A mathematical model of marine mucilage, the case of the liga on the Basque coast

Charles Pierre, Guy Vallet

A mathematical model of marine mucilage, the case of the *liga* on the Basque coast

Charles PIERRE* Guy VALLET†

January 9, 2023

Abstract

In this paper we are interested in modeling the production of mucus by diatoms under the constraint of a nutrient limitation. The initial questioning comes from the observation of the so-called "liga" on the Aquitaine coast. The biological origin of the phenomenon is presented and discussed based on the existing litterature. A mathematical model is proposed and its theoretical properties are analized: well-posedness and differentiability with respect to the model parameters. Finally, numerical experiments are provided, investigating the possibility of parameter identification for the model using chemostat-type experiments.
Keywords: Phytoplancton, ecological models, diatoms, mucilage, mathematical model, system of differential equations, parameter identification.
AMScodes : 37N25,92D25,92D40.

1 Introduction

The presence of pelagic muds composed of marine mucilage is well documented in the Adriatic sea (see *e.g.* [2, 19, 25, 51, 26]) and the sea of Marmara [1, 6], it is also observed on the Basque coast of the Atlantic ocean, where it is locally called *liga* [4]. Formerly occasionally observed on the Basque coast in spring these last few decades, the frequency, abundance and duration of liga have been steadily increasing since the early 2000's. This sharp increase of the phenomenon is reported by fishermen, whose nets get clogged by liga. This degradation of water quality shows a coastal ecosystem in poor health. Excessive marine mucilage production is related to blooms of phytoplankton algae, which can be harmful (in some cases leading to toxin releases) and can have negative impacts on the local ecosystem, as well as on fishing and aquaculture activities.

As indicated in [67], liga is a term of Basque origin derived from *"ligarda"*: a sticky liquid. Liga is a pelagic marine mucilage of thick and gelatinous aspect forming macroscopic aggregates. Aggregation is caused by *TEP* (Transparent Exopolymer Particles): extra-cellular organic particles, see [49]. The presence of such aggregates is highly correlated with a class of algae: diatoms, a unicellular phytoplankton species. Actually, large concentration of various species of diatoms are systematically present in liga samples as analyzed in [16, 67], the relationship between liga, TEP and diatom blooms will be discussed further in section 2.1.
Liga formation occurs in areas where salt and fresh waters mix together. Early liga production stage corresponds to pelagic colloidal aggregates, with the appearance of a mud as observed on fishing nets where it accumulates. Under stable anticyclonic conditions, it has time to mature to a pelagic mucous mass that may migrate in the water column. Depending on hydro-climatic

*charles.pierre@univ-pau.fr, LMAP UMR- CNRS 5142, IPRA BP 1155, 64013 Pau Cedex, France.
†guy.vallet@univ-pau.fr, LMAP UMR- CNRS 5142, IPRA BP 1155, 64013 Pau Cedex, France.

conditions and on its biological composition, it may either settle the sea bottom as a mucous mass or reach the surface as floating mosses. Liga forms temporary pelagic ecosystems containing a large diversity of microorganisms (viruses, bacteria, phyto and zoo-plankton). Two blooms of liga are usually observed every year: in spring and autumn. It disappears with the first winter storms but can reappear with a low intensity during winter anticyclonic conditions.

Liga production scenario is the following. Under suitable environmental conditions, a bloom (or an important increase) of diatoms occurs, followed by mucilage secretion. When the density of population and the mucus concentration are important enough, it can aggregate and form liga, see *e.g.* [49, Supplementary Material S2)]).
Research on mucilage production by diatoms such as [57] indicates that it might be induced by an imbalance in intracellular nutrient concentrations and be no more than a homeostatic mechanism to balance cell stoichiometry. This assumption is discussed in section 2.2 and an imbalance in nutrients C:N ratio is put forward.
The aim of this paper is to propose a model of liga production by diatom populations. The assumption of a C:N imbalance as a trigger of mucilage production is considered and a two nutrient model will be proposed based on carbon and nitrogen uptakes by diatoms.

This paper is organized as follows. Section 2 is devoted to a presentation of diatoms, their blooms and the causes of liga. A mathematical model of liga production is presented in section 3. This will be followed by a mathematical study of this model: existence and uniqueness of the solution, some qualitative information and the stability of the solution with respect to the parameters of the model. The last section will be devoted to the numerical study of the model. Firstly a biological setting of the model inducing mucilage production is introduced. Then the ability to perform a parameter identification of the model based on chemostat experiments is analyzed.

2 Diatoms

Diatoms are brown unicellular phytoplankton algae with a siliceous shell called frustule, a bipartite silica cell wall that encloses the eukaryotic protoplast. It protects the protoplast and provides routes for nutrient uptake, gaseous exchanges and extra-cellular secretions.
In most cases, diatom life cycle alternates a vegetative phase (lasting months or years) and a rejuvenation phase (few days) usually preceded by sexual reproduction [61, 63].
Diatoms have one or two cytoplasmic vacuoles with multiple roles. They have reserves that are easily mobilized storage in view of less favorable periods or shortages. The frustule provides a supporting structure to the vacuole and protects nutrients from competitors for subsequent cell divisions as noticed in [65].

2.1 Liga, TEP and diatom blooms

As mentioned in the introduction, mucilage aggregation is caused by TEP[1], see [49]. Diatom vacuoles constitute an important supply of nutrients, in particular $\beta - 1,3$-glucan chrysolaminarin, a carbohydrate, which, according to [60], contributes to the production of extracellular polymers. Indeed, diatoms exude *EPS* (Extra-cellular Polymeric Substances) constituting, under suitable conditions, precursors for the production of TEP, which (mixed with other organic matter particles) then transform to marine snow [35, 49, 54]. TEP are buoyant particles that should stay suspended in surface waters, and sinking is related to the size, the porosity and the density of aggregates. Ballasting can be due to incorporation of heavy minerals or atmospheric dusts [49], or modern pollution such as residue from spilled oil [14, 35]. In [9], the authors

[1]TEP: Transparent Exopolymeric Particles

noticed that diatoms also release poly-unsaturated aldehydes during blooms which contribute to increase TEP sizes and generate larger aggregates.

On the Basque coast, a first episode of liga usually appears in spring, a second in autumn, with mainly a bloom of diatoms [67]. It has been reported in [68] that the main species of diatoms found in bloom situations of liga are *Ceratoneis closterium* (always observed), *Leptocylindus danicus* (in June 2013), *Pseudo Nitzschia group B 2 & B1* (in October 2013) and *Thalassiosira gravida* (in march 2014). In case of a strong presence of liga, the following species have been found too: diatoms (in addition to those indicated above *Navicula spp.*, *Guinardia flaccida*, *Leptocylindrus minimus*, *Guinardia delicatula*, *Rhizosolenia setigera*); dinoflagelates (*Protoperidinium oceanicum*); gelatinous zooplankton (*Siphonophora Diphiidae*, *Oikopleura sp.*, *Sagitta sp*).

Diatom concentration is generally higher in the beginning of spring and in autumn, when nutrients start to be abundant again and when light intensity and day length are favorable. Furthermore, algal blooms are known to be often initiated by a diatom bloom, as observed in the Adriatic, where diatom *Cylindrotheca Closterium* is involved in extracellular carbohydrates production under nutrient limitation, see [2]. Several reasons can be pointed out for that. Firstly, [3] highlighted a functional cycle of urea conferring to diatoms an ecological advantage during periods of nitrogen limitations to other phytoplankton taxa [16, 46, 66]. Moreover, diatom brown color reveals the presence of carotenoid, utilized together with chlorophyll a and c, this ensures a large photosynthetic light harvesting. Eventually, diatoms are capable of C4 photosynthesis, allowing a more efficient utilization of available CO_2, in particular in less favorable conditions [61].

Diatom ability to grow more effectively than other species when environmental conditions get more favorable has also been investigated in terms of mathematical models and statistical data analysis. In [62, 65], it has been expressed by computing a growth rate which is higher for diatoms than other phytoplankton types. Based on this observation, Litchman *et. al.* [47] have been interested in modeling a coefficient of nutrient competitive uptake ability and estimated this coefficient for diatoms, coccolithophorids, dinoflagellates and green algae. By compiling a database of nutrient uptake and nutrient-dependent growth parameters, the authors have fitted the parameters of a model based on quota-Droop's law [27] and Michaelis-Menten-Monod [71] equation. Their analysis showed the ability of diatoms to respond more effectively to the return of good environmental conditions.

2.2 Stoichiometry and Redfield ratio

Redfield discovered in 1934 that the ratio of carbon, nitrogen and phosphorus (C:N:P) is nearly constant in oceans, in phytoplankton biomass as well as in dissolved nutrients. Concerning diatoms, the Redfield stoichiometry is usually given by $C : N : P = 106 : 16 : 1$ [13], even if it is known that it may fluctuate to become C:N:P = 163:22:1 [50, 43]. These values have to be understood as averages and may vary significantly locally. In particular, intracellular stoichiometry can strongly vary due to nutrient limitations [59].

Imbalance in stoichiometry can lead to many consequences. Algal growth rate can be impacted and abnormal secretions are induced as developed in [29]: *"When nutrient element X is in excess relative to requirements, consumers increase the specific rate of loss of that element (via excretion or egestion)"*. Resulting in particular in a synthesis and secretion of mucilages, as pointed out in the conclusion of [51], "*[...] the formation of mucilages is an alteration of the natural humification process that occurs when the organic matter degradation phase becomes slower than the synthesis phase. This alteration can [...] support the formation of refractory materials such as macro-aggregates and mucilages.*" Thus, together with a limitation of P stopping

cellular division, stoichiometry imbalance may be a hindrance to the development of bacteria in charge of the degradation of this organic matter (mucilage). The consequence is probably the development of liga.

2.2.1 Imbalance N:P

The excess of nitrogen in relation to phosphates is pointed out by [67] to be a possible trigger of liga production in a context of P limitation. A prevailing influence of nitrogen is also mentioned in situation of mucilage in the Adriatic sea in [17], in the Tasman Bay in [48], where a similar imbalance is observed but interpreted as an effect rather than a cause of TEP production.

In case of high level of N in comparaison to P, one is interested in polymeric secretions with a high content of N rather than P, whereas TEP have a composition poor in N. Meanwhile, CSP [2] are N-rich exopolymer particles that have recently been studied in [69]: the author notes that the C:N ratio for CSP is 3.8:1 (lower than Redfield's ratio of 6.6:1), though it can be estimated to be 26:1 for TEP.

If CSP can be understood as the excretion of an excess of N, it is remarked in [49] that contrary to TEP, CSP do not seem to significantly impact aggregation processes and may not play the same pivotal role in carbon cycling as TEP. In conclusion, abnormal TEP secretion inducing liga formation may rather be searched in a context of P limitation and of C:P or C:N imbalances.

2.2.2 Imbalance C:P or C:N

From a geological point of view, marine diatoms have been at the front line of the regulation of atmospheric partial pressure of carbon dioxide (pCO_2), in particular by reducing atmospheric carbon dioxide levels [20]. Accumulations of diatom-rich sediments extend back to the late Cretaceous. More recently, diatom-rich lake sediments are from the Eocene, with large deposits found in the Miocene [73]. Harwood *et. al.* [37] dated the emergence of planktonic diatom species at the end of the early cretaceous (Albian), a marine diatom extinction during the first part of the late cretaceous (Cenomanian/Turonian) and the beginning of pelagic diatom deposits at the early Campanian. However, these dates may be older [36, 53] and could be linked to local peaks of partial pressure of carbon dioxide identifiable during these periods [8]. Afterwards, diatom population growth declined during the Cenozoic era [56] with a global decreasing of the partial pressure of carbon dioxide perturbed by local peaks [10].

Oceans absorb huge quantities of inorganic carbon due to important anthropic emissions of CO_2 into the atmosphere, increasing the stoichiometry of carbon to nitrogen. This has implications for a variety of marine biological and biogeochemical processes, and underscores the importance of biologically driven feedbacks in the ocean to global climate change. It is pointed out in [58] that for the same uptake of inorganic nutrients, the production of organic matter increases significantly with higher CO_2 partial pressures, probably in the form of TEP.

Many authors related strong correlations between TEP production by plankton assemblages dominated by diatoms and cyanobacteria, and CO_2 concentration [7, 11, 23, 38, 39, 40, 49, 62, 70, 72] with different interpretations of the usefulness of these secretions: the excretion of an excess to rebalance Redfield's stoichiometry, buoyancy control, grazing repellents or to attract bacteria in charge of mucus degradation.

Following [49], the theoretical Redfield's ratio C:N:P=106:16:1 is strongly related to the

[2] CSP: Coomassie Stainable Particles

equation of production of phytoplankton biomass given by,

$$106\,CO_2 + 122\,H_2O + 16\,HNO_3 + H_3PO_4 \longrightarrow (CH_2O)_{106}(NH_3)_{16}(H_3PO_4) + 138\,O_2$$

whereas, without particular link to Redfield's ratio, the equation of production of TEP[3] is given by

$$106\,CO_2 + 105\,H_2O \longrightarrow C_{106}\,H_{201}O_{105} + 106\,O_2.$$

An imbalance C:P or C:N then could trigger TEP production by diatoms to rebalance their Redfield's ratio. This has been observed in [49]: under optimal growth conditions, elevated pCO_2 leads to increased TEP production by diatom *Skeletonema costatum*.

Of course, reality is more complex. It is in particular species dependent and in [49] the same experiment for diatom *Thalassiosira weissflogii* did not change TEP production (which however increased at sub-optimal irradiance and growth temperature). The variation of TEP production can have other origins: nutrient limitation (cellular carbon overflow), temperature (an increase of temperature can enhance extracellular releases), see [49].

3 Model statement

Models of marine ecosystem require to consider a large number of variables such as different types of phytoplankton and of grazers, of nutrients, environmental conditions or photosynthesis models. Its vocation is to describe biological diversity over a rather large period of time. We refer *e.ġ.* to [28, 32, 18] and [31] for such classic models in plankton ecology, which however do not model mucilage production.

Our aim is to propose a mathematical model that describes the evolution of a phytoplankton population during a period of bloom and the corresponding evolution of TEP production. Up to a prescribed threshold of TEP concentration, the first stage of liga (as described in introduction section) will be considered to be reached. Pelagic ecosystems being very complex, only the key features of our understanding of liga production will be kept, leading to an inevitable and drastic simplification of reality. In the light of what has been said in previous sections, diatom phytoplankton species will be considered to be at the origin of TEP production, in response to an imbalance of C:N stoichiometry during a bloom of population with P limitation. Predation will be neglected: during a diatom bloom, predators are expected to grow with a delay (ecological advantages of diatoms in terms of growth are developed in section 2.1). More generally let us emphasize again that the goal is to model the first stage leading to liga production and not to study the global ecosystem of liga.

Phytoplankton population models classically rely on *quota*. The basic tools are: Michaelis-Menten-Monod rule (see [71]) for the uptake of nutrients and Droop-Liebig's law (see [27]) to specify the impact of quotas and limitation on the phytoplankton growth.
Let us briefly recall these concepts. For a given species with concentration D and a given nutrient with concentration N, the associated *quota* is defined by $Q = \frac{N}{D}$.
The quota-Droop's law models a net-growth rate by $\mu_\infty\left(1 - \frac{Q_{\min}}{Q}\right) - m$ (where $Q_{\min}$ is the minimum quota, μ_∞ is the growth rate of species at infinite quota, m is the mortality rate).
The Michaelis-Menten-Monod equation defines a nutrient uptake rate by $V_{\max}\frac{N}{K+N}$ (where $V_{\max}$ is the maximum nutrient uptake rate, K is the half-saturation constant for nutrient uptake).

[3] under cover of the acceptance of its chemical writing

3.1 Discussion of some mathematical models

In this section are presented a selection of plankton models found in the literature. Without attempting to be comprehensive, we want to see how to adapt these models to go towards liga modeling.
In [45], a phytoplankton growth model relying on multiple nutrients is proposed, including nutrient limitation and stoichiometry information. In [5] a similar model is inserted in a more complex setting involving diffusion/convection, light, sinking rates and predators.
The mathematical model in [45] has been confronted to experimental data from chemostat[4] experiments in [45]. The authors concluded that their results are relatively satisfactory, except for phytoplankton stoichiometry during exponential growth at high nutrient levels. Two reasons can be raised for that. Firstly, no process of extra-cellular secretion is modeled, which would allow to eliminate a nutrient excess. Secondly, in the quota equilibrium equation, no consumption negative term for growth is present excepted cell-division.
In [55], a predator prey model is proposed including one algal and one grazer species and a nutrient in excess. A waste of up-taken nutrient not used for growth is added, as in [5], with a feedback from the producer biomass onto nutrient concentration. Meanwhile, no quota-decrease by cell division is taken into account in the model. It is not obvious at all that an excess of up-taken nutrient may return as available nutrients as proposed by [5, 55]. Consider for instance carbon: it is up-taken by the phytoplankton as inorganic carbon and then transformed, after photosynthesis, into polysaccharides and cannot be secreted back to the extra-cellular media as inorganic carbon to be up-taken again. Alternatively, being born in mind the importance of stoichiometry imbalances previously discussed in section 2.2, it seems necessary to insert in the model a possible secretion of a given nutrient in excess, especially when another nutrient is limited. This certainly helps to explain the problem of phytoplankton stoichiometry during exponential growth in [44].

Thus, it appears as necessary to add to the model one or more additional variables corresponding to secretion of EPS (Extracellular polymeric substances) when given nutrients are in surplus: TEP concerning C or CSP for N and C.

3.2 A two nutrient model

A chemostat experiment is considered with an inflow/outflow flushing rate given by $a > 0$. Diatom concentration (the biomass producer in our case) is denoted D. Two nutrients are considered, nitrogen and carbon, the associated concentration are N and C in the extra cellular media, and P_N and P_C in the phytoplankton. The associated quota are $Q_N = \frac{P_N}{D}$, $Q_C = \frac{P_C}{D}$ and M will denote EPS concentration.

Let the chemostat inlet have a nitrogen concentration N_{in}, the inflow/outflow nitrogen balance is $a(N_{\text{in}} - N)$. The diatom uptake rate in nitrogen is denoted τ_N and in the absence of other species in the chemostat, the evolution law for nutrient N is,

$$\frac{\mathrm{d}N}{\mathrm{d}t} = a(N_{\text{in}} - N) - \tau_N D$$

and similarly for C,

$$\frac{\mathrm{d}C}{\mathrm{d}t} = a(C_{\text{in}} - C) - \tau_C D$$

Denoting σ_N the consumption rate of N by diatoms (*e.g.* for growth) and by m_D the death-rate of diatoms, the concentration of N in the diatoms is ruled by

$$\frac{\mathrm{d}P_N}{\mathrm{d}t} = (\tau_N - \sigma_N)D - (a + m_D)P_N,$$

[4] chemostat: a system for the culture of microorganisms with controlled chemical composition.

and similarly for C,

$$\frac{\mathrm{d}P_C}{\mathrm{d}t} = (\tau_C - \sigma_C)D - (a + m_D)P_C.$$

The concentration in diatoms will be given by,

$$\frac{\mathrm{d}D}{\mathrm{d}t} = \tau_D D - (a + m_D)D,$$

where τ_D is the growth rate of the diatom population. Finally, mucilage concentration satisfies,

$$\frac{\mathrm{d}M}{\mathrm{d}t} = \tau_M D - aM,$$

where τ_M is the secretion rate of EPS.
In terms of quota, differentiating their definition $\frac{\mathrm{d}Q_N}{\mathrm{d}t} = \frac{1}{D}\frac{\mathrm{d}P_N}{\mathrm{d}t} - \frac{P_N}{D^2}\frac{\mathrm{d}D}{\mathrm{d}t}$,

$$\begin{aligned}\frac{\mathrm{d}Q_N}{\mathrm{d}t} =& \tau_N - \sigma_N - Q_N\tau_D,\\ \frac{\mathrm{d}Q_C}{\mathrm{d}t} =& \tau_C - \sigma_C - Q_C\tau_D,\end{aligned}$$

The uptake rate is given by Michaelis-Menten-Monod law as introduced in section 3.1,

$$\tau_N = \tau_N(t, N) = V^N_{\max}(t)\frac{N}{K_N(t) + N}, \tag{1}$$

$$\tau_C = \tau_C(t, C) = V^C_{\max}(t)\frac{C}{K_C(t) + C}. \tag{2}$$

In the sequel x^+ denotes the positive part of the real number x: $x^+ = \max(0, x)$. Cell division can take place if the quota of each nutrient is greater than a threshold: $Q_C \geq Q^C_{\min} > 0$ and $Q_N \geq Q^N_{\min} > 0$. Then, following [45], Droop's growth function coupled with Liebig's law yields

$$\tau_D = \tau_D(t, Q_C, Q_N) = \mu_D(t)\min\left((1 - \frac{Q^C_{\min}}{Q_C})^+,\ (1 - \frac{Q^N_{\min}}{Q_N})^+\right) \tag{3}$$

where μ_D is the maximal division rate at infinite quota. In *normal conditions*, the diatom stoichiometry ratio $C : N$ is a positive constant denoted α: diatoms will use α units of C for one unit of N. Considering that EPS release is based on the presence of extra C compared to N, *i.e.* if $Q_C > \alpha Q_N$ and, of course, if $Q_C > Q^C_{\min}$, the EPS release rate τ_M will be modeled by,

$$\tau_M = \tau_M(t, Q_C, Q_N) = \mu_M(t)\min\left((1 - \frac{Q^C_{\min}}{Q_C})^+,\ (1 - \frac{\alpha Q_N}{Q_C})^+\right), \tag{4}$$

with μ_M the maximal release rate of mucilage by diatoms.

Concerning nutrient consumption,

$$\sigma_N = \sigma_N(t, Q_C, Q_N) = \Theta_D(t)\min\left((1 - \frac{Q^C_{\min}}{Q_C})^+,\ (1 - \frac{Q^N_{\min}}{Q_N})^+\right), \tag{5}$$

where Θ_D the maximal consumption rate of diatoms for growth at infinite quota.
Biomass production for the growth of diatom population respects the stoichiometry α units of C for one unit of N, thus the carbon consumption rate is split into two parts, one for diatom growth and the second for mucilage production,

$$\sigma_C = \alpha\,\sigma_N + \sigma_M$$

with,

$$\sigma_M = \sigma_M(t, Q_C, Q_N) = \Theta_M(t) \min\left((1 - \frac{Q^C_{\min}}{Q_C})^+,\ (1 - \frac{\alpha Q_N}{Q_C})^+\right), \tag{6}$$

where Θ_M is the maximal consumption rate of diatoms for mucilage.

In short, we have the system:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = F(t, X), \tag{S}$$

where the model variable have been gathered in the column vector $X = (N, C, Q_N, Q_C, D, M)^T$ and where the function F summarizes as,

$$\begin{aligned}
\frac{\mathrm{d}N}{\mathrm{d}t} &= a(N_{\text{in}}(t) - N) - \tau_N(t, N)D \ :=\ F_1(t, X),\\
\frac{\mathrm{d}C}{\mathrm{d}t} &= a(C_{\text{in}}(t) - C) - \tau_C(t, C)D \ :=\ F_2(t, X),\\
\frac{\mathrm{d}Q_N}{\mathrm{d}t} &= \tau_N(t, N) - \sigma_N(t, Q_C, Q_N) - \tau_D(t, Q_C, Q_N)Q_N \ :=\ F_3(t, X),\\
\frac{\mathrm{d}Q_C}{\mathrm{d}t} &= \tau_C(t, C) - \alpha\,\sigma_N(t, Q_C, Q_N) - \sigma_M(t, Q_C, Q_N) - \tau_D(t, Q_C, Q_N)Q_C \ :=\ F_4(t, X),\\
\frac{\mathrm{d}D}{\mathrm{d}t} &= \tau_D(t, Q_C, Q_N)D - (a + m_D(t))D \ :=\ F_5(t, X),\\
\frac{\mathrm{d}M}{\mathrm{d}t} &= \tau_M(t, Q_C, Q_N)D - aM \ :=\ F_6(t, X).
\end{aligned}$$

Functions τ_N ,τ_C, τ_D, τ_M, σ_N and σ_C (given by (1) to (6) respectively), depend on $K_N(t)$, $K_C(t)$, $V^N_{\max}(t)$, $V^C_{\max}(t)$, $\mu_D(t)$, $\mu_M(t)$, $\Theta_D(t)$ and $\Theta_M(t)$. These parameters do *a priori* depend on the time t (since they are functions of lightness and temperature, see *e.g.* [24]), although they will be set to constants in numerical experiments. The same remark holds for $N_{\text{in}}(t)$, $C_{\text{in}}(t)$ and $m_D(t)$. The model finally also depends on the constant parameters a, $Q^N_{\min}$, $Q^C_{\min}$ and α.

4 Model mathematical analysis

For technical reasons, but fully compatible with biological modeling, the following assumptions are made on the parameters of system (S),

$$\exists \mathcal{M} > 0, \quad a, Q^C_{\min}, Q^N_{\min}, \alpha \in (\frac{1}{\mathcal{M}}, \mathcal{M}), \quad m_D \in C\Big([0, +\infty), (0, \mathcal{M})\Big), \tag{7}$$

$$N_{\text{in}}, C_{\text{in}}, V^N_{\max}, K_N, V^C_{\max}, K_C, \mu_D, \mu_M, \Theta_D, \Theta_M \in C\Big([0, +\infty), (\frac{1}{\mathcal{M}}, \mathcal{M})\Big). \tag{8}$$

On the contrary of the other parameter functions, the mortality rate m_D is allowed to vanish, which is relevant in the present context of a chemostat-type experiment.

Let us fix what we mean by a solution to the problem.

Definition 1. A solution to system (S) is any absolutely continuous function X with non negative values, solution to the ordinary differential equation $\frac{\mathrm{d}X}{\mathrm{d}t} = F(t, X)$ associated with the initial condition $X(0) = X_0$ where $X_0 \geq 0$.
For convenience, for $X \in \mathbb{R}^6$, $X \geq 0$ means that $X_i \geq 0$ for $1 \leq i \leq 6$.

Our main result is the following.

Theorem 1. Under assumptions (7) and (8),

1. there exists a unique maximal solution to system (S) in the sense of definition 1,
2. this unique solution is a global solution (defined for $t \in |0, +\infty)$),
3. if there exists a time $t_0 \geq 0$ such that $X_3(t_0) = Q_N(t_0) \geq Q^N_{\min}$ then $Q_N(t) \geq Q^N_{\min}$ for any $t \geq t_0$. Similarly, if $X_4(t_0) = Q_C(t_0) \geq Q^C_{\min}$ then $Q_C(t) \geq Q^C_{\min}$ for any $t \geq t_0$ (in other words, the solution remains *biologically relevant*).

Moreover, in the case of constant parameters, the following regularity with respect to the parameters holds. Denote

$$A = (a, N_{\text{in}}, C_{\text{in}}, V^N_{\max}, K_N, V^C_{\max}, K_C, Q^C_{\min}, Q^N_{\min}, \alpha, \mu_D, \mu_M, \Theta_D, \Theta_M, m_D),$$

a vector of constant parameters in $\mathcal{A} = (\frac{1}{\mathcal{M}}, \mathcal{M})^{15}$. Consider a given initial condition $X_0 \geq 0$ to problem (9).

4. For any $T > 0$, the unique solution $X(t, A)$, as a function of the time variable $t \in [0, T]$ and of the parameters $A \in \overline{\mathcal{A}}$ is weakly differentiable,

$$X \in W^{1,\infty}\left((0, T) \times \mathcal{A}\right)^6,$$

and $t \mapsto X(t, \cdot) \in C^0(\overline{\mathcal{A}})^6$ is continuously differentiable in t,

$$X \in C^1([0, T], C^0(\overline{\mathcal{A}})^6),$$

and,

$$X \in \text{Lip}([0, T], W^{1,\infty}(\mathcal{A})^6).$$

$\nabla_A X$ is the unique solution to the linear ordinary differential equation obtained by differentiation with respect to A of system (S),

$$\frac{\mathrm{d}}{\mathrm{d}t}\nabla_A X(t, A) = \frac{\partial F}{\partial X}\big(X(t, A), A\big)\nabla_A X(t, A) \; + \; \frac{\partial F}{\partial A}\big(X(t, A), A\big),$$

with the initial condition $(\nabla_A X)_0 = 0$.

The proof of Theorem 1 has been split into lemma 2, 3 & 4 for points 1 & 2 and lemma 5 for point 3. Then, point 4 is detailed in lemmas 7, 8 and 9.

Finally, in the last section 5, the differentiability properties enunciated in theorem 1 will be used to propose a method of parameter identification for model (9) based on a Gauss Newton method:

- a realistic choice of parameters is discussed, illustrated by direct simulations showing the model dynamics,
- then, a numerical methodology for identifying the problem parameters is presented, supported by numerical output statistics.

4.1 Well-posedness

To address the question of the positivity of the solutions of system (S), we study the Cauchy problem,

$$\frac{\mathrm{d}X}{\mathrm{d}t} = F(t, X^+), \quad X(0) = X_0. \tag{9}$$

It will be proven in this section that problem (9) is well posed with global solutions for all $t \geq 0$. Moreover its solutions will satisfy $X(t) \geq 0$ if $X_0 \geq 0$ and so will coincide with system (S) solutions.

In the sequel, we generically denote $C > 0$ a positive constant only depending on the

parameter $\mathcal{M}$ in assumptions (7) (8), whose value however may differ from one situation to another.

Let us note that for positive $Q^C_{\min}$ and $Q^N_{\min}$, and nonnegative Q_C and Q_N,

$$\min\left((1-\frac{Q^C_{\min}}{Q_C})^+,\ (1-\frac{Q^N_{\min}}{Q_N})^+\right)=\min\left(1-\frac{Q^C_{\min}}{\max(Q^C_{\min},Q_C)},\ 1-\frac{Q^N_{\min}}{\max(Q^N_{\min},Q_N)}\right),$$

and, noticing that $(1-\dfrac{Q^C_{\min}}{Q_C})^+=0$ for any $0\le Q_C\le Q^C_{\min}$, that

$$\min\left((1-\frac{Q^C_{\min}}{Q_C})^+,\ (1-\frac{\alpha Q_N}{Q_C})^+\right)=\min\left(1-\frac{Q^C_{\min}}{\max(Q^C_{\min},Q_C)},\ 1-\frac{\alpha Q_N}{\max(\alpha Q_N,Q_C)}\right),$$

with the convention that $\dfrac{\alpha Q_N}{\max(\alpha Q_N,Q_C)}=0$ if $Q_N=Q_C=0$. Then, the following function $\mathbb{A}$, defined *a priori* for any $(a,b,c,d)\in\mathbb{R}^4$ by

$$\mathbb{A}(a,b,c,d)=\min\left(1-\frac{a^+}{\max(a^+,b^+)},1-\frac{c^+}{\max(c^+,d^+)}\right)\ \text{ with convention }\ \frac{0}{0}=0, \tag{10}$$

is introduced for the study of problem (9), it satisfies the following lemma.

Lemma 1.

- For any $(a,b,c,d)\in\mathbb{R}^4$, $0\le\mathbb{A}(a,b,c,d)\le 1$.
- Let $\delta>0$, $(a_i,b_i,c_i,d_i)\in\mathbb{R}^4$ with $a_i,c_i\ge\delta$, $i=1,2$, then,

$$\Big|\mathbb{A}(a_1,b_1,c_1,d_1)-\mathbb{A}(a_2,b_2,c_2,d_2)\Big|\le 2\frac{|a_1-a_2|+|c_1-c_2|}{\delta}+\frac{|b_1-b_2|+|d_1-d_2|}{\delta}.$$

- Let $\delta>0$, $(a_i,b_i,c_i)\in\mathbb{R}^4$ with $a_i\ge\delta$, $i=1,2$, then,

$$\Big|\mathbb{A}(a_1,b_1,c_1,b_1)-\mathbb{A}(a_2,b_2,c_2,b_2)\Big|\le 2\frac{|a_1-a_2|+|c_1-c_2|+|b_1-b_2|}{\delta}.$$

Proof. Consider $(a_i,b_i,c_i,d_i)\in[\delta,+\infty[\times\mathbb{R}^+\times[\delta,+\infty[\times\mathbb{R}^+$ for a given positive δ, $i=1$ and 2. Using the following inequality on the min and max functions, $|\min(x,y)-\min(a,b)|\le|x-a|+|y-b|$ and $|\max(x,y)-\max(a,b)|\le|x-a|+|y-b|$, one has

$$\begin{aligned}
&\Big|\mathbb{A}(a_1,b_1,c_1,d_1)-\mathbb{A}(a_2,b_2,c_2,d_2)\Big|\\
=&\Big|\min\left(1-\frac{a_1}{\max(a_1,b_1)},1-\frac{c_1}{\max(c_1,d_1)}\right)-\min\left(1-\frac{a_2}{\max(a_2,b_2)},1-\frac{c_2}{\max(c_2,d_2)}\right)\Big|\\
\le&\Big|\frac{a_1}{\max(a_1,b_1)}-\frac{a_2}{\max(a_2,b_2)}\Big|+\Big|\frac{c_1}{\max(c_1,d_1)}-\frac{c_2}{\max(c_2,d_2)}\Big|\\
\le&\Big|\frac{(a_1-a_2)\max(a_2,b_2)+a_2[\max(a_2,b_2)-\max(a_1,b_1)]}{\max(a_1,b_1)\max(a_2,b_2)}\Big|+\Big|\frac{c_1}{\max(c_1,d_1)}-\frac{c_2}{\max(c_2,d_2)}\Big|\\
\le&\frac{|a_1-a_2|}{\max(a_1,b_1)}+\frac{|\max(a_2,b_2)-\max(a_1,b_1)|}{\max(a_1,b_1)}+\frac{|c_1-c_2|}{\max(c_1,d_1)}+\frac{|\max(c_2,d_2)-\max(c_1,d_1)|}{\max(c_1,d_1)}\\
\le&\frac{2}{\delta}\Big[|a_1-a_2|+|c_1-c_2|\Big]+\frac{1}{\delta}\Big[|b_1-b_2|+|d_1-d_2|\Big].
\end{aligned}$$

In case of signed variables, the inequality holds with a_i^+,b_i^+,c_i^+,d_i^+ (i=1,2) on the right hand side and the result holds since $|x^+-y^+|\le|x-y|$ for any real x,y.

Now consider the second inequality, first for non negative variables. It is here assumed that $a_i\ge\delta$. In the particular case where $b_i>a_i$, one has $b_1,b_2\ge\delta$ and the second inequality is induced by the first one. Now if $b_i\le a_i$ then $\mathbb{A}(a_i,b_i,c_i,b_i)=0$, we therefore only have to investigate the case $b_1\le a_1$ and $b_2>a_2$ (the fourth case being similar),

$$\Big|\mathbb{A}(a_1,b_1,c_1,b_1)-\mathbb{A}(a_2,b_2,c_2,b_2)\Big|=\mathbb{A}(a_2,b_2,c_2,b_2)\le 1-\frac{a_2}{b_2},$$

and, since $b_1 - a_1 \leq 0$,

$$1 - \frac{a_2}{b_2} = \frac{b_2 - b_1 + (b_1 - a_1) + a_1 - a_2}{b_2} \leq \frac{|b_2 - b_1| + |a_1 - a_2|}{\delta}.$$

The case of signed variables is similar as above. □

The well-posedness of the Cauchy problem (9) is stated in the following lemma.

Lemma 2. With assumptions (7) and (8), the function $(t, X) \mapsto F(t, X^+)$ is continuous and locally Lipschitz with respect to its second variable X on $[0, +\infty) \times \mathbb{R}^6$.

As a consequence, Picard–Lindelöf theorem yields a result of existence and uniqueness of the maximal solution to the Cauchy problem (9), for any initial condition $X_0 \in \mathbb{R}^6$.

Proof. It will be denoted here $F_i^+ : (t, X) \mapsto F_i(t, X^+)$.

$$\begin{aligned} F_1(t, X_1^+) - F_1(t, X_2^+) &= aN_2^+ - aN_1^+ + \tau_N(t, N_2^+)D_2^+ - \tau_N(t, N_1^+)D_1^+ \\ &= a(N_2^+ - N_1^+) + D_2^+(\tau_N(t, N_2^+) - \tau_N(t, N_1^+)) \\ &\quad + \tau_N(t, N_1^+)(D_2^+ - D_1^+). \end{aligned}$$

Thanks to assumptions (7), $\tau_N(t, N^+)$ in (1) is bounded and globally Lipschitz in N on $[0, +\infty) \times [0, +\infty)$ with a constant only depending on $\mathcal{M}$. Then,

$$\left|F_1(t, X_1^+) - F_1(t, X_2^+)\right| \leq C(1 + |D_2|)\left(|N_1 - N_2| + |D_1 - D_2|\right),$$

and F_1^+ is continuous and locally Lipschitz in X on $[0, +\infty) \times \mathbb{R}^6$, as well as F_2^+ and F_5^+ with a similar proof.

Let us now consider $F_3^+ : (t, X) \mapsto F_3(t, X^+)$, we recall that $F_3(t, X^+) = \tau_N(t, N^+) - \sigma_N(t, Q_C^+, Q_N^+) - \tau_D(t, Q_C^+, Q_N^+)Q_N^+$. We just stated that the first term is globally Lipschitz in N. With definition (10) of $\mathbb{A}$ we get,

$$\begin{aligned} \sigma_N(t, Q_C^+, Q_N^+) &= \Theta_D(t)\mathbb{A}(Q_{\min}^C, Q_C, Q_{\min}^N, Q_N), \\ \tau_D(t, Q_C^+, Q_N^+) &= \mu_D(t)\mathbb{A}(Q_{\min}^C, Q_C, Q_{\min}^N, Q_N). \end{aligned}$$

With assumption (8), Θ_D and μ_D are bounded and with lemma 1: $\mathbb{A}(Q_{\min}^C, Q_C, Q_{\min}^N, Q_N) \leq 1$ and,

$$\begin{aligned} &\left|\mathbb{A}(Q_{\min}^C, Q_{C,1}, Q_{\min}^N, Q_{N,1}) - \mathbb{A}(Q_{\min}^C, Q_{C,2}, Q_{\min}^N, Q_{N,2})\right| \\ &\qquad \leq C\left(|Q_{C,1} - Q_{C,2}| + |Q_{N,1} - Q_{N,2}|\right). \end{aligned}$$

It follows that,

$$|\sigma_N(t, Q_{C,1}, Q_{N,1}) - \sigma_N(t, (Q_{C,2}, Q_{N,2})| \leq C\left(|Q_{C,1} - Q_{C,2}| + |Q_{N,1} - Q_{N,2}|\right),$$

and,

$$\begin{aligned} &\left|\tau_D(t, Q_{C,1}^+, Q_{N,1}^+)Q_{N,1}^+ - \tau_D(t, Q_{C,2}^+, Q_{N_2}^+)Q_{N,2}^+\right| \\ &\quad \leq |\tau_D(t, Q_{C,1}^+, Q_{N,1}^+)|\,|Q_{N,1}^+ - Q_{N,2}^+| + \left|\tau_D(t, Q_{C,1}^+, Q_{N,1}^+) - \tau_D(t, Q_{C,2}^+, Q_{N_2}^+)\right|\,|Q_{N,1}^+| \\ &\quad \leq C\,|Q_{N,1} - Q_{N,2}| + C\left(|Q_{C,1} - Q_{C,2}| + |Q_{N,1} - Q_{N,2}|\right)|Q_{N,1}|. \end{aligned}$$

Altogether, this prove that F_3^+ is continuous and locally Lipschitz in X on $[0, +\infty) \times \mathbb{R}^6$.

We now consider F_4^+, we recall that $F_4^+(t, X) = g_4(t, X) - \sigma_M(t, Q_C^+, Q_N^+)$ with $g_4(t, X^+) := \tau_C(t, C^+) - \alpha\sigma_N(t, Q_C^+, Q_N^+) - \tau_D(t, Q_C^+, Q_N^+)Q_C^+$.

From what has been previously developed , g_4 is locally Lipschitz in X, and it remains to study the term,

$$\sigma_M(t, Q_C^+, Q_N^+) = \Theta_M(t)\mathbb{A}(Q_{\min}^C, Q_C, \alpha Q_N^+, Q_C).$$

As Θ_M is bounded with assumption (8), this latter is Lipschitz continuous in X thanks to the last point of lemma 1. So, F_4^+ is continuous and locally Lipschitz in X on $[0, +\infty) \times \mathbb{R}^6$, as well as F_6^+ with a similar proof. □

Let us now prove positivity properties for the solutions of problem (9): it implies the equivalence between formulations (S) and (9) for non-negative initial conditions X_0.

Lemma 3. With assumptions (7) and (8), the unique maximal solution X to problem (9) for an initial condition $X_0 \geq 0$ remains non-negative: $X(t) \geq 0$ as long as it exists.
This is the unique solution of system (S) in the sense of definition 1.

Proof. Consider a solution X to problem (9) for $X_0 \geq 0$. Let us prove that for any of its coordinate i, $X_i^-(t) = 0$ where $X_i^- = \max(0, -X_i)$ is the negative part of X_i. Thanks to the chain rule in Sobolev spaces $\dfrac{\mathrm{d}X_i^-}{\mathrm{d}t} = -\dfrac{\mathrm{d}X_i}{\mathrm{d}t}\ 1_{\{X_i \leq 0\}} = -F_i(\cdot, X^+)\ 1_{\{X_i \leq 0\}}$ for a.e. time t. Now the key argument is to notice in system (S) that $F_i(\cdot, X^+) \geq 0$ if $X_i \leq 0$. It follows that $\frac{\mathrm{d}X_i^-}{\mathrm{d}t} \leq 0$, $X_i(0)^- = 0$ by assumption and so $X_i(t)^- = 0$. □

We now investigate global existence for solutions to problem (9).

Lemma 4. With assumptions (7) and (8), the maximal solution to problem (9) is a global solution.

Proof. The Euclidian norm on $\mathbb{R}^n$ is denoted $\|\cdot\|$. Consider a solution $X(t)$ to problem (S) for an initial condition $X_0 \geq 0$:

$$\frac{1}{2}\frac{\mathrm{d}\|X(t)\|^2}{\mathrm{d}t} = \frac{\mathrm{d}X(t)}{\mathrm{d}t} \cdot X(t) = F(t, X(t)) \cdot X(t),$$

and, by a sign argument (since $X(t) \geq 0$ with lemma 3),

$$\begin{aligned}
&F(t, X(t)) \cdot X(t) \leq aN_{\text{in}}N + aC_{\text{in}}C + V_{\max}^N \frac{NQ_N}{K_N + N} + V_{\max}^C \frac{CQ_C}{K_C + C} \\
&+ \mu_D \mathbb{A}(Q_{\min}^C, Q_C, Q_{\min}^N, Q_N)D^2 + \mu_M \mathbb{A}(Q_{\min}^C, Q_C, \alpha Q_N, Q_C)DM \\
\leq &a\|N_{\text{in}}\|_\infty^2 + aN^2 + a\|C_{\text{in}}\|_\infty^2 + aC^2 + \frac{\|V_{\max}^N\|_\infty}{\min_t K_N(t)} NQ_N + \frac{\|V_{\max}^C\|_\infty}{\min_t K_C(t)} CQ_C \\
&+ \|\mu_D\|_\infty D^2 + \|\mu_M\|_\infty DM \leq \frac{1}{2} C \left(1 + \|X(t)\|^2\right),
\end{aligned}$$

where C is a constant depending on $\mathcal{M}$ in assumptions (7) (8). Then Gronwall lemma yields,

$$|X(t)|^2 \leq e^{Ct} \left(|X_0|^2 + Ct\right). \tag{11}$$

Assume that the maximal solution $X(t)$ is not a global one. Thus, there exists $T > 0$ such that $\lim_{t \to T^-} ||X(t)|| = +\infty$, which contradicts the above inequality. □

4.2 Biologically relevant solutions

With lemma 3 and 4, system (S) associated with a non-negative initial condition has a unique solution which is non-negative and global. It is stated here that the quota remain larger than the minimum quota if this is true at initial time, *i.e.* biologically relevant initial condition are associated with biologically relevant solutions.

Lemma 5. With assumptions (7) and (8), consider the solution to problem (S) for an initial condition $X_0 \geq 0$.
If there exists t_0 such that $Q_N(t_0) \geq Q_{\min}^N$, then $Q_N(t) \geq Q_{\min}^N$ for any $t \geq t_0$.
Similarly, if there exists t_0 such that $Q_C(t_0) \geq Q_{\min}^C$, then $Q_C(t) \geq Q_{\min}^C$ for any $t \geq t_0$.

Remark 1. Note that
i) if $Q_N \leq Q_{\min}^N$ in $(0,T)$ then $Q_N(t) = Q_N(0) + \displaystyle\int_0^t \frac{V_{\max}^N(s)N(s)}{K_N(s)+N(s)}\,\mathrm{d}s$ if $t \leq T$.
ii) if $Q_C \leq Q_{\min}^C$ in $(0,T)$ then $Q_C(t) = Q_C(0) + \displaystyle\int_0^t \frac{V_{\max}^C(s)C(s)}{K_C(s)+C(s)}\,\mathrm{d}s$ if $t \leq T$.
iii) if $Q_N \leq Q_{\min}^N$ or $Q_C \leq Q_{\min}^C$ in $(0,T)$ then $D(t) = D(0)e^{-(a+m_D)t}$ in $(0,T)$.
Therefore, in case of N or C limitation, the diatoms are no longer able to divide and diatom population decreases and goes to 0 as long as this situations continues.

Proof. The proof is similar as for lemma 3. Consider the component $X_3 = Q_N$ of $X(t)$,

$$\begin{aligned}
&\frac{\mathrm{d}(Q_N - Q_{\min}^N)^-}{\mathrm{d}t} = -1_{\{Q_N < Q_{\min}^N\}}\frac{\mathrm{d}Q_N}{\mathrm{d}t},\\
&= -1_{\{Q_N < Q_{\min}^N\}}\Big[V_{\max}^N(t)\frac{N}{K_N(t)+N} - [\Theta_D(t) + \mu_D(t)Q_N]\mathbb{A}(Q_{\min}^C, Q_C, Q_{\min}^N, Q_N)\Big]\\
&= -1_{\{Q_N < Q_{\min}^N\}}V_{\max}^N(t)\frac{N}{K_N(t)+N} \leq 0.
\end{aligned}$$

Assume that $Q_N(t_0) \geq Q_{\min}^N$ for $t_0 \geq 0$. since $(Q_N - Q_{\min}^N)^-$ is absolutely continuous, the integration of the above inequality for $t \geq t_0$ yields,

$$(Q_N - Q_{\min}^N)^-(t) \leq (Q_N - Q_{\min}^N)^-(t_0) = 0,$$

and so $Q_N(t) \geq Q_{\min}^N$. Similarly,

$$\begin{aligned}
&\frac{\mathrm{d}(Q_C - Q_{\min}^C)^-}{\mathrm{d}t} = -1_{\{Q_C < Q_{\min}^C\}}\frac{\mathrm{d}Q_C}{\mathrm{d}t}\\
&= -1_{\{Q_C < Q_{\min}^C\}}\Big[V_{\max}^C(t)\frac{C}{K_C(t)+C} - [\alpha\Theta_D(t) + \mu_D(t)Q_C]\mathbb{A}(Q_{\min}^C, Q_C, Q_{\min}^N, Q_N)\\
&\qquad\qquad - \Theta_M(t)\mathbb{A}(Q_{\min}^C, Q_C, \alpha Q_N^+, Q_C)\Big]\\
&= -1_{\{Q_C < Q_{\min}^C\}}V_{\max}^C(t)\frac{C}{K_C(t)+C} \leq 0,
\end{aligned}$$

which leads to the same result. □

4.3 Stability with respect to parameters

This section deals with the stability of system (S) solutions with respect to its parameters. The case of constant parameters is studied on a fixed time interval $[0,T]$. As in theorem 1, the set of parameters is denoted by a vector $A \in \mathcal{A} = (\frac{1}{\mathcal{M}}, \mathcal{M})^{15}$, with $\mathcal{M}$ the bound in assumptions (7) (8),

$$A = (a, N_{\text{in}}, C_{\text{in}}, V_{\max}^N, K_N, V_{\max}^C, K_C, Q_{\min}^C, Q_{\min}^N, \alpha, \mu_D, \mu_M, \Theta_D, \Theta_M, m_D) \in \mathcal{A}.$$

Through this section,

- an initial condition $X_0 \geq 0$ and a final time $T > 0$ are fixed,
- $F : (Y, A) \in \mathbb{R}^6 \times \mathcal{A} \mapsto F(Y, A) \in \mathbb{R}^6$ denotes the function defining system (S),
- for any $t \in [0,T]$ and $A \in \mathcal{A}$, set $X(t,A)$ the value at time t of the solution to the Cauchy problem $\dfrac{\mathrm{d}}{\mathrm{d}t}Y = F(Y,A)$ associated with the initial condition $Y(0) = X_0$,

- with the previous lemma, this function is well defined, $X(t,A) \geq 0$ and with inequality (11), there exists a constant $\mathcal{K} > 0$ only depending on T and $\mathcal{M}$ such that $0 \leq X(t,A) \leq \mathcal{K}$,
- we then consider the function $X : [0,T] \times \mathcal{A} \to \mathbb{R}^6$ and we consider the usual notations: $X(t,\cdot) : \mathcal{A} \to \mathbb{R}^6$ is the function $A \in \mathcal{A} \mapsto X(t,A)$ and $X(\cdot,A) : [0,T] \to \mathbb{R}^6$ is the function $t \in [0,T] \mapsto X(t,A)$.

As a consequence, the fonction F of system (S) can be restricted to $(-\mathcal{K},\mathcal{K})^6 \times \mathcal{A}$, or its closure $[-\mathcal{K},\mathcal{K}]^6 \times \overline{\mathcal{A}}$ when a compact set is needed.

Lemma 6. There exists $\mathcal{L} > 0$, only depending on T and $\mathcal{M}$, such that the fonction F is $\mathcal{L}$-Lipschitz continuous on $[-\mathcal{K},\mathcal{K}]^6 \times \overline{\mathcal{A}}$.

Proof. It has already been proven in lemma 2 that F is locally Lipschitz in X, it is therefore globally Lipschitz in X on the compact $[-\mathcal{K},\mathcal{K}]^6 \times \overline{\mathcal{A}}$. Let us prove that it is also globally Lipschitz in A: the variable X is assumed to be in $[-\mathcal{K},\mathcal{K}]^6$ and two parameters $A_1, A_2 \in \overline{\mathcal{A}}$ are considered.
The function τ_N in (1) satisfies,

$$\big|\tau_N(X,A_1) - \tau_N(X,A_2)\big| \leq \mathcal{K}\mathcal{M}|V_{\max}^{N,1} - V_{\max}^{N,2}| + \mathcal{M}^3\mathcal{K}|K_N^1 - K_N^2|$$

and the same inequality holds for τ_C in (2) by changing N into C. Thus, F_1 and F_2 are globally Lipschitz in A. With lemma 1, the function τ_D in (3) satisfies,

$$\big|\tau_D(X,A_1) - \tau_D(X,A_2)\big| \leq |\mu_D^1 - \mu_D^2| + 2\mathcal{M}^2|Q_{\min}^{C,1} - Q_{\min}^{C,2}| + 2\mathcal{M}^2|Q_{\min}^{N,1} - Q_{\min}^{N,2}|,$$

and a similar inequality holds for σ_N in (5), consequently F_3 and F_5 are globally Lipschitz in A. Using again lemma 1, the function τ_M in (4) satisfies,

$$\big|\tau_M(X,A_1) - \tau_M(X,A_2)\big| \leq |\mu_M^1 - \mu_M^2| + 2\mathcal{M}^2|Q_{\min}^{C,1} - Q_{\min}^{C,2}|$$

and a similar inequality holds for σ_M in (6), consequently F_4 and F_6 are globally Lipschitz in A, which ends the proof. □

As a direct consequence, we have the following first regularity result.

Lemma 7. Set $\mathcal{Q} = (0,T) \times \mathcal{A}$, X and $\partial_t X$ are bounded Lipschitz-continuous functions on $\overline{\mathcal{Q}}$, and therefore belong to $W^{1,\infty}(\mathcal{Q})^6$ [30, Thm. 4.5].
In particular, the application $A \in \mathcal{A} \mapsto X(\cdot,A) \in [C^1([0,T])]^6$ is Lipschitz-continuous.

Proof. Firstly X and $\partial_t X$ are bounded on $[0,T] \times \mathcal{A}$,

$$\|X(t,A)\| \leq \mathcal{K}, \quad \|\partial_t X(t,A)\| \leq \|F\|_\infty,$$

then, for any $A, B \in \mathcal{A}$ and $s,t \in [0,T]$,

$$\|X(t,A) - X(s,A)\| \leq \int_s^t \|F(X(\tau,A),A)\| d\tau \leq \|F\|_\infty (t-s)$$

and so X is Lipschitz in t. Moreover, with lemma 6

$$\|\partial_t X(t,A) - \partial_t X(s,A)\| = \|F(X(t,A),A) - F(X(s,A),A)\| \leq \mathcal{L}\|F\|_\infty (t-s),$$

and $\partial_t X$ is also Lipschitz in t. We now prove that X is also Lipschitz in A,

$$\begin{aligned}\|X(t,A) - X(t,B)\| &= \| \int_0^t F(X(s,A),A) - F(X(s,B),B)\,\mathrm{d}s\| \\ &\leq \mathcal{L}\int_0^t \|X(s,A) - X(s,B)\| + \|A - B\|\,\mathrm{d}s,\end{aligned}$$

and therefore, Gronwall's lemma yields for ant $t \in [0,T]$,

$$\|X(t,A) - X(t,B)\| \leq \mathcal{L}Te^{\mathcal{L}T}\|A-B\|.$$

This implies that $\partial_t X$ is Lipschitz in t,

$$\begin{aligned}\|\partial_t X(t,A) - \partial_t X(t,B)\| &= \|F(X(t,A),A) - F(X(t,B),B)\| \\ &\leq \mathcal{L}\Big[\|X(t,A) - X(t,B)\| + \|A-B\|\Big],\end{aligned}$$

which ends the proof. □

Higher regularity is obtained for $X:\ A \mapsto X(\cdot,A)$ with values in $C(\overline{\mathcal{A}})$ which is differentiable in $[0,T]$ as now stated.

Lemma 8. The function X satisfies, $X \in C^1([0,T], C^0(\overline{\mathcal{A}})^6) \cap \mathrm{Lip}([0,T], W^{1,\infty}(\mathcal{A})^6)$.

Proof. By previous lemma, for any $p \in [1,+\infty)$, $X, \partial_t X \in L^\infty(0,T,W^{1,p}(\mathcal{A})^6)$.
Note that since $L^\infty(\mathcal{A})$ is not separable, the measurability of X (resp. $\partial_t X$) as a mapping from $(0,T)$ to $W^{1,\infty}(\mathcal{A})^6$ may fail.
But $X \in W^{1,\infty}(0,T,W^{1,p}(\mathcal{A})^6)$ for any finite p, and, if $0 \leq s < t \leq T$, one has ([12]) that

$$\begin{aligned}\|X(t,\cdot) - X(s,\cdot)\|_{W^{1,p}(\mathcal{A})} \leq \int_s^t \|\partial_t X(\tau,\cdot)\|_{W^{1,p}(\mathcal{A})} d\tau &\leq (t-s)^{\frac{p-1}{p}} \|\partial_t X\|_{L^p(0,T,W^{1,p}(\mathcal{A}))} \\ &\leq (t-s)^{\frac{p-1}{p}} \|\partial_t X\|_{W^{1,p}(\mathcal{Q})} \leq (t-s)^{\frac{p-1}{p}} \|\partial_t X\|_{W^{1,\infty}(\mathcal{Q})} [19|\mathbb{A}|]^{\frac{1}{p}}\end{aligned}$$

Since $X(t,\cdot), X(s,\cdot) \in W^{1,\infty}(\mathcal{A})^6$, $\lim_{p\to+\infty} \|X(t,\cdot) - X(s,\cdot)\|_{W^{1,p}(\mathcal{A})} = \|X(t,\cdot) - X(s,\cdot)\|_{W^{1,\infty}(\mathcal{A})}$, and passing to the limit $p \to +\infty$ yields

$$\|X(t,\cdot) - X(s,\cdot)\|_{W^{1,\infty}(\mathcal{A})} \leq \|\partial_t X\|_{W^{1,\infty}(\mathcal{Q})}(t-s).$$

Thus, X is a bounded Lipschitz-continuous function from $[0,T]$ to $W^{1,\infty}(\mathcal{A})^6$.
By selecting p such that $W^{1,p}(\mathcal{A}) \hookrightarrow C^0(\overline{\mathcal{A}})$, one gets for any t, h such that $t, t+h \in [0,T]$,

$$\begin{aligned}&\|\frac{X(t+h,.) - X(t,.)}{h} - \partial_t X(t,.)\|_{C(\overline{\mathcal{A}})} = \frac{1}{|h|}\| \int_t^{t+h} \partial_t X(\tau,.) - \partial_t X(t,.)\, d\tau\|_{C(\overline{\mathcal{A}})} \\ \leq &\frac{1}{|h|} \int_{(t,t+h)} \|\partial_t X(\tau,.) - \partial_t X(t,.)\|_{C(\overline{\mathcal{A}})}\, d\tau \leq \mathcal{L}\|F\|_\infty h.\end{aligned}$$

So, X as a function with values in $C(\overline{\mathcal{A}})$ is differentiable in $[0,T]$, with derivative $\partial_t X$ and the result in lemma 8 holds. □

We are now interested in the differentiation of the function $X(t,A)$ with respect to A. Since F is not a continuously differentiable function, one cannot apply the implicit function theorem and our arguments will be based on the chain rule in Sobolev spaces. In oder to proceed, F first need to be extended to a Lipschitz function on $\mathbb{R}^{6+15}$. This can be obtained by Kirszbraun theorem, but for technical reasons, it will be done by replacing each coordinate x of A in F by $g(x) = \max(\frac{1}{\mathcal{M}}, \min(\mathcal{M}, x))$ and each coordinate x of X by $h(x) = \min(\mathcal{K}, x^+)$ (or a regularization of these functions). Doing so, each F_k $(k = 1, ..., 6)$ is a globally Lipschitz-continuous, piecewise C^1 function in the sense of Murat *et al* [52]: "*there exists a finite Borel-partition* $(P^i)_{i\in I_k}$ *of* $\mathbb{R}^{6+15}$ *and the same number of globally Lipschitz-continuous* C^1 *functions* $\widetilde{F}_k^i$ *on* $\mathbb{R}^{6+15}$ *such that* $\widetilde{F}_k^i = F_k$ *in* P^i."

Thus, [52, Thm. 2.1], since $X \in W^{1,p}(\mathcal{Q})^6$ for any finite p, one gets that $F_k(X,\cdot) \in W^{1,p}(\mathcal{Q})$ with the chain rule: a.e. in $\mathcal{Q}$,

$$\nabla_{(t,A)} F_k(X,\cdot) = \sum_{i\in I_k} 1_{P^i}(X)(DF_k)(X)\nabla_{(t,A)}X.$$

Moreover, $Y \in W^{1,p}(\mathcal{Q})^6 \mapsto F_k(Y,\cdot) \in W^{1,p}(\mathcal{Q})$ is sequentially continuous for the weak convergences and continuous for the strong topologies.
As the equality $\partial_t X(t,A) = F\big(X(t,A),A\big)$ holds in $W^{1,p}(\mathcal{Q})$, one gets that

$$\begin{aligned}\partial_t \nabla_A X(t,A) &= \nabla_A \partial_t X(t,A) = \nabla_A F\big(X(t,A),A\big)\\ &= \Big(\sum_{i\in I_k} 1_{P^i}\big(X(t,A)\big)(DF_k)\big(X(t,A),A\big)\nabla_A X(t,A)\Big)_k\end{aligned}$$

a priori in $L^p(\mathcal{Q})^{6\times 15}$ since ∇ is a continuous linear map from $W^{1,p}(\mathcal{Q})^{6\times 15}$ onto $L^p(\mathcal{Q})^{6\times 15}$, then in $W^{1,p}(\mathcal{Q})^{6\times 15}$ thanks to the above regularity given by the chain-rule.
For any $Z \in (a, N_{\text{in}}, C_{\text{in}}, V^N_{\max}, K_N, V^C_{\max}, K_C, Q^C_{\min}, Q^N_{\min}, \alpha, \mu_D, \mu_M, \Theta_D, \Theta_M, m_D)$,

$$\partial_Z F_k\big(X(t,A),A\big) = \mathbb{G}_{k,Z}\big(X(t,A),A\big) + \mathbb{H}_k\big(X(t,A),A\big)\partial_Z X(t,A)$$

where $\mathbb{G}_k = (\mathbb{G}_{k,a}, \cdots, \mathbb{G}_{k,\Theta_M})^T$ denotes $D_A F_k$ and $\mathbb{H}_k$ denotes $D_X F_k$ are bounded Borel functions.
Thus, back to the evolution problem,

$$\begin{aligned}\partial_t \nabla_A X(t,A) &= \Big(\mathbb{G}_k\big(X(t,A),A\big) + \mathbb{H}_k\big(X(t,A),A\big)\nabla_A X(t,A)\Big)_k\\ &= D_A F\big(X(t,A),A\big) + D_X F\big(X(t,A),A\big)\nabla_A X(t,A) \qquad (12)\end{aligned}$$

which ODE is solvable ([22, Caratheodory Conditions for Time-Varying Vector Fields])
Finally remark that, since we restricted our study to the case of a constant initial condition X_0, the initial condition for the above ODE is $\nabla_A X(0,A) = 0$ leading to the following lemma.

Lemma 9. $\nabla_A X(\cdot,A)$ is the unique absolute-continuous function in $[0,T]$, solution to the linear ODE (12) associated with the initial condition $(\nabla_A X)_0 = 0$.

5 Numerical simulations: an inverse problem

We wish in this section to investigate the ability of determining system (S) parameters (supposed to be constant) with data obtained from chemostat experiments. Up to now such data are not available: a target model will be considered instead and numerical experiments will be used to measure the capacity of recovering the target model parameters. For this, virtual chemostat data are constructed by numerically solving system (S) and recording its variable states with a given sampling rate.
Attention has been paid, as much as possible, to build a realistic and biologically relevant target problem where TEP production is encountered. So far, it is not possible to find a set of parameter for system (S) in the literature for a given species of diatoms: some parameters are known (usually within a given range) for a specific species, other parameters can be evaluated indirectly. This is detailed in section 5.1, variable units and parameters units and values are reported in tables 1 and 2 respectively.
The parameter identification algorithm is presented in section 5.2 together with its numerical evaluation.

Table 1: Variable units for system (S) (Xeq. is for *Xanthum gum equivalent*, a standard to express TEP concentration).

Variable	description	unit
C	concentration in carbon nutrient	μ mol/L
N	concentration in nitrogen nutrient	μ mol/L
Q_C	cellular quota in carbon	$10^{-9}\mu$ mol/cell
Q_N	cellular quota in nitrogen	$10^{-9}\mu$ mol/cell
D	diatom concentration	10^9 cell/L
M	TEP concentration	g Xeq./L

5.1 The target problem

In this section is presented a *target problem* for parameter identification with physical variable units and with parameters in agreement with biological measurements.
Problem (S) is considered with the variable units given in table 1 and with the parameters given in table 2.

5.1.1 Parameter definition

Parameter estimation for model (S) is not obvious: several parameters have been found in the literature (details and precise references follow) but parameter values differ from a species to another and sometimes depend on the conditions and on the authors. A global set of parameters for a given species of diatom is not available and we will rather focus on searching order of magnitude: without explicit mention, parameters are set to intermediate values inside range of values. Moreover some specific parameters could not be found in the literature and self estimations then will be considered.

Table 2: Parameters for system (S)

Variable	description	value	unit
a	chemostat dilution rate	0.59	day^{-1}
C_{in}	input C concentration	2000	μ mol/L
N_{in}	input N concentration	15	μ mol/L
$V^C_{\max}$	max. C uptake rate	400	$10^{-9}\,\mu$ mol/cell/day
$V^N_{\max}$	max. N uptake rate	70	$10^{-9}\,\mu$ mol/cell/day
K_C	C uptake half saturation constant	1.5	μ mol/L
K_N	N uptake half saturation constant	1.25	μ mol/L
m_D	diatom mortality rate	0.1	day^{-1}
$Q^C_{\min}$	min C cellular quota	0.5	$10^{-9}\,\mu$ mol/cell
$Q^N_{\min}$	min N cellular quota	10	$10^{-9}\,\mu$ mol/cell
μ_D	max. diatom growth rate	1.24	day^{-1}
μ_M	max. TEP production rate	8.2	10^{-9} g Xeq./cell/day
Θ_D	max. N consumption rate for diatom growth	4.5	$10^{-9}\mu$ mol/cell/day
Θ_M	max. C consumption rate for TEP production	1000	$10^{-9}\mu$ mol/cell/day
α	$C:N$ stoichiometric ratio	16	without unit

Nitrogen uptake. Nitrogen is up-taken by diatoms among three inorganic sources: ammonium, nitrate and nitrite. Only nitrate uptake is considered here, as it usually is the main

source of nitrogen, see *e.g.* [64]. To increase the effects of nitrogen limitation during bloom, a low value of $N_{\text{in}} = 15\ \mu\text{mol/L}$ is set, relatively to the range of nitrate input concentration for chemostat experiments of 14-240 μmol/L given in [43]. In [41, table A46] are reported half saturation constants for nitrate uptake for various species of marine diatoms which range from 0.45 to 1.87 μmol/L. Maximal nitrate uptake rate for various species of diatoms can be found in [41, table A60] ranging from 72 to 384 $10^{-9}\,\mu$mol/cell/day and a low value of $V^N_{\max} = 70\ 10^{-9}\,\mu$mol/cell/day is set here to enhance the effects of nitrate limitation after a bloom. Minimal cellular quota in nitrate $Q^N_{\min} = 45.4\ 10^{-9}\,\mu$mol/cell is used in [43], however this value is too high to allow TEP production and a smaller value of $Q^N_{\min} = 10\ 10^{-9}\,\mu$mol/cell will be arbitrary considered here.

Carbon uptake. Carbon nutrient is available either as dissolved CO_2 or as HCO_3^-: dissolved CO_2 has a smaller and variable concentration, between 5 and 25 μmol/L whereas HCO_3^- has a stable and much higher concentration of 2000 μmol/L, we will then only consider HCO_3^- as source of carbon with an input concentration of C_{in}=2000 μmol/L. In [15], the half saturation constant for HCO_3^- uptake has been evaluated to 1-2 μmol/L and the maximal HCO_3^- uptake rate to 175 - 230 $\mu\text{mol}/(\text{mg Chl}_\text{a})/\text{h}$ for diatom *Thalassiossira Weissflogii* (Chl_a standing for type a chlorophyll). With [41, table A30] the cellular content of Chl_a is of $8.1\ 10^{-11}$ mg for *Thalassiossira* species and we get a range of 340 to 450 $10^{-9}\,\mu$mol/cell/day for $V^C_{\max}$. The minimal cell quota in carbon $Q^C_{\min}$ could not have been found in the literature. We propose an estimation based on [42] where it is reported that diatoms are able to accumulate inorganic carbon only when its external concentration is low (when it is of 200 μmol/L the cell concentration in inorganic carbon is of 400 μmol/L), but meanwhile unable to accumulate inorganic carbon when its external concentration is at its normal state of 2000 μmol/L (then the cell concentration in inorganic carbon remains above 1470 μmol/L). Assuming a cell volume of 750 μm^3, see *e.g.* [21], this suggests that $Q^C_{\min} \geq 0.3\,10^{-9}\,\mu$mol/cell and that $Q^C_{\min} \leq 1.5\,10^{-9}\,\mu$mol/cell.

Other parameters. The chemostat dilution rate is set to a =0.59 day^{-1} as for the chemostat numerical simulations in [43], the C:N Redfield ratio is $\alpha = 16$ and the mortality rate is set to 0.1 day^{-1}, a generic value for phytoplankton species given in [41, table A72]. The maximal growth rate for diatom *Thalassiossira Weissflogii* is set to $\mu_D = 1.24\ \text{day}^{-1}$ as measured in [34], and chemostat experiments of TEP production in [33] show a maximal TEP production rate of $\mu_M = 8.2\ 10^{-9}$ g Xeq./cell/day for a marine diatom in the pacific ocean.
The two remaining parameters Θ_D and Θ_M could not be found in the literature. The maximal nitrate consumption rate for diatom growth Θ_D can be estimated from the diatom maximal growth rate $\mu_D = 1.24\ \text{day}^{-1}$. In [41, table A13] is reported the carbon biomass for diatom *Thalassiossira Weissflogii* which ranges between $250 - 450\ 10^{-9}\,\mu$mol C/cell, thus the maximal C consumption rate for diatom growth is of order $250 - 350 \times 0.24\ 10^{-9}\,\mu$mol C/cell/day and using C:N Redfield ratio for diatom growth $\alpha = 16$ we propose the estimation $\Theta_D = 250 - 450 \times 0.24/16\ 10^{-9}\,\mu$mol N/cell/day = 3.75-6.75 $10^{-9}\,\mu$mol N/cell/day.
Finally no satisfactory estimation of Θ_M could be made and it will be set to $\Theta_M = 1000\ 10^{-9}\mu$mol/cell/day leading to a maximal TEP concentration of order 0.1 g Xeq./L which is consistent with the experiments in [33].

5.1.2 Numerical simulation

The evolution of system (S) for the parameters in table 2 is depicted on figures 2 and 1. System (S) is solved on a period of time of 50 days and the solution is recorded every days. At initial state the chemostat concentrations in nutrients $C(0)$ and $N(0)$ are set to C_{in} and N_{in}, diatom cellular quota $Q_C(0)$ and $Q_N(0)$ are set to their associated minimal quota $Q^C_{\min}$ and $Q^N_{\min}$, the

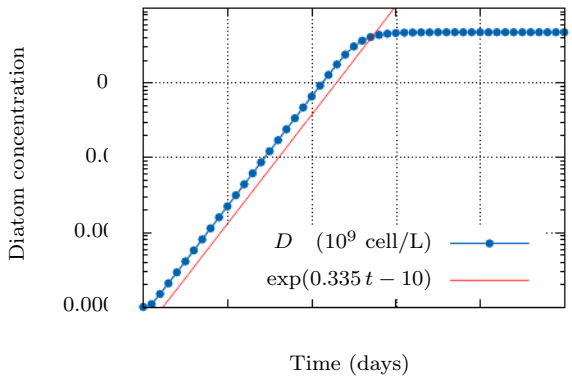

Figure 1: Evolution of diatom population in log scale

initial diatom concentration is $D(0) = 10^5$ cell/L (roughly 1/500 of diatom population observed after the bloom) and no TEP is initially present in the chemostat ($M(0) = 0$).

Figure 1 specifically shows the exponential growth of diatom population up to a final plateau, with a growth rate numerically evaluated to 0.335.

Figure 2 shows the global evolution of system (S) for its six variables C, N, Q_C, Q_N, D and M.
In the 10 first days, nutrient supplies are sufficiently abundant to feed diatoms (whose population is multiplied by 20) and nutrient concentrations remain almost equal to the input concentration, meanwhile cellular quotas quickly increase to reach a first plateau. From day 10 to day 30 diatom population is again multiplied by 20, nutrient supplies are still abundant enough to feed the diatom though a slight decrease in nutrient concentrations is observed, cellular quotas stay at their first plateau level. Between day 20 and day 30 is observed the end of the exponential growth of diatom population which now increases by a factor 7 only: nitrate becomes limiting, its concentration falls to roughly 1/30 of the input concentration stopping diatom growth, nitrate cellular quota decreases to reach a second plateau; meanwhile carbon nutrient remains abundant, carbon uptake is no longer balanced by cellular growth and cellular quota in carbon increases again. In the following period, between days 30 and 40, carbon cellular quota reaches a peak and stop increasing: TEP production is initiated when $Q_C/Q_N > \alpha$, carbon uptake is balanced by TEP production and Q_C reaches its second plateau level. In the last period of time, TEP reaches its maximal concentration and the system stays in an equilibrium state.

5.2 Parameter identification

System (S) is rewritten as,

$$\frac{\mathrm{d}X}{\mathrm{d}t} = F(P, X), \quad X(0) = X^0, \tag{13}$$

with variable $X = (X_i)_{1 \le i \le 6} = (C, N, Q_C, Q_N, D, M)$, with the initial condition X^0 introduced in section 5.1 ($X^0 = (C_{\text{in}}, N_{\text{in}}, Q^C_{\text{min}}, Q^N_{\text{min}}, D(0) = 10^5\,\text{cell/L}, M(0) = 0)$) and where $P \in \mathbb{R}^p$ are the model parameters assumed to be constant: the vector P is composed of the 15 model parameters excepted a, C_{in}, N_{in} and α which are considered as problem data, so that $P \in \mathbb{R}^p$ with $p = 11$.
The target solution X^{tg} is the solution of problem (13) with the parameters P^{tg} in table 2. The same final time $T = 50$ j as on figure 2 is considered. A sampling period $\delta t = T/m$ (for an integer m) is introduced and the associated sampling time instants are the $t_j = j\delta t$ for

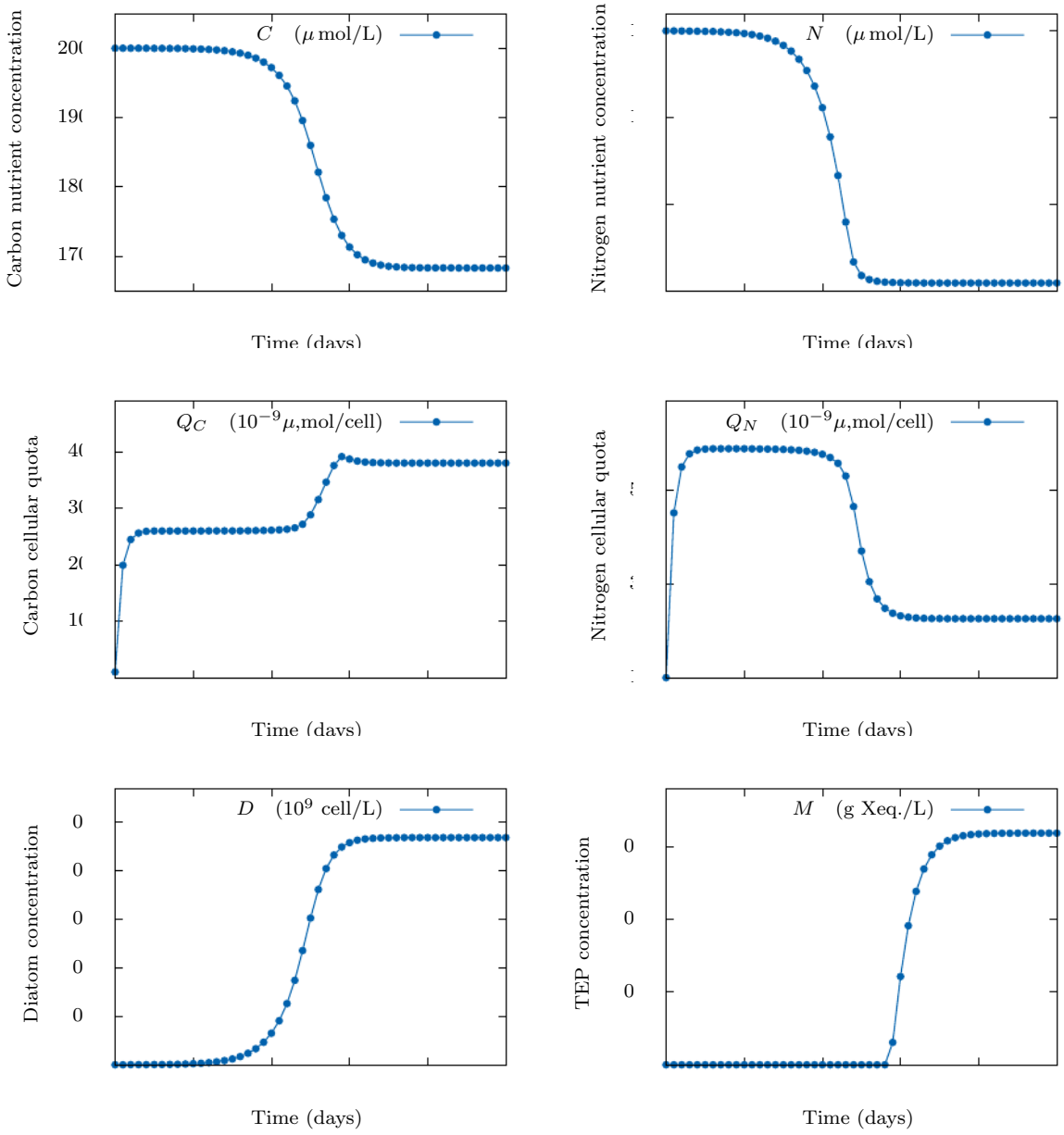

Figure 2: Evolution of system (S) with the parameters in table 2.

$j = 1, \cdots, m$. Our goal will be to minimise the following cost functional $\mathcal{L}$,

$$\mathcal{L}(P) = \sum_{j=1}^{m} \sum_{i=1}^{6} r_{ij}(P)^2, \quad r_{ij}(P) := X_i(t_j) - X_i^{\text{tg}}(t_j) \tag{14}$$

which obviously has value 0 when $P = P^{\text{tg}}$.
The key argument is that, following theorem 1, it is possible to differentiate $X(t)$ with respect to P and thus to compute a gradient for the functional $\mathcal{L}$. More precisely, by differentiating the ODE in 13, $\dfrac{\partial X}{\partial P}$ satisfy the following ODE,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial X}{\partial P}\right) = \frac{\partial F}{\partial P}(X, P) + \frac{\partial F}{\partial t}(X, P)\frac{\mathrm{d}X}{\mathrm{d}X}, \quad \frac{\partial X}{\partial P}(0) = 0, \tag{15}$$

where it has to be noticed that the initial condition $\dfrac{\partial X}{\partial P}(0) = 0$ follows from the fact that the same initial condition $X(0) = X^0$ is imposed for $X(t)$ independently with P.
Then, the derivative $\dfrac{\partial X}{\partial P}(t)$ will be computed numerically by solving the coupled ODE system 13 15. Thus, the evaluation of $\nabla\mathcal{L}(P)$ requires the resolution of the ODE system 13 15 over the time interval $[0, T]$, the storage of $X(t_j)$ and of $\partial X/\partial P(t_j)$ at every sampling time instants t_j to get,

$$\nabla\mathcal{L}(P) = 2\sum_{j=1}^{m}\sum_{i=1}^{6}\left(X_i(t_j) - X_i^{\mathrm{tg}}(t_j)\right)\frac{\partial X_i}{\partial P}(t_j).$$

The computation of $\nabla\mathcal{L}(P)$ will be held using the Runge Kutta 4 algorithm and a time step of 0.01 day ensuring a very accurate computation of $X(t)$ and $\partial X/\partial P(t)$, for which the error has been evaluated to be below 1E-7 for the L^∞ norm based on a numerical convergence analysis.

5.3 Minimisation algorithm

The Gauss Newton method The Gauss Newton method to solve a least square problem is briefly recalled. Consider the functional to be minimized, for $P \in \mathbb{R}^p$,

$$\mathcal{L}(P) = \mathbf{r}^T\mathbf{r}, \quad \mathbf{r} = (r_i(P))_{1\le i\le n}.$$

It is minimized starting from an initial guess P^0 and computing successive approximations $P^{i+1} = P^i + \delta$, where the increment δ is computed at each step as the solution of the linear system $J^TJ\delta = -J^T\mathbf{r}(P^i)$, where $J \in \mathbb{R}^{n\times p}$ is the Jacobian matrix of $\mathbf{r}$ at point P_i. Note that the system matrix $J^TJ \in \mathbb{R}^{p\times p}$ is symmetric positive and moreover definite in practice.
The algorithm is stopped when $\|\mathbf{r}\|_2 \le \varepsilon$, or more precisely in our case when

$$\frac{\mathcal{L}(P)^{1/2}}{|X^{\mathrm{tg}}|_2} \le \mathrm{tol}, \quad |X^{\mathrm{tg}}|_2^2 := \sum_{j=1}^{m}\sum_{i=1}^{6} X_i^{\mathrm{tg}}(t_j)^2 \tag{16}$$

for a given tolerance tol.

The advantage of Gauss Newton algorithm is that it is able to deliver very accurate numerical solutions together with very fast convergence properties. However its main drawback is to induce instabilities, especially when the initial guess is not accurate enough, leading to numerical blow-ups in the resolution of ODE (13). A modified Gauss-Newton algorithm can be considered, as detailed below, to improve the method stability.

Modified Gauss Newton method At every step of the Gauss Newton method, the increment δ is now considered as a descent direction. ODE (13) is numerically solved with parameters $P^i + h_k\delta$ with a decreasing h_k starting at $h_0 = 1$ in order to detect the presence of instabilities and more precisely of a numerical blow-up. Once a safe value of $h_k \in (0, 1]$ has been determined, a golden section search is done to evaluate a best descent step $h \in (0, h_k]$ and the present algorithm step ends by setting $P^{i+1} = P^1 + h\delta$.

Instabilities remain problematic even with the modified version of the Gauss Newton method. Alternatively a classical gradient algorithm can also be used. Such an algorithm is stable but after a few steps (2 or 3 steps in general) it stops moving, delivering numerical solutions of very poor accuracy. It can however be considered in order to improve the initial guess quality.

The minimisation method used in practice None of the previous methods (Gauss Newton, modified Gauss Newton and gradient method) being fully satisfactory, a combination

of these methods will be considered. The first remark is that the Gauss Newton method is very accurate and fast once the initial guess P^0 is close enough to the target solution P^{tg}. Then, the following combination of methods is used to build an accurate initial guess, and the algorithm ends with a Gauss Newton with a small tolerance set to tol $= 10^{-10}$ in (16) and a maximal number of iterations set to 50.
First of all we avoid performing directly the parameter identification on the whole time period $[0, T]$ and instead try to predict P on a shorter time period $[0, T_{1/2}]$ (with $T_{1/2} = 20$ days). Then these predicted parameters are used as an initial guess for a second parameter identification now on the time period $[0, T_{1/2} + 1]$ and so on up to $T_{1/2} + k = T$.
At every stage we start with a few steps (namely 2) of the gradient method. Then the modified Gauss Newton method is applied, with at most 50 iterations and with a rough tolerance of 10^{-4} and eventually the Gauss Newton method is launched with also a maximal number of iterations of 50 and a tolerance of 10^{-4}.

5.4 Numerical results

The algorithm presented in the previous section is evaluated by setting the initial guess components to $P_i^0 = P_i^{\text{tg}}(1 + \theta_i)$ where the θ_i are randomly generated numbers of given amplitude ε, $\theta_i \in [-\varepsilon, \varepsilon]$. Parameter identification is launched with this initial guess and we record:

1. whether it succeeded or failed to find a solution (failure being associated with a blow-up in the resolution of ODE (13),
2. the final residual defined in (16),
3. the maximal relative error $\|P^{\text{num}} - P^{\text{tg}}\|_\infty / \|P^{\text{tg}}\|_\infty$ between the numerically determined parameters P^{num} and the target parameters.

For a given amplitude ε, a number of 200 such experiments are done and the results are statistically presented in tables 3, 4 and 5 for three different sampling period of 1 day, half a day and a quarter of a day respectively.

Amplitude ε	1 %	5 %	10 %	15 %	20 %	5 %	50 %
Success rate (in %)	87	88	72.5	55.5	42	32.5	19.5
Mean residual	1E-7	2E-7	9E-7	3E-7	0.4	0.2	0.5
Mean max. error (in %)	5E-5	3E-4	2E-2	1.0	12.	29.	26.

Table 3: Numerical results for a sampling period of 1 day

Amplitude ε	1 %	5 %	10 %	15 %	20 %	5 %	50 %
Success rate (in %)	76.5	85	67.5	57.5	49.5	37.5	11.5
Mean residual	6E-5	5E-6	6E-6	5E-6	4E-5	0.4	0.2
Mean max. error (in %)	1E-3	1E-4	1E-3	2E-3	7E-2	13.	19.

Table 4: Numerical results for a sampling period of 0.5 day

A first fact is that, even for a small perturbation amplitude of 1 %, there is no full guaranty of converging to a numerical solution because of instability's. The algorithm success rate is quite large for small perturbation amplitudes but decreases and become quite small for a larger amplitudes of 25 %. The quality of the numerical solution can be easily controlled by the

Amplitude ε	1 %	5 %	10 %	15 %	20 %	5 %	50 %
Success rate (in %)	84	90	76.5	65	49	37	14.5
Mean residual	1E-6	1E-6	2E-6	2E-6	1E-6	0.6	0.3
Mean max. error (in %)	3E-4	3E-4	5E-4	3E-2	1E-2	14.	26.

Table 5: Numerical results for a sampling period of 0.25 day

final residual: once the numerical solution is associated to a very small residual (below 1E-5, which usually occurs when the algorithm converges) then the maximal relative error between the numerical solution and the target parameters is very small, most generally below 1E-2 %. We point out that this figure is a max error and that in practice most of the parameters are determined with an accuracy below 1E-10 % whereas few of them are less accurately predicted. When the residual is higher, of order 1 or more, the parameters are less accurately predicted but still with a correct order of magnitude. Again, we point out that only a few of the parameters are not accurately predicted when most of them, with $\varepsilon = 50\%$ *e.g.*, are captured with a relative error below 1 %. This seems to indicate the presence of many local minima for the functional $\mathcal{L}$ around P^{tg}.
As expected, results are better when decreasing the sampling period, however, the benefits are not very important and a sampling period of 1 day seems to be relevant for this present application.

As a conclusion, parameter identification of model (S) seems to be feasible with chemostat experiment data. The presented algorithm allows very accurate parameter identification but requires a first initial guess of high quality because of remaining strong unsuitability features. It seems necessary to go to a coupled approach, first using a generic minimisation method for the functional $\mathcal{L}$ in (14) (such as a genetic algorithm *e.g.*) able to predict a good initial guess, and then to end the parameter identification with a Gauss Newton type method. This is let to further developments on the subject.

6 Acknowledgments

The authors would like to thank Pr. Xavier Mari from IRD (Institute of Research for Development) for valuable discussions and references he provided.

The authors would like to thank the National Label of Excellence I-SITE: Energy and Environment Solutions (E2S - UPPA) for the financial support given to this work.

References

[1] Y. Aktan, A. Dede, and P. S. Ciftci. Mucilage event associated with diatoms and dinoflagellates in Sea of Marmara, Turkey. *Harmful Algae News*, IOC-UNESCO, 1-3, 2008.

[2] T. Alcoverro, E. Conte and L. Mazzella. Production of mucilage by the Adriatic epipelic diatom cylindrotheca closterium (bacillariophyceae) under nutrient limitation. *Journal of Phycology*, 36(6): 1087–1095, 2000.

[3] A. E. Allen, C. L. Dupont, M. Oborník, A. Horák, A. Nunes-Nesi, J. P. McCrow, H. Zheng, D. A. Johnson, H. Hu, A. R. Fernie and C. Bowler. Evolution and metabolic significance of the urea cycle in photosynthetic diatoms. *Nature*, 473(7346): 203–207, 2011.

[4] I. Auby and N. Neaud-Masson. Identification des composants d'une substance dénommée localement "liga" se déposant sur certains engins de pêche au large de Saint Jean de Luz.

Report of Institut français de recherche pour l'exploitation de la mer (ifremer), 2001. https://archimer.ifremer.fr/doc/00076/18681/

[5] M. E. Baird and S. M. Emsley. Towards a mechanistic model of plankton population dynamics. *Journal of Plankton Research*, 21(1): 85–126, 1999.

[6] N. Balkis-ozdelice, T. Durmuş M. Balci. A Preliminary Study on the Intense Pelagic and Benthic Mucilage Phenomenon Observed in the Sea of Marmara. *International Journal of Environment and Geoinformatics*, 8(4): 414–422, 2021.

[7] J. Barcelos e Ramos, K. G. Schulz, C. Brownlee, S. Sett and E. B. Azevedo. Effects of Increasing Seawater Carbon Dioxide Concentrations on Chain Formation of the Diatom *Asterionellopsis glacialis*. *PLoS ONE* 9(3): e90749, 2014.

[8] A. Barral, B. Gomez, F. Fourel, V. Daviero-Gomez and C. Lécuyer. CO_2 and temperature decoupling at the million-year scale during the cretaceous greenhouse. *Scientific Reports*, 7: 8310, 2017.

[9] A. Bartual, I. V. Cera, S. Flecha and L. Prieto. Effect of dissolved polyunsaturated aldehydes on the size distribution of transparent exopolymeric particles in an experimental diatom bloom *Mar Biol.*, 164: 120, 2017.

[10] D. J. Beerling and D. L. Royer. Convergent cenozoic co2 history. *Nature Geoscience*, 4(7): 418–420, 2011.

[11] T. Brembu, A. Muhlroth, L. Alipanah L and AM. Bones. The effects of phosphorus limitation on carbon metabolism in diatoms. *Phil. Trans. R. Soc. B*, 372: 20160406, 2017.

[12] H. Brézis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert.* North-Holland Mathematics Studies. 5. Notas de matematica (50). Amsterdam-London: North-Holland Publishing Comp.; New York: American Elsevier Publishing Comp., Inc. 183 p., 1973.

[13] M.A. Brzezinski. The Si: C: N ratio of marine diatoms: interspecific variability and the effect of some environmental variables *Journal of Phycology*, 21(3): 1985, 347-357.

[14] A. B. Burd, J. P. Chanton, K. L. Daly, S. Gilbert, U. Passow and A. Quigg. The science behind marine-oil snow and MOSSFA: Past, present, and future. *Progress in Oceanography*, 187:102398, 2020.

[15] S. Burkhardt, G. Amoroso, U. Riebesell, and D. Sültemeyer. CO_2 and $HCO3^-$ uptake in marine diatoms acclimated to different CO_2 concentrations. *Limnology and Oceanography*, 46(6):1378–1391, 2001.

[16] A. Bussard. *Capacités d'acclimatation des diatomées aux contraintes environnementales.* PhD thesis, Muséum national d'histoire naturelle - Sciences de la Nature et de l'Homme , 2015.

[17] E. Buzzelli, R. Gianna, E. Marchiori and M. Bruno. Influence of nutrient factors on production of mucilage by amphora coffeaeformis var. perpusilla. *Continental Shelf Research*, 17(10): 1171 – 1180, 1997.

[18] M. Cadier. *Diversité des communautés phytoplanctoniques en relation avec les facteurs environnementaux en mer d'Iroise : approche par la modélisation 3D.* PhD thesis, LEMAR - Laboratoire des Sciences de l'Environnement Marin, tel.archives-ouvertes.fr/tel-01383247, 2016.

[19] S. Calvo, R. Barone and L. Flores. Observations on mucus aggregates along sicilian coasts during 1991–1992. *Science of The Total Environment*, 165(1): 23 – 31, 1995.

[20] P. Cermeño. The geological story of marine diatoms and the last generation of fossil fuels *Perspectives in Phycology*, 3(2): 53–60, 2016.

[21] J. Chen and D. Thornton. The effect of temperature and growth rate on TEP production by *Thalassiosira Weissflogii*. *Journal of phycology*, 47:S64–S64, 2011.

[22] J. Cortes Discontinuous dynamical systems *IEEE Control Systems Magazine*, 28(3): 36–73, 2008.

[23] R. Danovaro, S. Fonda Umani and A. Pusceddu. Climate Change and the Potential Spreading of Marine Mucilage and Microbial Pathogens in the Mediterranean Sea. *PLoS ONE*, 4(9): e7006, 2009.

[24] A. Dauta. Conditions de développement du phytoplancton. Étude comparative du comportement de huit espèces en culture. II. Rôle des nutriments : assimilation et stockage intracellulaire *Annls Limnol.*, 18 (3): 263-292, 1982.

[25] D. Degobbis, R. Precali, C. R. Ferrari, T. Djakovac Attilio Rinaldi, I. Ivancic, M. Gismondi and N. Smodlaka. Changes in nutrient concentrations and ratios during mucilage events in the period 1999-2002. *Science of the Total Environment*, 353:103–114, 2005.

[26] P. Del Negro, E. Crevatin, C. Larato, C. Ferrari, C. Totti, M. Pompei, M. Giani, D. Berto and S. Fonda Umani. Mucilage microcosms. *Science of The Total Environment*, 353(1):258 – 269, 2005.

[27] M. Droop Vitamin B12 and marine ecology. IV. The kinetics of uptake, growth and inhibition in Monochrysis lutheri *Journal of the Marine Biological Association of the United Kingdom.* 48(3): 689–733, 1968

[28] S. Dutkiewicz, M. J. Follows and J. G. Bragg. Modeling the coupling of ocean ecology and biogeochemistry. *Global Biogeochem. Cycles*, 23: GB4017, 2009.

[29] J. Elser, R. Sterner, E. Gorokhova, W. Fagan, T. Markow, J. Cotner, J. Harrison, S. Hobbie, G. Odell and L. Weider. Biological stoichiometry from genes to ecosystems. *Ecology Letters*, 3(6):540–550, 2000.

[30] L. C. Evans and R. F. Gariepy. *Measure Theory and Fine Properties of Functions*, Revised Edition. Chapman and Hall/CRC, 2015.

[31] K. J. Flynn and A. Mitra. Why plankton modelers should reconsider using rectangular hyperbolic (michaelis-menten, monod) descriptions of predator-prey interactions. *Frontiers in Marine Science*, 3:165, 2016.

[32] M. J. Follows, S. Dutkiewicz, S. Grant and S. W. Chisholm. Emergent biogeography of microbial communities in a model ocean. *SCIENCE*, 315:1843–1846, 2007 (and Supporting Online Material).

[33] T. Fukao, K. Kimoto, and Y. Kotani. Effect of temperature on cell growth and production of transparent exopolymer particles by the diatom *Coscinodiscus granii* isolated from marine mucilage. *Journal of Applied Phycology*, 24(2):181–186, 2012.

[34] N. García, J. A. López-Elías, A. Miranda, M. Martínez-Porchas, N. Huerta, and A. García. Effect of salinity on growth and chemical composition of the diatom *Thalassiosira weissflogii* at three culture phases. *Latin American Journal of Aquatic Research*, 40(2):435–440, 2012.

[35] J. L. Genzer, M. Kamalanathan, L. Bretherton, J. Hillhouse, C. Xu, P. H. Santschi and A. Quigg. Diatom aggregation when exposed to crude oil and chemical dispersant: Potential impacts of ocean acidification *PLoS ONE 15(7): e0235473*, 2020.

[36] V. Girard, S. Saint Martin, J.-P. Saint Martin, A. R. Schmidt, S. Struwe, V. Perrichot, G. Breton and D. Néraudeau. Exceptional preservation of marine diatoms in upper Albian amber. *Geology*, 37(1):83–86, 2009.

[37] D. M. Harwood, V. A. Nikolaev and D. M. Winter. Cretaceous records of diatom evolution, radiation, and expansion. *The Paleontological Society Papers*, 13:33–59, 2007.

[38] C. Heemann. *Phytoplanktonexsudation in Abhängigkeit von der Meerwasserkarbonatchemie.* Thesis, Univ. Bremen, 2002.

[39] M. Hein and K. Sand-Jensen. Co2 increases oceanic primary production. *Nature*, 388(6642):526–527, 1997.

[40] D. O. Hessen and T. R. Anderson. Excess carbon in aquatic organisms and ecosystems: Physiological, ecological, and evolutionary implications. *Limnology and Oceanography*, 53(4):1685–1696, 2008.

[41] S. Joergensen, M. Friis, and J. Henriksen. Handbook of environmental data and ecological parameters. 1979.

[42] A. M. Johnston and J. A. Raven. Inorganic carbon accumulation by the marine diatom phaeodactylum tricornutum. *European Journal of Phycology*, 31(3):285–290, 1996.

[43] C. A. Klausmeier, E. Litchman, T. Daufresne and S. A. Levin. Optimal nitrogen-to-phosphorus stoichiometry of phytoplankton. *Nature*, 429(6988):171–174, 2004.

[44] C. A. Klausmeier, E. Litchman and S. A. Levin. Phytoplankton growth and stoichiometry under multiple nutrient limitation. *Limnol. Oceanogr.*, 49(4, part 2): 463–1470, 2004.

[45] T. Legovic and A. Cruzado. A model of phytoplankton growth on multiple nutrients based on the Michaelis-Menten-Monod uptake, Droop's growth and Liebig's law. *Ecological Modelling*, 99: 19–31, 1997.

[46] O. Levitan, J. Dinamarca, E. Zelzion, D. S. Lun, L. T. Guerra, M. Kyung Kim, J. Kim, B. A. S. Van Mooy, D. Bhattacharya and P. G. Falkowski. Remodeling of intermediate metabolism in the diatom Phaeodactylum tricornutum under nitrogen stress. *PNAS*, 112 (2): 412–417, 2015.

[47] E. Litchman, C. Klausmeier, O. Schofield and P. Falkowski. The role of functional traits and trade-offs in structuring phytoplankton communities: scaling from cellular to ecosystem level. *Ecol Lett.*, 10(12):1170–1181, 2007.

[48] L. MacKenzie, I. Sims, V. Beuzenberg and P. Gillespie Mass accumulation of mucilage caused by dinoflagellate polysaccharide exudates in Tasman Bay, New Zealand *Harmful Algae*, 1: 69–83, 2002.

[49] X. Mari, U. Passow, C. Migon, A. B. Burd and L. Legendre Transparent exopolymer particles : Effects on carbon cycling in the ocean *Progress in Oceanography*, 151: 13–37, 2017.

[50] A. C. Martiny, J. A. Vrugt and M. W. Lomas. Concentrations and ratios of particulate organic carbon, nitrogen, and phosphorus in the global ocean. *Scientific Data*, 1(1):140048, 2014.

[51] M. Mecozzi, E. Pietrantonio, V. Di Noto and Z. Papai. The humin structure of mucilage aggregates in the Adriatic and Tyrrhenian seas: hypothesis about the reasonable causes of mucilage formation. *Marine Chemistry*, 95:255–269, 2005.

[52] F. Murat and C. Trombetti. A chain rule formula for the composition of a vector-valued function by a piecewise smooth function. *Bollettino dell'Unione Matematica Italiana*, 8, 6-B(3):581–595, 2003.

[53] J. M. Nakov, T. M. Beaulieu and A. J. Alverson. Accelerated diversification is related to life history and locomotion in a hyperdiverse lineage of microbial eukaryotes (diatoms, bacillariophyta). *New Phytologist*, 219:462–473, 2018.

[54] U. Passow. Formation of transparent exopolymer particles, TEP, from dissolved precursor material *Mar Ecol Prog Ser*, 192:1–11, 2000.

[55] A. Peace, H. Wang and Y. Kuang. Dynamics of a Producer–Grazer Model Incorporating the Effects of Excess Food Nutrient Content on Grazer's Growth *Bull Math Biol*, 76:2175–2197, 2014.

[56] J. Renaudie. Quantifying the cenozoic marine diatom deposition history: links to the C and Si cycles. *Biogeosciences*, 13(21):6003–6014, 2016.

[57] C. S. Reynolds. Variability in the provision and function of mucilage in phytoplankton: facultative responses to the environment. *Hydrobiologia*, 578(1):37–45, 2007.

[58] U. Riebesell, K. G. Schulz, R. G. J. Bellerby, M. Botros, P. Fritsche, M. Meyerhöfer, C. Neill, G. Nondal, A. Oschlies, J. Wohlers and E. Zöllner. Enhanced biological carbon consumption in a high co2 ocean. *Nature*, 450(7169):545–548, 2007.

[59] S. A. Sañudo-Wilhelmy, A. Tovar-Sanchez, F. X. Fu, D. G. Capone, E. J. Carpenter and D. A. Hutchins. The impact of surface-adsorbed phosphorus on phytoplankton Redfield stoichiometry. *Nature*, 432(7019):897–901, 2004.

[60] O. Sayanova, V. Mimouni, L. Ulmann, A. Morant-Manceau, V. Pasquet, B. Schoefs and J. A. Napier Modulation of lipid biosynthesis by stress in diatoms. *Phil. Trans. R. Soc. B*, 372:20160407, 2017.

[61] S. Scala and C. Bowler. Molecular insights into the novel aspects of diatom biology. *CMLS, Cell. Mol. Life Sci.*, 58: 1666–1673, 2001.

[62] M. Schapira. *Dynamique spatio-temporelle de Phaeocystis globosa en Manche orientale : effets de la turbulence et des apports sporadiques en sels nutritifs.* PhD thesis, UMR 8013 ELICO: Ecosystèmes Littoraux et Côtiers, Université de Lille 1, 2005.

[63] J. Seckbach and J. P. Kociolek *The Diatom World* Cellular origin, life in extreme habitats and astrobiology, Volume 19, 2011.

[64] J. L. Serra, M. J. Llama, and E. Cadenas. Nitrate utilization by the diatom *Skeletonema costatum*: Ii. regulation of nitrate uptake. *Plant Physiology*, 62(6):991–994, 1978.

[65] V. Smetacek. Diatoms and the ocean carbon cycle. *Protist*, 150(1):25 – 32, 1999.

[66] S. R. Smith, C. L. Dupont, J. K. McCarthy, J. T. Broddrick, M. Oborník, A. Horák, Z. Füssy, J. Cihlář, S. Kleessen, H. Zheng, J. P. McCrow, K. K. Hixson, W. L. Araújo, A. Nunes-Nesi, A. Fernie, Z. Nikoloski, B. O. Palsson and A. E. Allen. Evolution and regulation of nitrogen flux through compartmentalized metabolic networks in a marine diatom. *Nature Communications*, 10(1):4552, 2019.

[67] N. Susperregui. "Liga", mucilages marins sur la côte basque *Dossier de presse Comité Interdépartemental des Pêches Maritimes et des Elevages Marins des Pyrénées Atlantiques - Landes*, Juin 2019.

[68] N. Susperregui. Personal communication.

[69] D. C. O. Thornton. Coomassie Stainable Particles (CSP) : Protein Containing Exopolymer Particles in the Ocean *Front. Mar. Sci.*, 5, p. 206, 2018.

[70] A. Torstensson, M. Hedblom, M. Mattsdotter Björk, M. Chierici, and A. Wulff. Long-term acclimation to elevated pCO_2 alters carbon metabolism and reduces growth in the antarctic diatom *Nitzschia lecointei. Proceedings of the Royal Society B: Biological Sciences*, 282(1815): 20151513, 2015.

[71] D.H. Turpin Physiological mechanisms in phytoplankton resource competition *Growth and reproductive strategies of freshwater phytoplankton, Cambridge University Press, 1988*

[72] Z. Wei, T. Xuexi, Y. Yingying, Z. Xin Zhang and Z. Xinxin. Elevated pCO_2 Level Affects the Extracellular Polymer Metabolism of Phaeodactylum tricornutum. *Front. Microbiol.*, 11: Article 339, 2020.

[73] P. Zahajská, S. Opfergelt, S. C. Fritz, J. Stadmark and D. J. Conley. What is diatomite? *Quaternary Research*, 96:48–52, 2020.