 CFU could be achieved. The remaining 900 μL of the initial solution was centrifuged at 14000rpm for 20 min at 4 °C to retain the resulting pellet and freeze it at -20 °C for future molecular analysis (qPCR and 16 rRNA gene sequencing).

3.1.5 Algae growth parameters

After sampling the algal biofilm, the algal disc was weighed on a precision balance (SCALTEC SBA 32, SCALTEC, Spain) to obtain the wet weight. Subsequently, the algal discs were placed in plastic bags and scanned for surface measurement using image processing with FIJI software (Schindelin et al., 2012). For image processing, the image to be analyzed was first calibrated (the scanned images were scanned with a millimeter ruler), transformed into a binary image, and the area was measured with the "Analyze Particles" function.

3.1.6 Physico−chemical analysis of water

The pH of the aqueous medium was measured with a pH meter (Sension+ PH3, HACH, United States). Nitrates were measured following the method described by Rice, Baird, Eaton, Clesceri, et al., 2012. Briefly, this measurement method is based on the absorbance of the nitrate ion in an aqueous sample, previously filtered through a 0.45 μm filter, at 220nm and taking into account the absorbance error observed at 275nm due to the presence of organic matter. Finally, the phosphates were measured using the method described in Grasshoff, Kremling, & Anderson, 1999. This method is based on a color reaction produced by ascorbic acid when it reduces a complex formed by ammonium molybdate, antimony potassium tartrate, and soluble phosphorus of the solution in an acidic environment. This reductive reaction produces a blue color proportional to the phosphorus concentration of the sample, achieving a maximum absorbance value of 880nm. For both nitrate and phosphate measurements, a UV spectrophotometer (UV-1800, SHIMADZU, Japan) was used.

3.2 Computational modeling

3.2.1 Structural identifiability analysis

The structural identifiability problem relates to the possibility of finding a unique solution for the parameters with perfect data (noise-free and continuous in time). Structural identifiability is therefore associated with

the model equations and possibly the type of experiments, but it is independent of the parameter values.

In the case of Lotka-Volterra models, there are two obvious reasons for assessing structural identifiability: first, model parameters have biological meaning; second, lack of identifiability would lead to a wrong map of interactions.

There is no universally valid method for the analysis of structural identifiability of general nonlinear dynamic models; but a combination of the generating series approach with identifiability *tableaus* is suitable for nonlinear dynamic models in systems biology (Chis, Banga, & Balsa-Canto, 2011). This method is implemented in the GenSSI2 toolbox (Ligon et al., 2018), which is used in this work.

The underlying idea of the generating series approach is that observables, algae and microbial abundances, can be expanded in series with respect to time and inputs around a given time point (t_0), and that the uniqueness of the series coefficients guarantees the structural identifiability of the model. The series coefficients are computed by means of successive Lie derivatives of the observation function. The identifiability *tableaus* correspond to the Jacobian of the Lie derivatives with respect to the model parameters and help to decide the global or local structural identifiability of the model (Balsa-Canto et al., 2010; Chis et al., 2011).

If the Jacobian has full rank, then the model will be at least locally identifiable. If not, i.e. the tableau presents empty columns, the corresponding parameters may be unidentifiable. Note that since the number of series coefficients may be infinite, unidentifiability may not be fully guaranteed unless higher-order series coefficients are demonstrated to be zero.

3.2.2 Parameter estimation

Parameter estimation is used to calculate the value of model parameters: specific growth rates, carrying capacities, and interaction coefficients, which minimize the distance between the experimental data and the predictions of the model. This distance is usually quantified using the weighted least-squares cost function:

$$J() = \sum_{e=1}^{n_e} \sum_{s=1}^{n_s^{e,i}} \left[x_s^{e,i}() - xm_s^{e,i} \right]^2 \tag{3.5}$$

where $xm^{e,i}$ corresponds to the vector of n_s data on the abundance of species i in the experiment e; $x^{e,i}$ corresponds to the model predictions for each sampling time s.

Parameter estimation is then formulated as a nonlinear optimization problem, where the decision variables are the parameters, and the objective is to minimize $J()$ subject to the system dynamics, the

Lotka-Volterra equations, and algebraic constraints that define the feasible region for the parameters.

To automate the estimation of parameters and subsequent analyzes performed in this work, we use the AMIGO2 toolbox (Balsa-Canto et al., 2016). To solve the model, we used a method of the backward differentiation formula, CVODES (Hindmarsh et al., 2005b). The global optimizer *Enhanced Scatter Search* (eSS, (Egea et al., 2009)) was used to find the optimal parameter values in reasonable computational time.

3.2.3 AMIGO2 script

To build the mathematical models, two scripts[1] were programmed in MATLAB for each light condition: the first incorporates all biological data, i.e., number of experiments, experimental conditions (initial and stimulation conditions), number of sampling times, sampling times, duration of experiment, experimental data and the experimental error as computed from experimental replicates. Table 1 summarizes the information required in the AMIGO2 toolbox:

Table 1: Inputs required by AMIGO2 to introduce biological data for each experiment.

inputs structure element	Meaning
inputs.exps.n_exp	Number of experiments or replicas
inputs.exps.n_obs{}	Number of observed quantities per experiment
inputs.exps.obs_names{}	Name of each observable
inputs.exps.obs{}	Observation equations
inputs.exps.u_interp{}='sustained'	Stimuli profile
inputs.exps.exp_y0{}	Initial conditions for each experiment
inputs.exps.n_s{}	Number of sampling times
inputs.exps.t_s{}	Sampling times. By default this value is equidistant
inputs.exps.t_f{}	Experiment duration
inputs.exps.t_con{}	Initial and final time
inputs.exps.data_type	Type of data. In this model is $'real'$.
inputs.exps.noise_type	Experimental noise type. By default is $'homo\_var'$.
inputs.exps.exp_data{}	Values of the experimental data
inputs.exps.error_data{}	Values of the experimental error

[1]Scripts and biological data can be found in: https://github.com/julianapereira99/Ecological_modelling_light_Ulva

Table 2a: Inputs required by AMIGO2 to formulate the model and the parameter estimation problem.

	inputs structure element	Meaning
Path	inputs.pathd.results_folder	Name of the folder to keep results in Results.
	inputs.pathd.short_name	Short name to identify figures and reports.
	inputs.pathd.runident	Identifier required not to overwrite previous results.
Model	inputs.model.input_model_type	Model description type. In this case we used $'charmodelC'$.
	inputs.model.n_st	Number of states.
	inputs.model.n_par	Number of parameters.
	inputs.model.n_stimulus	Number of stimuli.
	inputs.model.names_type	Type of names used to define the model. By default is used $'custom'$.
	inputs.model.st_names	Names of the states.
	inputs.model.par_names	Names of model parameters.
	inputs.model.eqns	Model equations.
	inputs.model.par	Nominal value for model parameters. If known, parameters are fixed to their nominal values; otherwise their values will be updated by the optimiser.
Model Unknowns	inputs.PEsol.id_global_theta	Parameters to be estimated
	inputs.PEsol.global_theta_min	Lower bounds for estimated parameters.
	inputs.PEsol.global_theta_max	Upper bounds for estimated parameters.
	inputs.PEsol.global_theta_guess	Initial guess for optimisation.
Cost function	inputs.PEsol.PEcost_type	'lsq' for least squares or 'llk' for log-likelihood
	inputs.PEsol.lsq_type	Least squares normalisation. Q_I for identity matrix.

Table 2b: Inputs required by AMIGO2 to to numerically solve the parameter estimation problem.

		inputs structure element	Meaning
Numerical Methods	Simulation	inputs.ivpsol.ivpsolver	Initial value problem (IVP) solver algorithm. By default $'cvodes'$ is used for C.
		inputs.ivpsol.senssolver	Sensitivity solver. By default $'cvodes'$ is used for C.
		inputs.ivpsol.rtol	Relative IVP tolerance.
		inputs.ivpsol.atol	Absolute IVP tolerance.
		inputs.ivpsol.ivp_maxnumsteps	Maximum steps allowed for the IVP solver.
	Optimization	inputs.nlpsol.nlpsolver	Non-linear programming (NLP) solver. By default is used the metaheuristic solver $'ess'$.
		inputs.nlpsol.eSS.log_var	Indexes of the parameters to be considered in log scale.
		inputs.nlpsol.eSS.maxeval	Maximum number of cost function evaluations.
		inputs.nlpsol.eSS.maxtime	Maximum CPU time invested (in seconds).
		inputs.nlpsol.eSS.local.solver	Local solver method used ($'dhc'$).
		inputs.nlpsol.eSS.local.finish	Last call local method ($'fminsearch'$).

Table 2c: Inputs required by AMIGO2 to select display options.

	inputs structure element	Meaning
Display of Results	inputs.plotd.plotlevel	Regards the number of analyses and figures shown. 'min' was selected.
	inputs.plotd.figsave	Select (or not) saving figures.
	inputs.plotd.n_t_plot	Time steps for smooth simulation.

The second script includes the definition of the model states, parameters, equations and formulates the parameter estimation problem (parameters to be estimated from data, parameter estimation cost function) and the numerical methods to be used for simulation and optimisation. Table 2a presents the inputs required by AMIGO2 to formulate the model and the parameter estimation problem; Table 2b presents the inputs required to formulate the numerical solution to the problem and Table 2c presents the display options offered in AMIGO2.

3.2.4 Command Window

The scripts described above are used to define the MATLAB structure $input$, which can then be used to perform several tasks in AMIGO2::

- **AMIGO_Startup.** Function used to open the AMIGO2 tool.
- **AMIGO_Prep.** Pre-processing function. This function interprets the $inputs$ structure, creating the necessary files for the following tasks. Depending on the type of model, functions that will be used for model simulation and sensitivity analysis are generated. For example, in the particular case of using $'charmodelC'$ C code is generated and mexed so as to be automatically called for simulation. It should be noted that the use of the C code greatly improves the efficiency of the parameter estimation problem.
- **AMIGO_SModel.** This function is used to simulate the model for given experimental conditions and parameter values. The figures will show the dynamics of all states in the model, regardless of whether they can be measured.
- **AMIGO_SData.** This function simulates the model for the given experimental conditions and parameter values and represents model simulations against experimental data, i.e. only measured states will be shown.
- **AMIGO_PE.** This function is devoted to parameter estimation. By formulating a non-linear programming problem, this function will find the unknown parameters in the model through data fitting, i.e. by minimizing the distance between the model predictions and the experimental data. In this work, the least squares difference was considered.

4 Results

4.1 Data analysis: Data set 1

The first data set (Data set 1) included the size (surface) and weight of the algae, the bacterial microbiome and *Phaebacter* over time. *Phaebacter* was incorporated into the experimental system after the second sampling time. Values were reported at seven sampling times. As a first step, we adapted the data to modeling: computed mean and standard deviation of replicates and removed outliers. The generalized Lotka-Volterra model was implemented in AMIGO2 taking into account the late incorporation of *Phaeobacter*. Estimation of parameters revealed serious problems. We observed a high correlation between the parameters, large confidence intervals and extremely large carrying capacity values were required to reproduce the experimental data. We decided to revisit the experimental data and explore the identifiability properties of the model.

In terms of experimental data, we realized that the medium was being renewed every three days. This was a serious drawback for Lotka-Volterra modeling. Note that the addition of nutrients to the medium implies an increase in the carrying capacity; while the carrying capacity K_i is a constant in the model. In principle, one could consider including a time-dependent carrying capacity to be modified each time nutrients were added to the medium, but the model was poorly identifiable and the incorporation of additional degrees of freedom was disregarded.

The analysis revealed the need to perform new experiments.

4.2 Experimental design: Data set 2

This experiment was designed to be compatible with the Lotka-Volterra model. Therefore, only culture medium was added at the beginning of the experiment and there was a change in medium when *Phaeobacter* was added to the medium. The experimental design can be seen in Figure 4. The whole experiment is documented in detail in subsection 3.1.

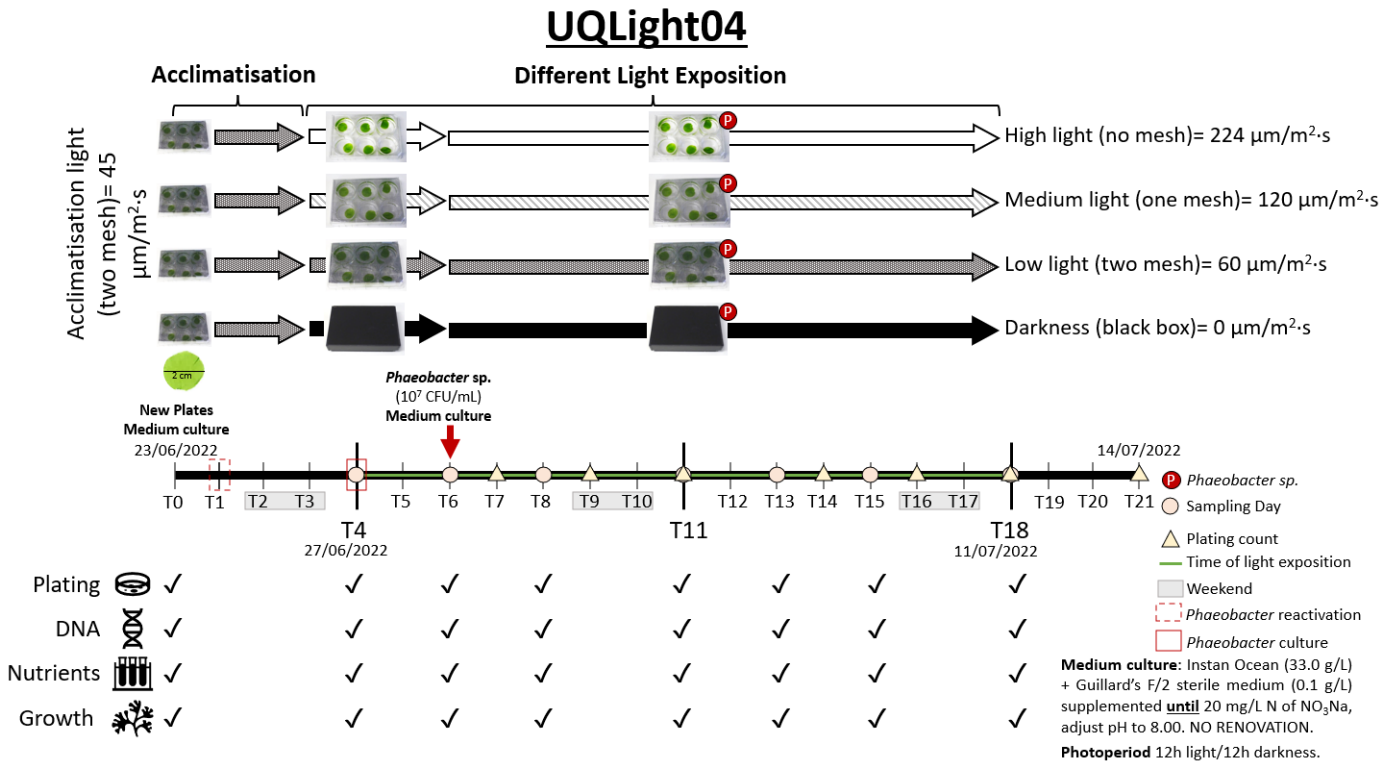


Figure 4: Scheme representation of the experimental design described in 3.1. UQLight04 was a experiment carried out to study the effect of light intensity on the colonisation and maintenance of *Phaeobacter* on *Ulva onhoi*. In this study, 6-well plates and algal discs with an initial size of 2 cm in diameter were used. This figure shows the different light conditions (High Light, H: 224 $\mu m\ m^{-2}s^{-1}$, Meddium Light, M: 120 $\mu m\ m^{-2}s^{-1}$, Low Light, L: 60 $\mu m\ m^{-2}s^{-1}$, Darkness, D: 0 $\mu m\ m^{-2}s^{-1}$) used after the acclimatisation period (45 $\mu m\ m^{-2}s^{-1}$) as well as the time of addition of the probiotic bacteria (circle with P). In addition, the chronogram of the experiment is included, showing the different sampling days (circle), the days of plate counting (triangle), the days with exposure to different light intensities (green line), the days of re-activation (dotted rectangle), culture (rectangle) and inoculation of *Phaeobacter* (arrow) and the different analyses carried out (microbiological: plating and DNA sampling, nutrient analysis, and algae growth). Finally, this figure legend also includes the composition of the medium used for the algae culture, and the photoperiod used (12h:12h).

The experiment started on 23 June 2022 (T0). All plates with algae discs were prepared (process described in Subsection 3.1.1) and placed to acclimate in a refrigerator at a temperature of 20°C (Figure 5). These multi-well plates were identified with the designated light intensity (high (H), medium (M), low (L) and darkness (D)) and replicates (A, B, C and D).

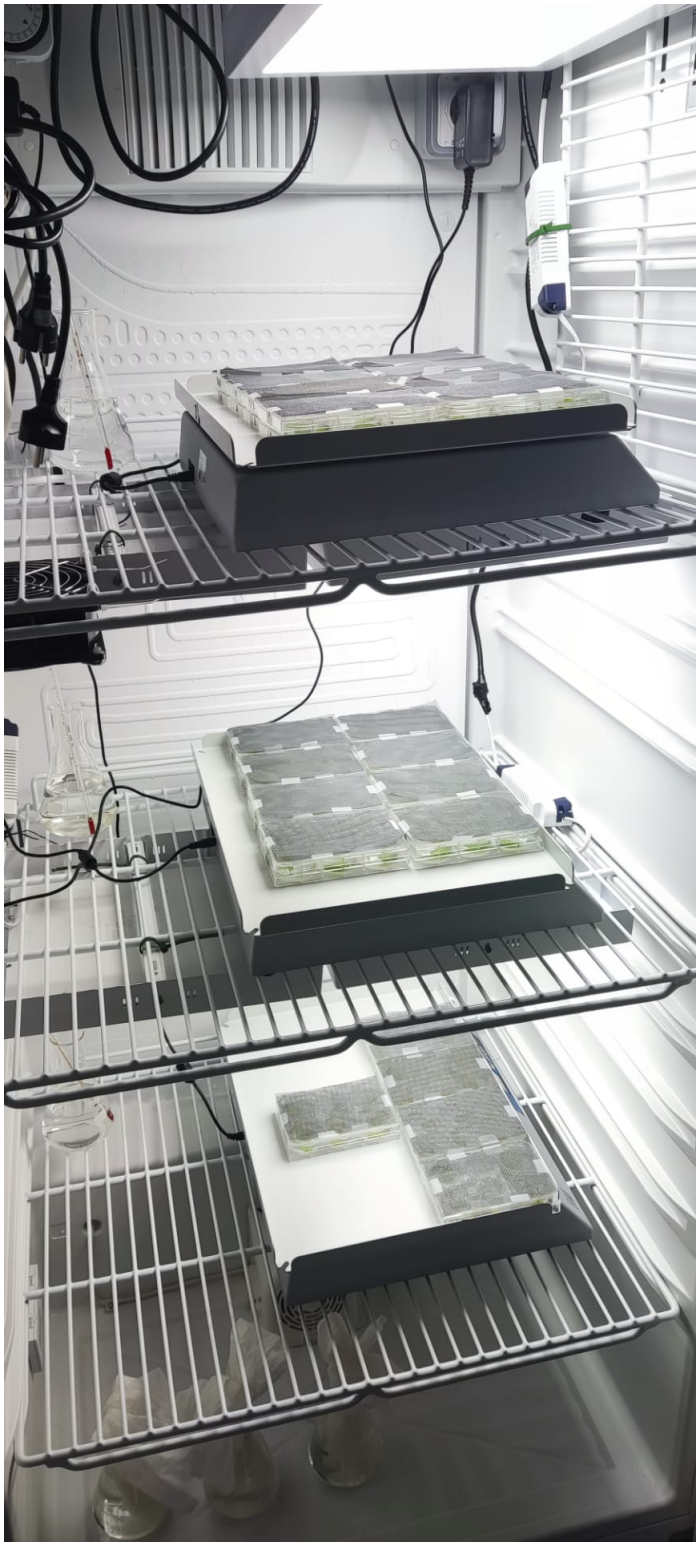


Figure 5: Refrigerator at 20°C where the algae are inside their respective plates to acclimatise.

The samples were collected in:

- **Eppendorfs** used to keep microbial communities swabbed from the surface of algae, as well as to prepare serial dilutions with sterile seawater (SSW) for microbial analysis. They were identified with the T (day), light intensity, and the corresponding replicate and dilution.
- **Algae bags** used to store algae discs in a freezer at -20°C for further image analysis. They were identified with T (day), light intensity, and corresponding replica, date, and name of the experience (in this case UQLight04).
- **Petri dishes with marine agar** used to grow microbial colonies in an incubator. They were identified with T (day), light intensity, and the corresponding replicate, date, and dilution.

The sampling was then initiated, where, in each T (day), 1 algae was swabbed from each multi-well plate for 1 min to collect the bacteria on the algal surface. Subsequently, this swab was placed in an eppendorf tube with $1mL$ of SSW and vortexed, with the aim of obtaining resuspension of the associated algae microbiota. Serial dilutions were performed as many times as necessary, so that after plating the dilution on a marine Agar plate, a number of colonies between 30 and 300 CFU could be achieved 3 days after plating (Figure 6). As for the algae, after being swabbed, they were dried with paper and then weighed. Subsequently, they were placed in bags and scanned with a ruler at the bottom for later measurement using Fiji software.

4.3 Algae growth

Figure 7 shows the growth of the algae during the experiment under different conditions of light intensity. The inspection of the figure shows the following:

- The growth of algae is greater at higher light intensities;
- Algae at high and medium light intensities were thinner and more fragile than algae without light. The brighter the algae, the thinner and fragile it is.
- No growth is perceptible when the algae are deprived of light.

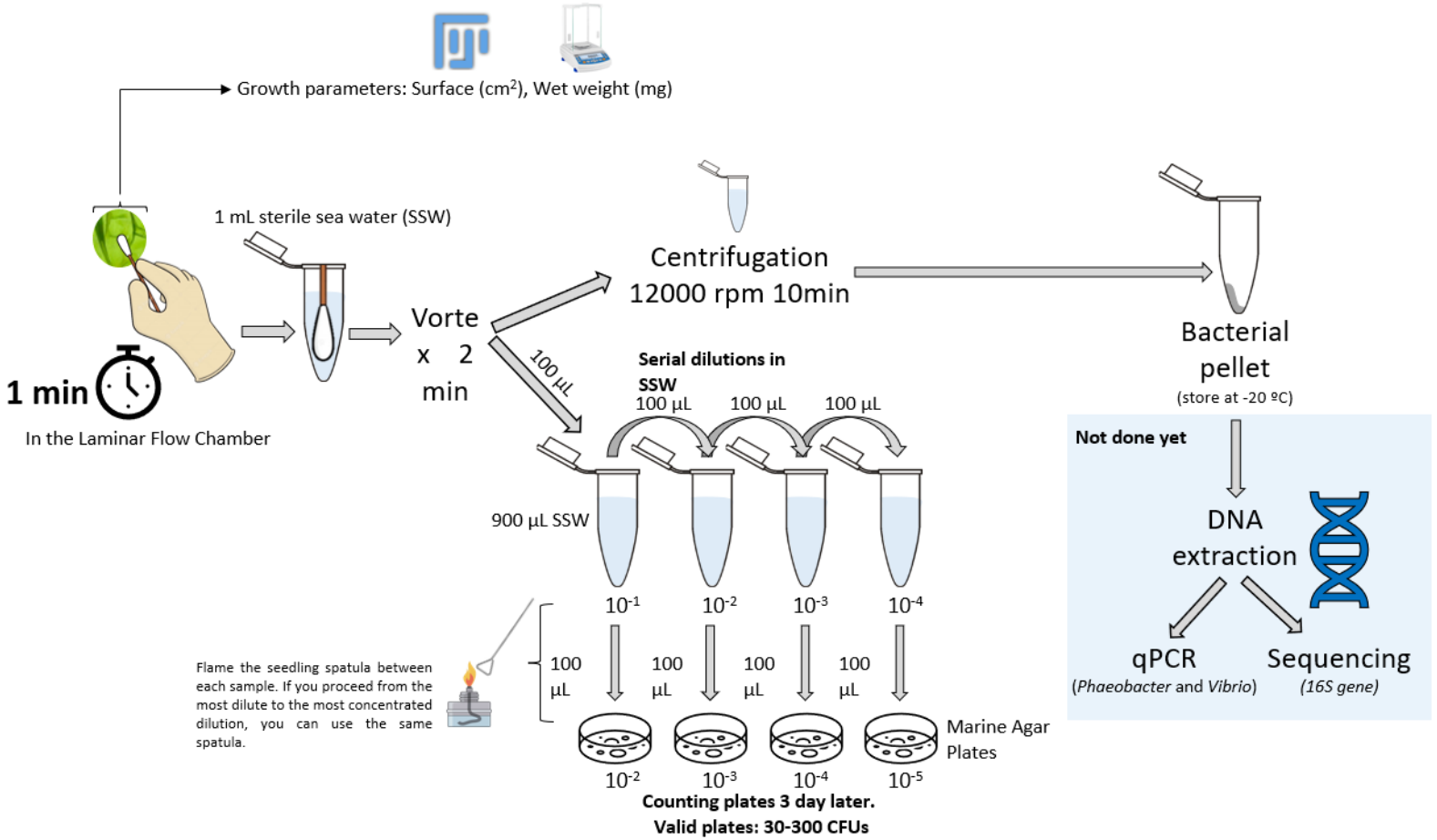


Figure 6: Schematic summary of the treatment of the algae discs on each sampling day. After rubbing the surface of the algae for 1 min, the swab used was immersed in 1 mL of sterile seawater and vortexed for 2 min to resuspend the collected biofilm. Then, 100 μL of the suspension was used to make serial dilutions which were plated on Marine agar for subsequent colony plating. The rest of the suspension was centrifuged to obtain the pellet from which DNA will be extracted for qPCRs and sequencing. The algal disc was weighed with a precision scale and scanned to measure its surface using FijI software. Both bacterial pellet and algae were stored in a -20 °C freezer.

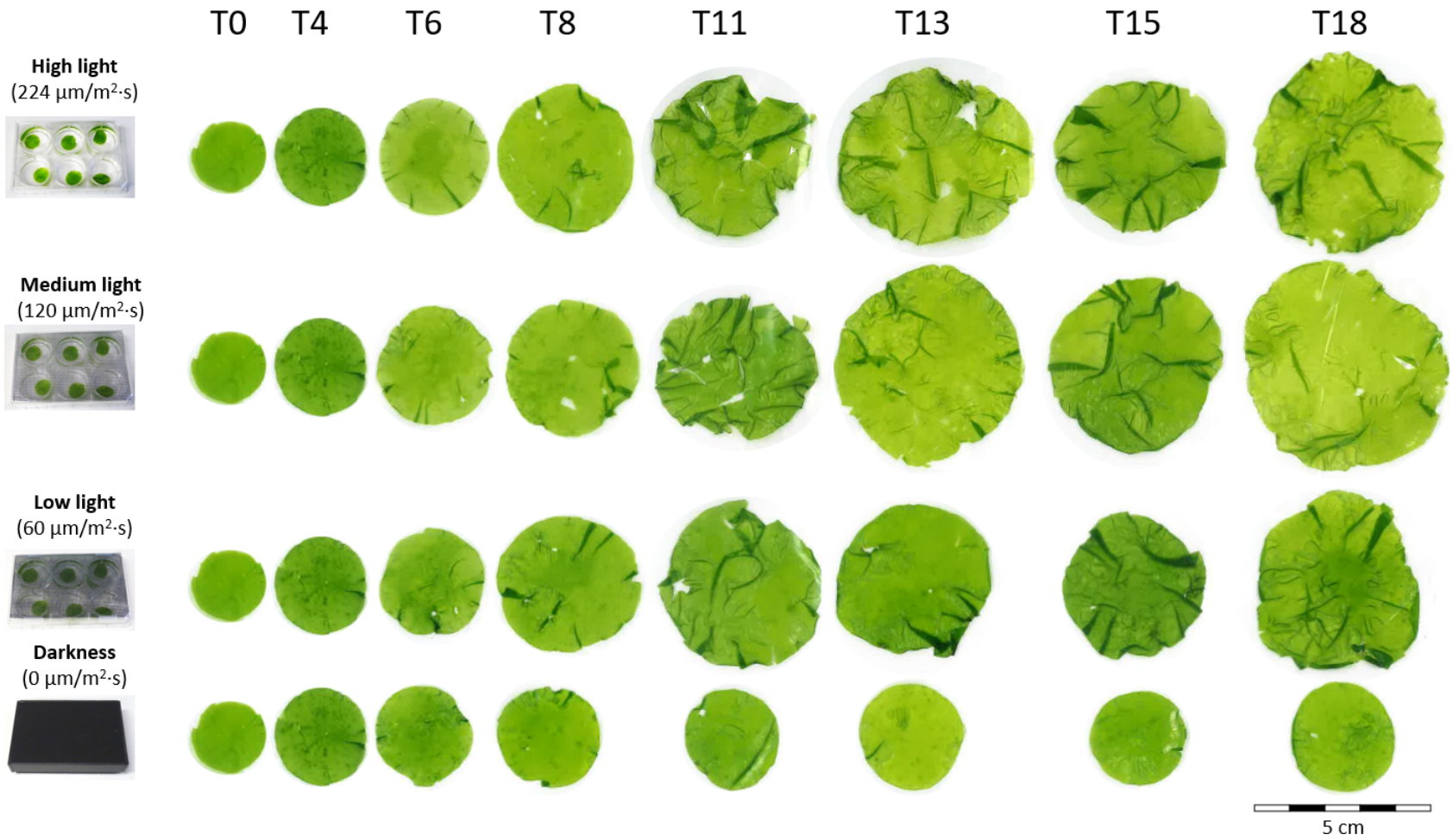


Figure 7: Demonstrates the algae growth across an 18-day culture period under different light intensities. Photosynthetic photon flux densities (PPFD): High = 224 $\mu m\ m^{-2}s^{-1}$, Medium = 120 $\mu m\ m^{-2}s^{-1}$, Low = 60 $\mu m\ m^{-2}s^{-1}$ and Darkness = 0 $\mu m\ m^{-2}s^{-1}$.

Figure 8 shows algal discs exposed to different light intensities, with the exception of condition D, algae experienced remarkable growth throughout the experiment. It is also shown that in the presence of light and without adding nutrients, between T11 and T15 algae stop growing and after T15 the surface starts to grow again, demonstrating a secondary growth. The wet weight also reveals that the weight of the algae tends to follow this trend. Note that this behavior was not observed at high light.

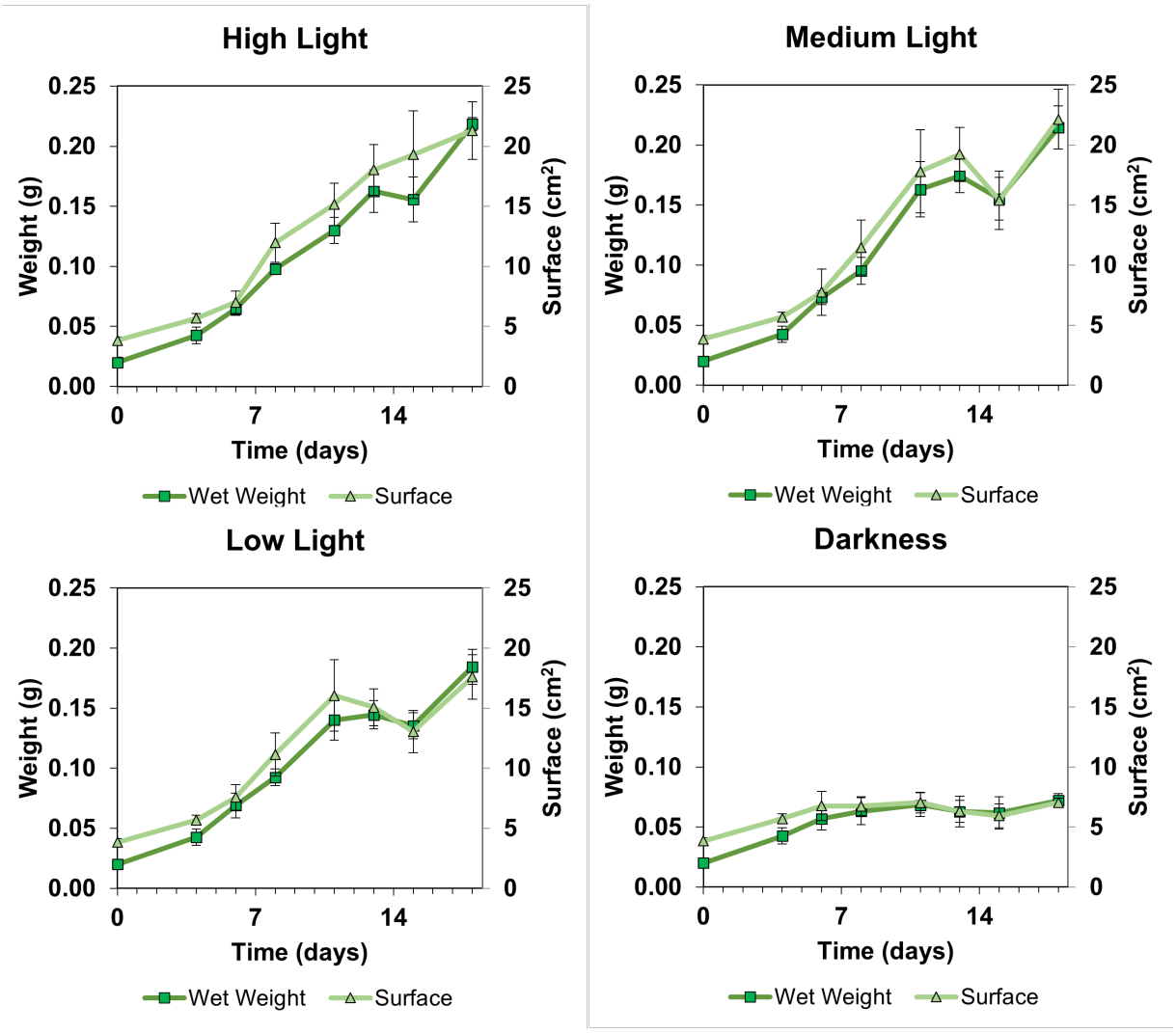


Figure 8: Wet Weight and Surface of *Ulva ohnoi* across an 18-day culture period under different light intensities. High Light = 224 $\mu m\ m^{-2}s^{-1}$, Medium Light = 120 $\mu m\ m^{-2}s^{-1}$, Low Light = 60 $\mu m\ m^{-2}s^{-1}$ and Darkness = 0 $\mu m\ m^{-2}s^{-1}$. Each point represents an average (n = 4) of the surface and wet weight of the algae in different sampling days and the error bar represents the standard deviation.

4.4 Microbial analysis

Figure 9 shows that during the acclimation period, the total bacteria increased significantly (about 3 log(CFU/cm2)) under all conditions. In T6, *Phaeobacter* was added, and in T8 the first reading was taken after this addition. The maximum total bacteria peak was reached in T8 under all light conditions. Figure 9 shows that the value of *Phaeobacter* is very close to the value of total bacteria. This would mean that after the addition of *Phaeobacter*, there was a decrease in all other bacteria on the algae surface, together with a significant growth in the concentration of *Phaeobacter*. After this maximum peak at T8, the following was observed:

- In High PPFD, after T8, *Phaeobacter* decreases, and the total bacteria almost stabilizes. This indicates that the environmental conditions on the *Ulva* surface are not suitable for the maintenance of *Phaeobacter*, which loses its ecological niche.
- In Medium PPFD, after T8, *Phaeobacter* shows almost the exact same behavior as in HL condition. However, there is a stabilization between T13 and T15. After that, *Phaeobacter* starts to decrease again. This decrease after T15 matches the secondary growth of the algae observed in Figure 8.
- In Low PPFD, after T8, *Phaeobacter* also decreases, except between T13 and T15.
- In darkness, after T8, *Phaeobacter* also decreases. However, after T13 starts to increase. This would indicate that darkness facilitates *Phaeobacter* to maintain its ecological niche on the surface of *Ulva*.

For all conditions, after T8 there is little variability in the value of the total number of bacteria. However, the highest variability occurs under High PPFD conditions.

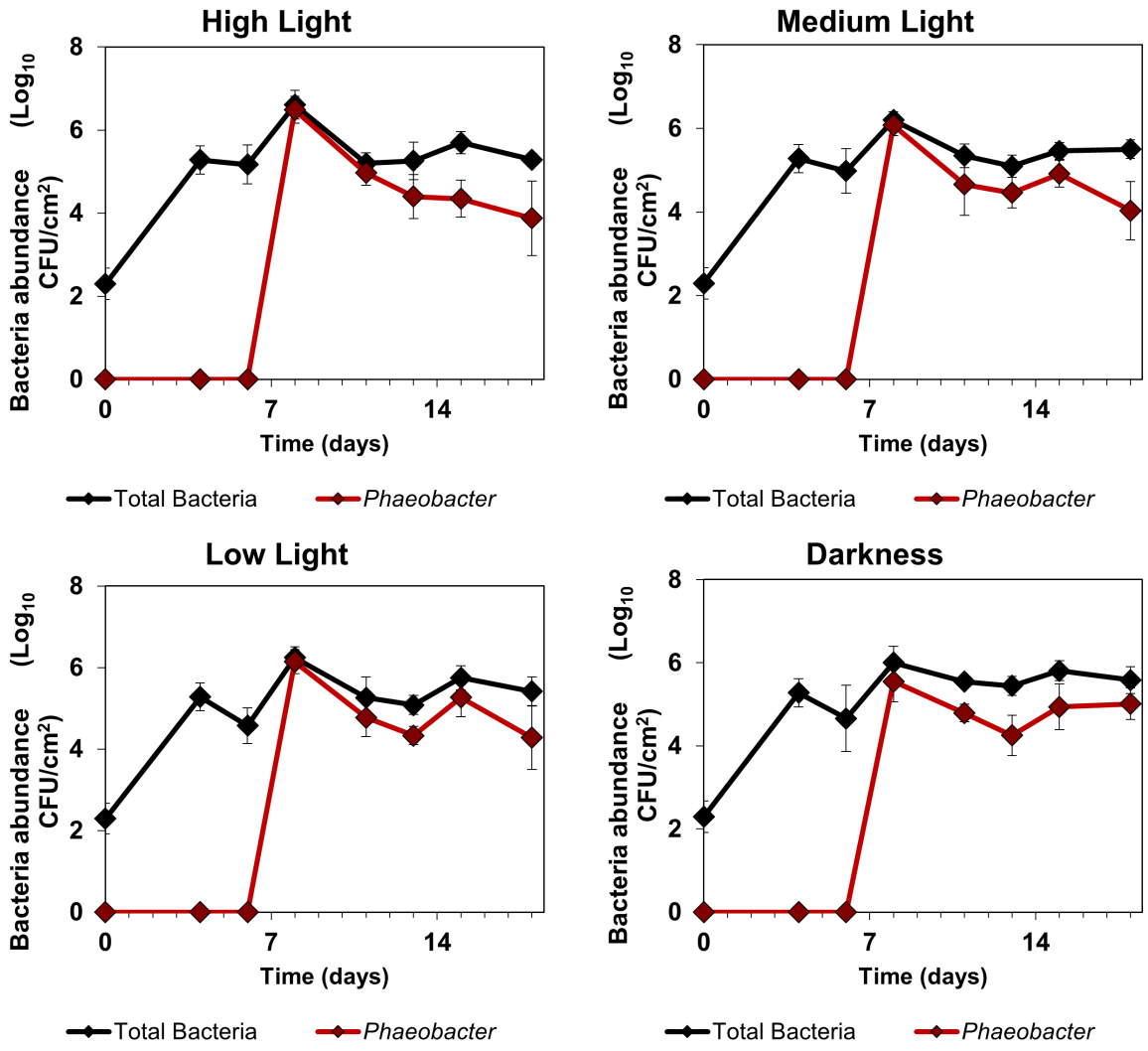


Figure 9: Total bacteria and *Phaeobacter* on the surface of *U. ohnoi* discs exposed to different light intensities for 18 days. High PPFD = 224 $\mu m\ m^{-2}s^{-1}$, Medium PPFD = 120 $\mu m\ m^{-2}s^{-1}$, Low PPFD = 60 $\mu m\ m^{-2}s^{-1}$ and Darkness = 0 $\mu m\ m^{-2}s^{-1}$. Each point represents an average (n = 4) of the total bacteria colony counts and *Phaeobacter* colony counts in different in different sampling days and the error bar represents the standard deviation.

4.5 Model equations

Lotka-Volterra equations were formulated to describe the interactions between algae, bacteria and *Phaeobacter* (Figure 10). It should be noted that we selected the generalized version, assuming logistic growth, instead of exponential growth. Additionally, we use a logarithmic scale of ten bases to ensure that all variables in the model are of the same order of magnitude. The final equations read as follows:

$$\frac{d_{N_A}}{dt} = \frac{1}{\ln 10}\mu_A \left(1 - 10^{(N_A - K_A)} - \left(\alpha_{AB}10^{(N_B - K_A)}\right) - \left(\alpha_{AP}10^{(N_P - K_A)}\right)\right) \tag{4.6}$$

$$\frac{d_{N_B}}{dt} = \frac{1}{\ln 10}\mu_B \left(1 - 10^{N_B - K_B} - \left(\alpha_{BA}\frac{N_A}{10^{K_B}}\right) - \left(\alpha_{BP}10^{N_P - K_B}\right)\right) \tag{4.7}$$

$$\frac{d_{N_P}}{dt} = \frac{1}{\ln 10}\mu_P \times \left(1 - 10^{N_P - K_P} - \left(\alpha_{PA}\frac{N_A}{10^{K_P}}\right) - \left(\alpha_{PB}10^{N_B - K_P}\right)\right) \tag{4.8}$$

where N_A corresponds to base ten logarithm of the number of cells for the algae, N_B and N_P with the base ten logarithm of the amount of bacteria present in the algae (microbiome) and *Phaeobacter*, respectively. The explanations for each parameter can be found in Table 3:

Table 3: Shows all parameters of equations (4.6), (4.7) and (4.8) and their meaning.

Parameters	Meaning
μ_A	Algae growth rate.
K_A	Algae carrying capacity.
α_{AB}	Interaction coefficient, effect of the bacteria over the algae.
α_{AP}	Interaction coefficient, effect of *Phaeobacter* over the algae.
μ_B	Bacteria growth rate
K_B	Bacteria carrying capacity.
α_{BA}	Interaction coefficient, effect of algae over bacteria.
α_{BP}	Interaction coefficient, effect of *Phaeobacter* over total bacteria.
μ_P	*Phaeobacter* growth rate.
K_P	*Phaeobacter* carrying capacity.
α_{PA}	Interaction coefficient effect of algae over *Phaeobacter*.
α_{PB}	Interaction coefficient effect of bacteria over *Phaeobacter*.

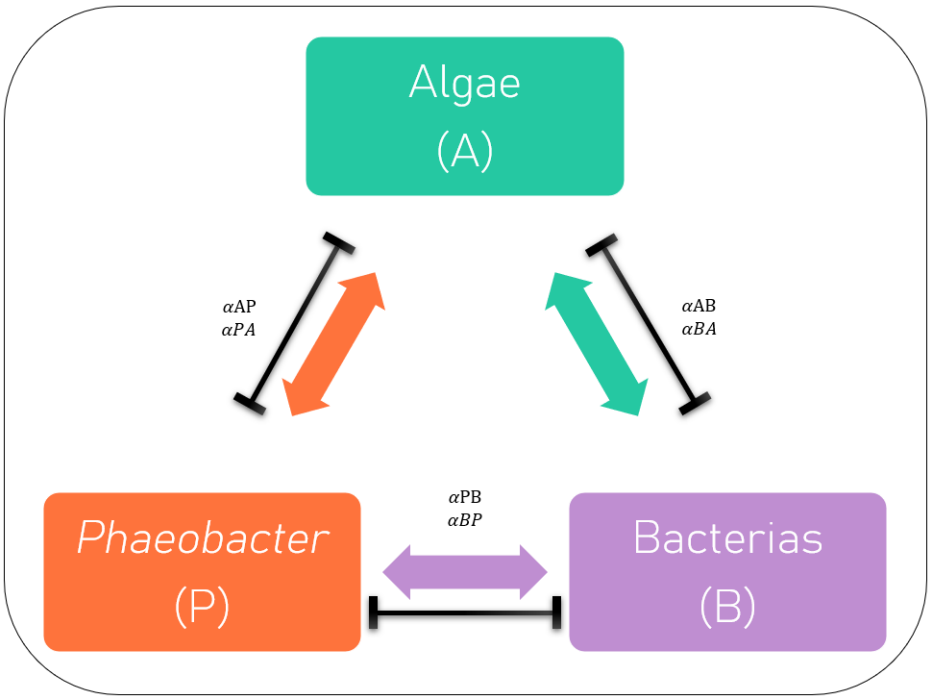


Figure 10: Schematic representation of all possible interactions between algae, bacteria and *Phaeobacter*.

4.6 Structural identifiability analysis of the model

We used GenSSI2 toolbox to analyze whether the model parameters can be uniquely identified taking into account the observable function, i.e. the variables that can be measured in a single experiment. Note that in this particular case all three variables are being measured. The number of cells is approximated taking into account the measured surface, the medium size of the cell for *Ulva ohnoi* ($6\mu m \times 6\mu m$), and the fact that the width of the algae incorporates 2 cells.

The analysis reveals that the model is locally structurally identifiable. Global structural identifiability would only be guaranteed if it was possible to grow the different species individually. However, this is not possible in this particular example. *Phaeobacter* is an heterotrophic bacteria and does not grow in the algae medium unless the algae is there. Of course, this may also be the case for other bacteria present in the algae microbiome.

We explored numerically the consequences of the model being locally structurally identifiable. To do so, we assumed given values for growth rates, carrying capacities, and interaction coefficients and generated synthetic data by simulating the Lotka-Volterra model. The ideal case of abundant (50 sampling times) noiseless data was considered. To generate the data, the AMIGOSData function was used. We then used

the synthetic data generated to estimate the model parameters.

Table 4 presents the results for this illustrative example. Two cases were considered, a first case in which the bounds for parameter estimation are rather large (often the case in real practice), and a second case in which the search space is restricted. In both cases, AMIGOPE with eSS was used to improve global convergence. The results show that even in this ideal scenario in which abundant noiseless data is available, the results can be wrong when the bounds are unknown. Note that the maps of interactions recovered in the first and second cases would differ, while the simulations would be almost the same. Figure 11 presents the difference between the model predictions as obtained with the suboptimal solution and the real parameter values. The maximum difference is lower than 3×10^{-3}, much lower than the usual experimental error. Both models, with different interactions, would be indistinguishable in real practice.

Table 4: Parameters used for synthetic data generation and parameters recovered in parameter estimation with abundant noiseless data. Estimates 1: correspond to the parameter estimation using larger bounds ($[0, 10]$ for growth rates and carrying capacities and $[-10, 10]$ for interaction coefficients. Estimates 2: correspond to the parameter estimation using stricter bounds ($[0, 10]$ for growth rates and carrying capacities and $[-2, 2]$ for interaction coefficients. Highlighted in red those cases in which the sign of the interaction is wrong.

Parameters	Real Value	Estimates 1 Large bounds	Estimates 2 Strict bounds
μ_A	0.4	0.400	0.399
K_A	8	7.939	7.998
α_{AB}	0.5	-0.727	0.469
α_{AP}	0.5	0.669	0.504
μ_B	0.3	0.300	0.300
K_B	6	6.003	6.001
α_{BA}	-0.1	-0.099	-0.099
α_{BP}	-0.2	-0.201	-0.200
μ_P	0.5	0.499	0.499
K_P	7	7.001	7.001
α_{PA}	0.08	0.076	0.079
α_{PB}	-1	-0.966	-0.990

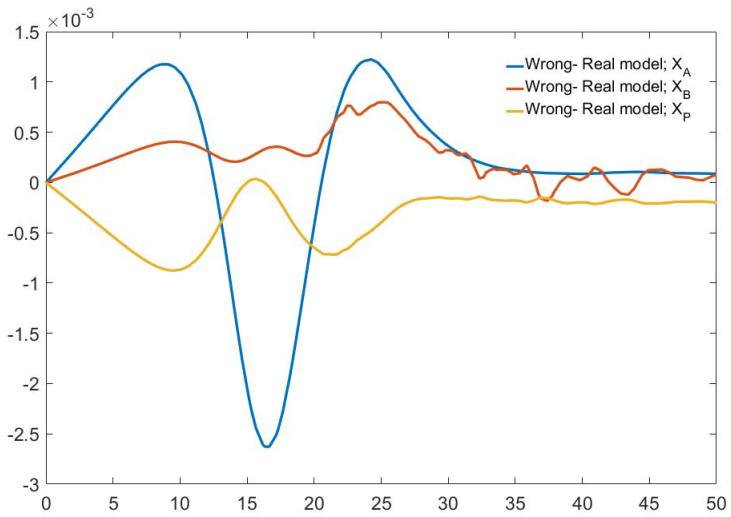


Figure 11: Difference between model predictions as obtained with the suboptimal parameters and the real parameters.

4.7 Practical identifiability analysis of the model

The same experiment was performed for a more realistic scenario, similar to what we have in practice, with limited noisy data. Table 5 presents the results obtained for the case of having **seven** sampling times with 5% experimental noise. The estimated parameter values correspond to the restricted search space, with bounds $[0, 10]$ for growth rates and carrying capacities, and $[-2, 2]$ for the interaction coefficients.

The quality of fit is very good (R^2=0.95) (see Figure 12, but the recovered map of interactions is not correct; nor the order of magnitude of some of the interactions.

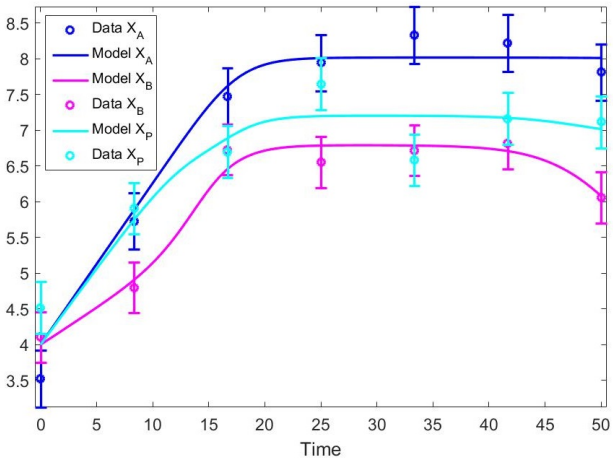


Figure 12: Best fit as obtained for limited noisy experimental data in a synthetic example.

Table 5: Parameters used for synthetic data generation and parameters recovered in parameter estimation with abundant noiseless data. Estimates correspond to the parameter estimation using strict bounds ($[0, 10]$ for growth rates and carrying capacities and $[-2, 2]$ for interaction coefficients. Highlighted in red those cases in which the sign of the interaction is wrong. Highlighted in blue those cases in which the order of magnitude is wrong.

Parameters	Real Value	Estimates Strict bounds
μ_A	0.4	0.520
K_A	8	7.923
α_{AB}	0.5	1.810
α_{AP}	0.5	-1.968
μ_B	0.3	0.235
K_B	6	6.343
α_{BA}	-0.1	0.189
α_{BP}	-0.2	-1.482
μ_P	0.5	0.490
K_P	7	6.733
α_{PA}	0.08	-0.029
α_{PB}	-1	-1.232

4.8 Model complexity reduction: hypotheses

These identifiability problems shown in previous sections could be addressed, at least partially, using a specific multi-experiment design in which species can grow individually and under mixed conditions (data not shown). However, despite the experimental efforts made (see the discussion for further details), the required experiments cannot be performed for the system under consideration. Therefore, we decided to simplify the model.

Taking into account that bacteria cannot grow in the selected medium as it does not contain carbon sources. We can assume that μ_B and μ_P are zero for this medium. Furthermore, the carrying capacity would be zero in that medium as the bacteria will eventually die. With these assumptions, the model reads:

$$\frac{dN_A}{dt} = \frac{1}{ln(10)}\mu_A(1 - 10^{(N_A - K_A)} - \tilde{a}_{AB}10^{N_B} - \tilde{a}_{AP}10^{N_P}) \quad (4.9)$$

$$\frac{dN_B}{dt} = \frac{1}{ln(10)}(-a_{BA}10^{N_A} - a_{BP}10^{N_P}) \quad (4.10)$$

$$\frac{dN_P}{dt} = \frac{1}{ln(10)}(-a_{PA}10^{N_A} - a_{PB}10^{N_B}) \quad (4.11)$$

where $\tilde{a}$ represent normalized interactions $\tilde{a} = a/10^{KA}$ so as to avoid interaction coefficients of different orders of magnitude.

GenSSi2 shows that this model is again only locally identifiable. However, it becomes structurally globally identifiable when fixing one interaction coefficient. Taking this into account, we define two different maps of interactions (see Figures 13) that are plausible in view of the experimental data.

1. First hypothesis (H1): *Phaeobacter* has no impact on algae, $\alpha_{AP} = 0$.

2. Second hypothesis (H2): algae has no impact on *Phaeobacter*, $\alpha_{PA} = 0$.

4.8.1 Parameter estimation

Parameter estimation was implemented as a multi-experiment problem with global and local parameters. Global parameters are those that take the same value for all experiments, whereas local parameters change from experiment to experiment. Carrying capacities depend on the medium, and thus these parameters are assumed to be global. Note that for the dark experiment, we assumed (and observed) that there is practically no growth for algae; thus $\mu_A \in [0, 10^{-7}]$.

For the interaction coefficients, we assumed two possibilities: A) The algae growth rate depends on the illumination; thus μ_A is considered a local parameter. All other parameters are considered global. Thus, independent of the illumination. B) μ_A, α_{BA} and α_{PA} are considered to depend on the illumination. Therefore, they are considered as local parameters. The following Table 5 summarizes the least squares values achieved during the parameter estimation for all cases. Note that the number of parameters to be estimated n_θ differs for each candidate model.

The parameter estimation problem was solved several times (at least 10 times) to ensure convergence to the best possible solution. If the optimal value of the parameters activates the bounds, the size of the

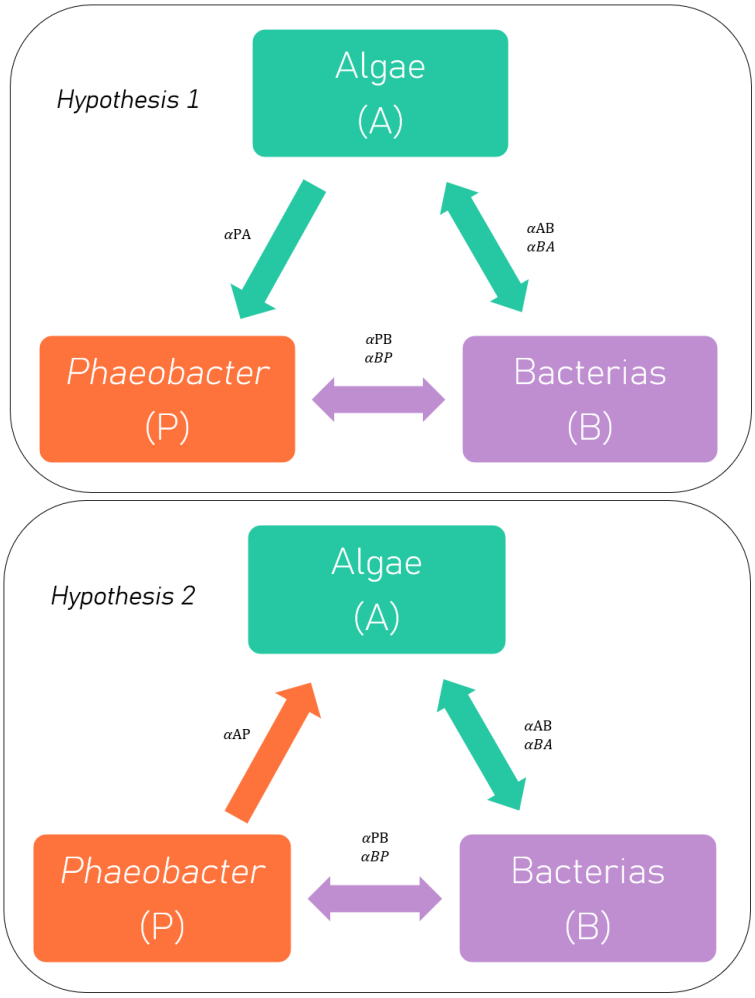


Figure 13: Maps of interactions for two different hypotheses.

search space was increased. In the optimal solutions reported, the optimal parameters do not activate the bounds.

Table 5: Presents candidate models with the corresponding hypotheses, global and local parameters, number of parameters, and least squares value at the optimum.

Hypothesis	Global parameters	Local parameters	n_θ	LSQ
H1-A	$K_A\ \alpha_{AB}\ \alpha_{BA}\ \alpha_{BP}\ \alpha_{PA}\ \alpha_{PB}$	μ_A	10	4.566
H1-B	$K_A\ \alpha_{AB}\ \alpha_{PB}$	$\mu_A\ \alpha_{BA}\ \alpha_{PA}$	16	1.959
H2-A	$K_A\ \alpha_{AB}\ \alpha_{BA}\ \alpha_{BP}\ \alpha_{AP}\ \alpha_{PB}$	μ_A	10	5.248
H2-B	$K_A\ \alpha_{AB}\ \alpha_{PB}\ \alpha_{AP}$	$\mu_A\ \alpha_{BA}$	13	3.086

The most successful case corresponds to H1-B, so *Phaeobacter* has no impact on algae, $\alpha_{AP} = 0$ and the interaction coefficients α_{BA} and α_{PA} depend on the illumination. The following Table 6 presents the results obtained for the different parameters under various experimental conditions. HL, high PPFD;

ML, medium PPFD; LL, low PPFD; and D, PPFD.

Table 6: Presents the optimal parameter values for the model H1-B.

Results	HL	ML	LL	D
μ_A	$4.649 \cdot 10^{-3}$	$1.962 \cdot 10^{-2}$	$1.206 \cdot 10^{-2}$	$3.708 \cdot 10^{-11}$
K_A	8.264	8.264	8.264	8.264
α_{AB}	$7.913 \cdot 10^{-8}$	$7.913 \cdot 10^{-8}$	$7.913 \cdot 10^{-8}$	$7.913 \cdot 10^{-8}$
α_{AP}	0	0	0	0
α_{BA}	$-3.986 \cdot 10^{-11}$	$-4.739 \cdot 10^{-11}$	$-5.942 \cdot 10^{-11}$	$2.792 \cdot 10^{-11}$
α_{BP}	$4.039 \cdot 10^{-9}$	$4.039 \cdot 10^{-9}$	$4.039 \cdot 10^{-9}$	$4.039 \cdot 10^{-9}$
α_{PA}	$-2.319 \cdot 10^{-9}$	$-7.623 \cdot 10^{-10}$	$-6.719 \cdot 10^{-10}$	$-2.863 \cdot 10^{-9}$
α_{PB}	$3.801 \cdot 10^{-8}$	$3.801 \cdot 10^{-8}$	$3.801 \cdot 10^{-8}$	$3.801 \cdot 10^{-8}$

The following Figure 14 shows how local parameters depend on PPFD.

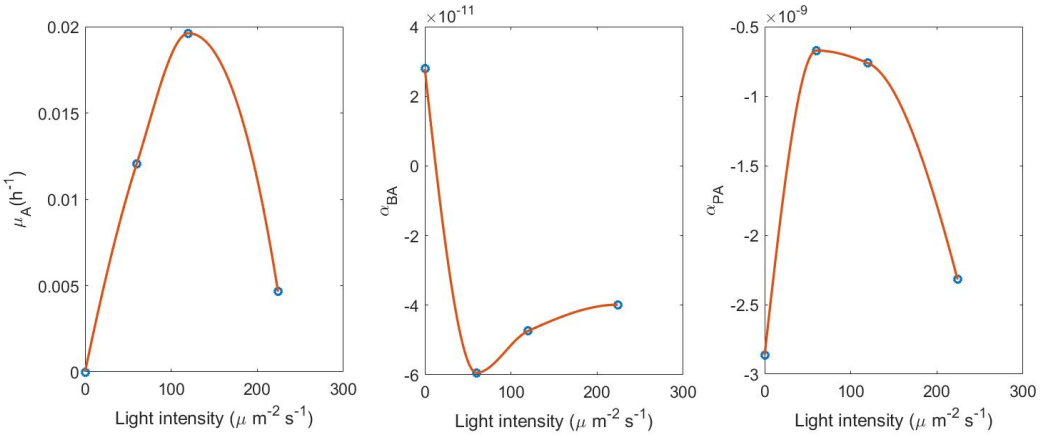


Figure 14: Optimal parameter values as function of the photosynthetic photon flux density (PPFD). Circles represent optimal parameter values at four tested conditions. Continuous line is a spline approximation as a function of the PPFD.

The results show that:

- α_{AB} and α_{BA} have opposite signs. $\alpha_{AB} > 0$ and $\alpha_{BA} < 0$ but four orders of magnitude lower in absolute value. The relationship is defined as ammensalism. Bacteria get benefit from the interaction, while algae are slightly harmed. Note that the situation changes at really low light intensities; in that case, both species compete.

- α_{BP} and α_{PB} have the same positive sign that indicates competition between species. Probably bacteria and *Phaeobacter* compete for the same resources. Bacteria tend to dominate the competition as $\alpha_{PB} > \alpha_{BP}$ in absolute value.

- In general, the overall effect of algae on bacteria is two orders of magnitude lower than that obtained for *Phaeobacter*. This would indicate that *Phaeobacter* gets more benefit from algae than the community of bacteria. This might be explained by considering that different species within the community of bacteria might cooperate and the dependence on algae would be lower.

- The algae growth rate is clearly influenced by the intensity of light. The results show a detrimental effect for high light intensities. Results also show that two different values of light intensity will lead to the same growth rate.

4.8.2 Optimization

The aim of optimization is to select a range of light intensities that offers the best compromise between algae production and *Phaeobacter* retention on the surface of the algae.

To optimize the process, a new script in AMIGO2 was implemented that allows the system dynamics to be simulated for different values of light intensity using the splines shown in Figure 14. A tight range of light intensity values was preselected $[115 - 130]$ $\mu m\ m^{-2}s^{-1}$ to guarantee a minimum desirable amount of algae in the number of cells after $8h$.

Figure 15 presents the results. As shown in the figures, the model predicts that algae production is rather insensitive to light intensity in the selected range, whereas retention of *Phaeobacter* varies significantly. A good compromise seems to be in the light intensities within the range $[127 - 130]$ $\mu m\ m^{-2}s^{-1}$.

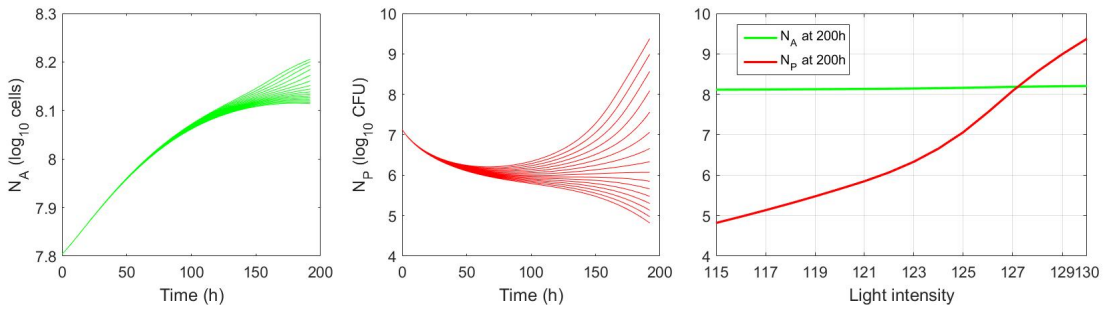


Figure 15: Light intensity optimization. Figure a) presents the dynamics of algae growth for a range of light intensities $[115 - 130]$ $\mu m\ m^{-2}s^{-1}$; b) shows the dynamics of *Phaeobacter*; c) shows the production at $8h$ versus the light intensity.

5 Discussion

This work addressed the modeling of interactions between algae, bacterial microbiome, and a probiotic bacterial species that produces an antibiotic antagonistic to fish pathogens (*Phaeobacter*). The aim was to elucidate the influence of light intensity in the interactions and thus the retention of the probiotic species. The results can be applied to the design of multitrophic inland aquaculture systems with recirculation (IMTA-RAS). The specific experimental system we considered combines: *Ulva ohnoi*, a macroalgae of commercial interest; the algae microbiome, here regarded as "bacteria", and probiotic bacteria *Phaeobacter spp.*.

The modeling was started using available experimental data and a classical Lotka-Volterra ecological model. The data consisted of the algae surface and the abundances of microbial species over time during 4 experiments under different light intensity conditions. Unfortunately, in the original experimental setup, the growth medium was being modified between sampling times. Thus, the hypotheses underlying the Lotka-Volterra model were not met. The addition of nutrients modifies the carrying capacity of the species. Note that measurements on nutrient consumption were not available so alternative models, such as those based on Monod equations, were not feasible. A new experiment was designed and performed to ensure that the data fit well with the classical ecological modeling.

Although many modelers use LV models to decipher interactions from "mixed" culture data, we discovered that this is indeed a challenge. Since global identifiability cannot be guaranteed, we explored this in more detail, concluding that a possible solution comes from performing individual experiments. And use all experiments together to estimate model parameters. Remarkably, for this particular system, only the "mixed" experiment is feasible, that is, we can grow algae and bacteria together, but it is not possible to grow *Phaeobacter* (and possibly most of the bacteria on the algae surface) in the algae growth medium. The team led by Dr. Pintado performed experiments to grow *Phaeobacter* in the algae medium by adding glycerol as a carbon source. However, when glycerol was added to the algae-bacteria system, the dynamics of bacterial growth changed completely, the turbidity of the medium increased rapidly as a result of bacterial growth, and the experiment results cannot be compared to those of the standard medium (data not shown).

The only choice we had is to reduce the model and solve the parameter estimation problem under two different hypotheses. Both models were able to recover the data. However, the most successful model was the one in which the amount of *Phaeobacter* does not affect the algae dynamics ($\alpha_{AP} = 0$). Note

that this result would need further confirmation. Taking into account that to experimentally test this result, we would need sterile algae, which is difficult if not impossible to achieve without damage, it would be desirable to explore the genomes of species to analyze the possibilities (if any) of the effect *Phaeobacter* on *Ulva ohnoi*.

This model also showed that 3 parameters (μ_A, α_{BA} and α_{PA}) are dependent on light intensity. Figure 14 shows how the increase in light intensity influences the behavior of these 3 light-dependent parameters. In all cases, the relationship with light intensity is non-linear. The specific growth rate of algae μ_A grows up to a certain level of light intensity, while after that it decreases. This would be in agreement with previous observations that showed that high light intensities induced a thin and fragile surface, or photoinhibition of algae (Bialevich, Zachleder, & Bišová, 2022). We also observed those negative impacts in our samples (see the pictures in Figure 7) and the ratio between the wet weight and the surface area of the algae (Figure 8). At High Light, between T13 and T18, although the algae had more surface area, its weight decreased, demonstrating how the algae became thinner. The same was demonstrated by Oca et al., 2019, where high light intensities and low stock density cause damage in *Ulva* and by Falkowski & LaRoche, 1991, in nature.

However, light also positively influences the interaction of algae-*Phaeobacter* (α_{PA}) and algae-bacteria microbiome (α_{BA}). This could be explained if some of the metabolites produced in the primary metabolism of algae are also responsible for *Phaeobacter* or the survival of the algae bacterial community. According to Roth-Schulze et al., 2018, the bacterial community on the algal surface is not defined by particular species, but by the function they perform. Roth-Schulze et al., 2018 determined that certain bacterial genes play a critical role in the interaction between the alga and its microbiome. About 70% form a stable core set of function genes in the bacterial communities and the remaining 30% are possibly involved in local or host-specific adaptations. *Phaeobacter gallaeciensis* is a bacterium found in a biofilm on the surface of *Ulva australis* (Rao, Webb, & Kjelleberg, 2006), so taking into account the results of our model, we hypothesize that *Phaeobacter gallaeciensis* may have some function genes that benefit *Ulva ohnoi*.

Figure 9 shows that after inoculation, the number of CFU/g of *Phaeobacter* shows a rapid increase and afterward a slow decrease; the dynamics depends on the light condition and, after T15 (ML and LL conditions), also depends on the secondary growth of the algae (see Section 4.4). This behavior may be due to: 1) competition between *Phaeobacter* and the microbiome present on the algal surface; 2) certain metabolites excreted by the algae have a negative influence on *Phaeobacter* concentration; or 3)

a combination of the two previous hypotheses. According to the results of the model, α_{BP} and α_{PB} are positive, thus showing competition between the algal microbiome and *Phaeobacter*. The model also showed that *Phaeobacter* is more dependent on the algae than the microbiome present on the algal surface, so hypotheses 2) and 3) are less likely in light of our model.

Rao et al., 2006 showed that on the surface of *Ulva australis*, *Phaeobacter* is numerically dominant and that it can invade and disperse into pre-established biofilms. Our experimental and modeling results show a completely different behavior in the interaction with *Ulva ohnoi*. Based on the assumption that there is competition between *Phaeobacter* and the algal microbiome, we hypothesize that: a) other bacteria with a functional gene(s) similar to (or the same) of *Phaeobacter* are better adapted to our experimental conditions; b) *Phaeobacter* cannot integrate into the existing community due to competitive interactions; c) *Ulva ohnoi* does not produce the metabolites necessary for *Phaeobacter* to maintain itself on the algal surface.

To decipher the real mechanism or mechanisms underlying this behavior, additional experiments and analyzes are required: metagenomics analysis of the algae microbiome (under development); to explore the metabolites that are being exchanged between the algae and the bacteria (under development); explore the biological functions that *Phaeobacter* may provide to *Ulva ohnoi*; and decipher the type of interactions between *Phaeobacter* and the microbiome, whether 1 to 1, 1 to several or 1 to all. Remarkably, finding functional genes and possibilities of interactions would benefit from the genome sequencing and metabolic reconstruction of at least the most relevant species.

We used the model to explore the range of light intensity values that would offer a good compromise between algae growth rate and retention of *Phaeobacter* in the system. To do so, we simulated the model for a range of light intensities in which a minimum algae growth rate was achieved. Figures 15a and b show the dynamics of the system under different light intensities. Figure 15c shows that, within the light range, the algae growth rate is rather insensitive to light intensity, while the final amount of *Phaeobacter* varies quite significantly, achieving the maximum values at higher intensities. Our model shows that a good compromise would be achieved in the range [127-130] $\mu m\ m^{-2}s^{-1}$. It is therefore suggested to perform an additional experiment under these conditions for model validation.

Model results need to be taken with caution, particularly those related to the dependence of the parameters on the light intensity. It is expected that, with additional experiments, we will be able to

confirm or improve the current model. In addition, nutrients uptake measurements will be made so as to increase the detail in the model and account for experiments in batch conditions (as those used in this work) and semi-batch (with medium changes over time). The possibility of combining experiments under different nutrient regimes will provide more information about the dynamics of the system and the interactions between species.

6 Conclusion

This work tackled the modeling of the role of light intensity into the interactions between algae - bacterial microbiome - *Phaeobacter*, a system of high interest in multi-trophic aquaculture. A classical Lotka-Volterra model was considered. Data was also obtained in the context of this work.

The structural and practical identifiability analyses, revealed that it is not possible to uniquelly identify model parameters. Therefore, biological hypothesis were used to proposed two identifiable reduced models. Model parameters were identified using a multi-experiment parameter estimation scheme as implemented in the AMIGO2 toolbox. The parameter estimation problem was solved using a global optimizer based on Scatter Search.

The best results corresponded to the model in which there is no interaction of *Phaeobacter* over the algae ($\alpha_{AP} = 0$) and that μ_A, α_{BA} and α_{PA} depend on light intensity. The modeling found that *Phaeobacter* competes with the bacterial microbiome that is found on the algal surface. We also used a spline to approximate the dependence of μ_A, α_{BA} and α_{PA} on the light intensity which resulted to be highly nonlinear.

Finally we formulated an optimization problem with the aim of calculating the range of light intensities that would provide a good compromise between algae growth and the probiotic *Phaeobacter*. Our results revealed that an intensity around [127-130] $\mu m\ m^{-2}s^{-1}$ would be well suited for our purposes.

In this work we hypothesize about possible biological mechanisms that would explain the type of interactions we have obtained through modeling. Ideally, performing one experiment under optimal conditions could help us to validate or refine the model. In addition, the detailed metagenomics study of the algae microbiome (under development) and the study on the metabolites that are being exchanged in the algae surface could provide information for a more detailed mechanistic model.

Remarkably, we are limited in the type of experiments that can be done for the purpose of modeling. Thus, we envision that having access to the genome sequences of the most abundant species and, if possible, to their metabolic reconstructions, would provide us with relevant additional information to decipher the types of functions that algae, bacterial microbiome and *Phaeobacter* are performing for each other.

References

Balsa-Canto, E., Alonso, A., & Banga, J. (2010). An iterative identification procedure for dynamic modeling of biochemical networks. *BMC Systems Biology*, *4:11*.

Balsa-Canto, E., Alonso-del Real, J., & Querol, A. (2020). Temperature shapes ecological dynamics in mixed culture fermentations driven by two species of the saccharomyces genus. *Frontiers in Bioengineering and Biotechnology*, *8*.

Balsa-Canto, E., Henriques, D., Gabor, A., & Banga, J. (2016). AMIGO2, a toolbox for dynamic modeling, optimization and control in systems biology. *Bioinformatics*, *32(21)*, 3357-3359.

Bialevich, V., Zachleder, V., & Bišová, K. (2022). The effect of variable light source and light intensity on the growth of three algal species. *Cells*, *11*(8), 1293.

Bregnballe, J., et al. (2010). A guide to recirculation aquaculture: an introduction to the new environmentally friendly and highly productive closed fish farming systems.

Brinkhoff, T., Bach, G., Heidorn, T., Liang, L., Schlingloff, A., & Simon, M. (2004). Antibiotic production by a roseobacter clade-affiliated species from the german wadden sea and its antagonistic effects on indigenous isolates. *Applied and environmental microbiology*, *70*(4), 2560–2565.

Burke, C., Steinberg, P., Rusch, D., Kjelleberg, S., & Thomas, T. (2011a). Bacterial community assembly based on functional genes rather than species. *Proceedings of the National Academy of Sciences*, *108*(34), 14288–14293.

Burke, C., Thomas, T., Lewis, M., Steinberg, P., & Kjelleberg, S. (2011b). Composition, uniqueness and variability of the epiphytic bacterial community of the green alga ulva australis. *The ISME journal*, *5*(4), 590–600.

Chiş, O., Banga, J. R., & Balsa-Canto, E. (2011). Genssi: a software toolbox for structural identifiability analysis of biological models. *Bioinformatics*, *27*(18), 2610–2611.

Chis, O., Banga, J. R., & Balsa-Canto, E. (2011). Structural identifiability of systems biology models: A critical comparison of methods. *Plos One*, *6*(11), e27755.

Chopin, T., et al. (2010). Integrated multi-trophic aquaculture.

Correia, M., Azevedo, I. C., Peres, H., Magalhães, R., Oliva-Teles, A., Almeida, C. M. R., & Guimarães, L. (2020). Integrated multi-trophic aquaculture: a laboratory and hands-on experimental activity to promote environmental sustainability awareness and value of aquaculture products. *Frontiers in Marine Science*, *7*, 156.

D'Alvise, P. W., Phippen, C. B., Nielsen, K. F., & Gram, L. (2016). Influence of iron on production of the antibacterial compound tropodithietic acid and its noninhibitory analog in phaeobacter inhibens.

Applied and environmental microbiology, *82*(2), 502–509.

De La Maza, M., & Yuret, D. (1994). Dynamic hill climbing. *AI expert*, *9*(26), 26.

Dréo J, T. E., Petrowski A, & P, S. (2006). *Metaheuristics for hard optimization. methods and case studies*. Springer.

D'alvise, P. W., Lillebø, S., Prol-Garcia, M. J., Wergeland, H. I., Nielsen, K. F., Bergh, Ø., & Gram, L. (2012). Phaeobacter gallaeciensis reduces vibrio anguillarum in cultures of microalgae and rotifers, and prevents vibriosis in cod larvae.

Edelstein-Keshet, L. (2005). *Mathematical models in biology*. SIAM.

Egan, S., Harder, T., Burke, C., Steinberg, P., Kjelleberg, S., & Thomas, T. (2013). The seaweed holobiont: understanding seaweed–bacteria interactions. *FEMS Microbiology Reviews*, *37*(3), 462–476.

Egea, J., Balsa-Canto, E., Garcia, M., & Banga, J. (2009). Dynamic optimization of nonlinear processes with an enhanced scatter search method. *Industrial & Engineering Chemistry Research*, *48*(9), 4388-4401.

Falkowski, P. G., & LaRoche, J. (1991). Acclimation to spectral irradiance in algae. *Journal of Phycology*, *27*(1), 8–14.

FAO. (2022a). *The state of world fisheries and aquaculture 2022: Towards blue transformation*. Author.

FAO. (2022b). *Thinking about the future of food safety - a foresight report*. Author.

Favero, M., McDade, J., Robertsen, J., Hoffman, R., & Edwards, R. (1968). Microbiological sampling of surfaces. *Journal of applied Bacteriology*, *31*(3), 336–343.

Fortes, M., & Lüning, K. (1980). Growth rates of north sea macroalgae in relation to temperature, irradiance and photoperiod. *Helgoländer Meeresuntersuchungen*, *34*(1), 15–29.

Fraga-Corral, M., Ronza, P., Garcia-Oliveira, P., Pereira, A., Losada, A., Prieto, M., ... Simal-Gandara, J. (2021). Aquaculture as a circular bio-economy model with galicia as a study case: How to transform waste into revalorized by-products. *Trends in Food Science & Technology*.

Gavina, M. K. A., Tahara, T., Tainaka, K.-i., Ito, H., Morita, S., Ichinose, G., ... Yoshimura, J. (2018). Multi-species coexistence in lotka-volterra competitive systems with crowding effects. *Scientific reports*, *8*(1), 1–8.

Gilpin, M., & Ayala, F. (1973). Global models of growth and competition. *Proc. Nat. Acad. Sci. USA*, *70*(12), 3590 - 3593.

Grasshoff, K., Kremling, K., & Anderson, L. (1999). Methods of seawater analysis. chapter 10.2.5: determination of dissolved inorganic phosphate. In (pp. 170–174). Weinheim, Germany.

Hindmarsh, A. C., Brown, P. N., Grant, K. E., Lee, S. L., Serban, R., Shumaker, D. E., & Woodward,

C. S. (2005a). Sundials: Suite of nonlinear and differential/algebraic equation solvers. *ACM Transactions on Mathematical Software (TOMS)*, *31*(3), 363–396.

Hindmarsh, A. C., Brown, P. N., Grant, K. E., Lee, S. L., Serban, R., Shumaker, D. E., & Woodward, C. (2005b). Sundials: Suite of nonlinear and differential/algebraic equation solvers. *ACM Trans. Math. Softw.*, *31*(3), 363–396.

Kinley, R. D., Martinez-Fernandez, G., Matthews, M. K., de Nys, R., Magnusson, M., & Tomkins, N. W. (2020). Mitigating the carbon footprint and improving productivity of ruminant livestock agriculture using a red seaweed. *Journal of Cleaner production*, *259*, 120836.

Lachnit, T., Fischer, M., Künzel, S., Baines, J. F., & Harder, T. (2013). Compounds associated with algal surfaces mediate epiphytic colonization of the marine macroalga fucus vesiculosus. *FEMS microbiology ecology*, *84*(2), 411–420.

Lavaud, R., Guyondet, T., Filgueira, R., Tremblay, R., & Comeau, L. A. (2020). Modelling bivalve culture-eutrophication interactions in shallow coastal ecosystems. *Marine Pollution Bulletin*, *157*, 111282.

Ligon, T., Fröhlich, F., Chiş, O., Banga, J., Balsa-Canto, E., & Hasenauer, J. (2018). Genssi 2.0: Multi-experiment structural identifiability analysis of sbml models. *Bioinformatics*, *34*(8), 1421-1423.

Longford, S. R., Tujula, N. A., Crocetti, G. R., Holmes, A. J., Holmström, C., Kjelleberg, S., ... Taylor, M. W. (2007). Comparisons of diversity of bacterial communities associated with three sessile marine eukaryotes. *Aquatic microbial ecology*, *48*(3), 217–229.

Mantri, V. A., Kazi, M. A., Balar, N. B., Gupta, V., & Gajaria, T. (2020). Concise review of green algal genus ulva linnaeus. *Journal of Applied Phycology*, *32*(5), 2725–2741.

Martens, T., Heidorn, T., Pukall, R., Simon, M., Tindall, B. J., & Brinkhoff, T. (2006). Reclassification of roseobacter gallaeciensis ruiz-ponte et al. 1998 as phaeobacter gallaeciensis gen. nov., comb. nov., description of phaeobacter inhibens sp. nov., reclassification of ruegeria algicola (lafay et al. 1995) uchino et al. 1999 as marinovum algicola gen. nov., comb. nov., and emended descriptions of the genera roseobacter, ruegeria and leisingera. *International Journal of Systematic and Evolutionary Microbiology*, *56*(6), 1293–1304.

Murray, J. D. (2002). *Mathematical biology i. an introduction* (Vol. 17). New York: Springer.

Nations, U. (n.d.). *World population projected to reach 9.8 billion in 2050, and 11.2 billion in 2100| united nations.*

Nelder, J., & R, M. (1965). A simplex method for function minimization. *Comput J*, *7*, 308-313.

Oca, J., Cremades, J., Jiménez, P., Pintado, J., & Masaló, I. (2019). Culture of the seaweed ulva ohnoi integrated in a solea senegalensis recirculating system: influence of light and biomass stocking

density on macroalgae productivity. *Journal of Applied Phycology*, *31*(4), 2461–2467.

Pauly, D., & Zeller, D. (2016). Catch reconstructions reveal that global marine fisheries catches are higher than reported and declining. *Nature communications*, *7*(1), 1–9.

Pintado, J., Ruiz, P., Cremades, J., & Wichard, T. (2022). *The effect of light on Phaeobacter gallaeciensis biofilms on U. ohnoi (ulvales, chlorophyta).*

Prol, M., Bruhn, J. B., Pintado, J., & Gram, L. (2009). Real-time pcr detection and quantification of fish probiotic phaeobacter strain 27-4 and fish pathogenic vibrio in microalgae, rotifer, artemia and first feeding turbot (psetta maxima) larvae. *Journal of Applied Microbiology*, *106*(4), 1292–1303.

Prol-García, M. J., Gómez, M., Sánchez, L., & Pintado, J. (2014). Phaeobacter grown in biofilters: A new strategy for the control of vibrionaceae in aquaculture. *Aquaculture Research*, *45*(6), 1012–1025.

Prol-García, M. J., & Pintado, J. (2013). Effectiveness of probiotic phaeobacter bacteria grown in biofilters against vibrio anguillarum infections in the rearing of turbot (psetta maxima) larvae. *Marine biotechnology*, *15*(6), 726–738.

Prol-García, M. J., Planas, M., & Pintado, J. (2010). Different colonization and residence time of listonella anguillarum and vibrio splendidus in the rotifer brachionus plicatilis determined by real-time pcr and dgge. *Aquaculture*, *302*(1-2), 26–35.

Qiu, X., Carter, C. G., Hilder, P. E., & Hadley, S. (2022). A dynamic nutrient mass balance model for optimizing waste treatment in ras and associated imta system. *Aquaculture*, *555*, 738216.

Rao, D., Webb, J. S., & Kjelleberg, S. (2006). Microbial colonization and competition on the marine alga ulva australis. *Applied and Environmental Microbiology*, *72*(8), 5547–5555.

Rice, E. W., Baird, R. B., Eaton, A. D., Clesceri, L. S., et al. (2012). *Standard methods for the examination of water and wastewater* (Vol. 10). American public health association Washington, DC.

Roth-Schulze, A. J., Pintado, J., Zozaya-Valdés, E., Cremades, J., Ruiz, P., Kjelleberg, S., & Thomas, T. (2018). Functional biogeography and host specificity of bacterial communities associated with the marine green alga ulva spp. *Molecular ecology*, *27*(8), 1952–1965.

Rybak, A. S. (2018). Species of ulva (ulvophyceae, chlorophyta) as indicators of salinity. *Ecological indicators*, *85*, 253–261.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., ... others (2012). Fiji: an open-source platform for biological-image analysis. *Nature methods*, *9*(7), 676–682.

Teles, A. O., Couto, A., Enes, P., & Peres, H. (2020). Dietary protein requirements of fish–a meta-analysis. *Reviews in Aquaculture*, *12*(3), 1445–1477.

UNICEF, et al. (2021). The state of food security and nutrition in the world 2021.

Volterra, V. (1926). Variazioni e fluttuazioni del numero d'individui in specie animali conviventi. In (p. 31-113).