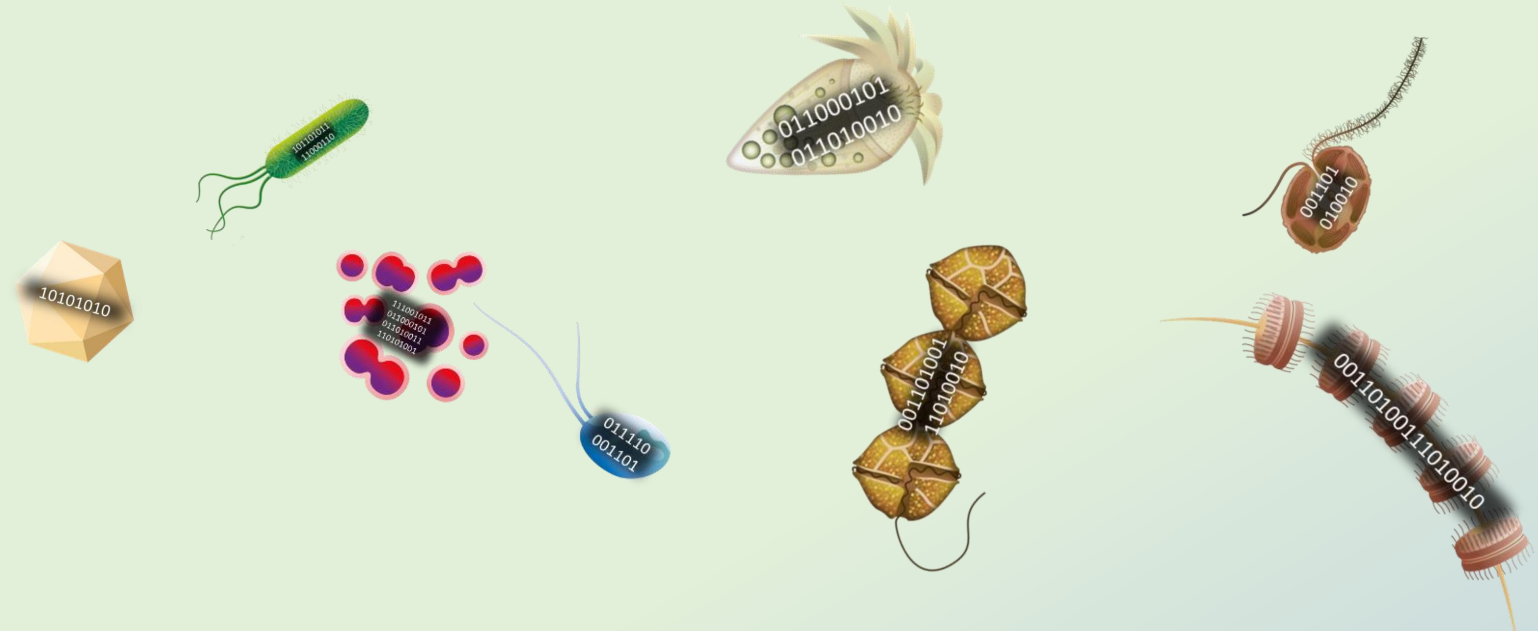

Simulating Plankton - getting it right in the era of Digital Twins of The Ocean

Core features of plankton digital twins

Flynn, Atkinson, Beardall, Berges, Boersma
Brunet, Calbet, Caron, Dam, Glibert, Hansen,
Jin, Lomas, Lønborg, Mayor, Meyer, Millette,
Mock, Mulholland, Poulton, Robinson, Rokitta,
Rost, Saiz, Scanlan, Schmidt, Sherr, Stoecker,
Svensen, Thiele, Thingstad, Unrein & Våge

v2

Simulating Plankton - getting it right in the era of Digital Twins of The Ocean

Core features of plankton digital twins

Flynn KJ[1*], Atkinson A[1], Beardall J[2], Berges JA[3], Boersma M[4], Brunet C[5], Calbet A[6], Caron D[7], Dam H[8], Glibert PM[9], Hansen PJ[10], Jin P[11], Lomas MW[12], Lønborg C[13], Mayor D[14], Meyer B[15], Millette NC[16], Mock T[17], Mulholland M[18], Poulton AJ[19], Robinson C[17], Rokitta S[15], Rost B[15], Saiz E[6], Scanlan D[20], Schmidt K[21], Sherr E[22], Stoecker DK[9], Svensen C[23], Thiele S[24], Thingstad TF[25], Unrein F[26], Våge S[25]

1) Plymouth Marine Laboratory, Prospect Place, Plymouth PL1 3DH, UK
2) Biological Sciences, Monash University, Victoria 3800, Australia
3) University of Wisconsin-Milwaukee, Milwaukee, WI 53211, USA
4) Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung Biologische Anstalt Helgoland, 27498 Helgoland, Germany
5) Stazione Zoologica Anton Dohrn, Naples, Italy
6) Institute of Marine Sciences, CSIC. Pg. Marítim 37-49, 08003. BCN, Spain
7) University of Southern California, Los Angeles, California, USA
8) Marine Sciences, University of Connecticut, CT 06340-6048, USA
9) Horn Point Laboratory, University of Maryland Center for Environmental Science, Cambridge, Maryland, USA
10) University of Copenhagen, Helsingør, Denmark
11) Guangzhou University, China
12) Bigelow Laboratory for Ocean Sciences, ME, USA
13) Department of Ecoscience, Aarhus University, Roskilde, Denmark
14) Exeter University, Exeter, UK
15) Alfred-Wegener-Institute – Helmholtz Centre for Polar and Marine Research, 27570 Bremerhaven, Germany
16) Virginia Institute of Marine Science, Gloucester Point, VA 23062, USA
17) University of East Anglia, Norwich Research Park, Norwich NR4 7TJ UK
18) Old Dominion University, Norfolk 23529 USA
19) Lyell Centre for Earth and Marine Science, Heriot-Watt University, Edinburgh, UK.
20) University of Warwick, Coventry, Warwickshire, UK
21) Earth & Environmental Sciences, University of Plymouth, PL4 8AA, UK
22) Eugene, Oregon, USA
23) Arctic and Marine Biology, UiT The Arctic University of Norway, Tromsø, Norway
24) Microbial Ecology, University of Vienna, Djerassiplatz 1, A-1030 Vienna, Austria
25) Biological Sciences, University of Bergen, Bergen 5020, Norway
26) INTECH, Buenos Aires 7130, Argentina

* KJF@PML.ac.uk

Citing this work

Flynn, K. J., Atkinson, A., Beardall, J., Berges, J. A., Boersma, M., Brunet, C., Calbet, A., Caron, D., Dam, H., Glibert, P. M., Hansen, P. J., Jin, P., Lomas, M. W., Lønborg, C., Mayor, D., Meyer, B., Millette, N., Mock, T., Mulholland, M., Poulton, A., Robinson, C., Rokitta, S., Rost, B., Saiz, E., Scanlan, D., Schmidt, K., Sherr, E., Stoecker, D. K., Svensen, C., Thiele, S., Thingstad, T. F., Unrein, F. & Våge, S. (2024). *Simulating Plankton - getting it right in the era of Digital Twins of The Ocean; core features of plankton digital twins*. Zenodo https://doi.org/10.5281/zenodo.10952868

Please note that this work may be subject to changes during its year of publication. The doi noted in the citation details will always link to the most up-to-date version. Please check for updates.

Acknowledgment to Funder

This work was funded by UKRI Natural Environment Research Council (NE/X010783/1; *Simulating Plankton - getting it right in the era of Digital Twins of The Ocean*), under the direction of the lead author of this work.

Author Contributions

Flynn designed the project and task. Others, notably Sherr, helped to improve the task description. All, except Flynn, undertook the task described in Methods, working individually. Flynn undertook the initial analysis of the Results and drafted the Discussion. All contributed to finalising the report.

Executive Summary

This work describes the outcomes from a subcomponent of a project funded by the NERC (UK) during 2023, with the overarching aim of facilitating the construction of the next generation of plankton simulation models by engaging with experts in real plankton physiology and ecology. Over 30 experts, covering plankton from viruses to krill, contributed to various facets of the project. They were selected specifically for their empirical interests; modellers per se were not included. This component had 32 contributors.

Contributors were asked to configuring aspirational plankton digital twins (PDT). To protect the intellectual property of the contributors, their PDT concepts are not presented. The PDT concepts were analysed with respect to the required abiotic and biotic components and input/output needs.

Most PDT concepts were for short duration (<1mo) laboratory or mesocosm scenarios, some with chemostat-like settings. Field scenarios were 1D (with depth). Most concepts required multiple plankton function types, with host-virus, prey-predator and competition (allelopath) interactions. There was also interest in single organism concepts for photophysiology studies. Concepts included scope for multiple stressor/resource descriptions, most commonly including temperature, light (irradiance, light:dark cycle, with some interests in spectra), and pH/pCO_2. While dissolved inorganics were the most requested nutrient form, over a third expressed interest in detailed dissolved organics (i.e., multiple DOM forms and DOC). Two-way interactions were of interest not only for inorganics (with nutrient regeneration), but also for pH, CO_2/O_2, dissolved organics and the fate of debris (including mortality).

Details of the plankton themselves were invariably expected to include biomass with variable stoichiometry (some including fatty acids), and also numeric abundance (often with explicit allometry). Interests in photopigments were dominated by chlorophyll, with some interest in other pigments such as carotenoids. About a third expressed a requirement for inclusion of allelopaths or toxins.

Almost all expressed a requirement for detailed physiological and/or behavioural descriptions for the plankton, including variable prey or host selectivity. Explicit descriptions of ontogeny and of vertical migration or other depth-linked aspects were also of interest. A few required explicit inclusion of omics-linkages within the PDT.

Input and outputs expectations included an interest in a rewind function during the simulation to facilitate ‘what-if?’ testing, and multiple time and x,y scatter plot options.

In short, the general expectation for a PDT, or at least for first-generation PDTs, describes an in silico laboratory in which a range of plankton types would be described to high levels of detail with respect to physiology and behaviour, in a chemically detailed but dimensionally simple (0D or 1D) setting for short-period incubations.

Contents

1. Introduction

Digital twins are computer-based, virtual or '*in silico*', representatives of real systems. In many ways they represent the state-of-the-art in simulation models coupled with a comprehensive interface supporting interactivity with the interests of the user. Digital twins of plankton, as descriptions of individual species through to communities, have scope to provide advanced tools for educational, research and ecosystem management (Flynn et al. 2022b). Their construction, while undoubtedly challenging, promises to also be highly beneficial for plankton research, by bringing different data and information streams together with a clear aim of summarising knowledge in a dynamic interactive form.

Building plankton digital twins requires collaborations between experts in empirical plankton science, modellers and computer scientists. Digital twins differ from traditional models in that they have no simple set usage; it is largely up to the user to define the settings of the simulation. Flexibility is thus important, to identify what types of interest may be expressed by the user. The work described here was part of a project exploring next generation plankton models and digital twin developments (Flynn 2024). Here contributors (as experts in empirical plankton science) were invited to describe the salient details required to support construction of a digital twin model for a named plankton group, genus, or species, or (because in reality, each 'single' species actually requires interactions with other species) a named collection of organisms. Working from a broad template description, contributors worked through the abiotic and biotic components they would wish to control or see as outputs. The collated data were used to inform an analysis of expectations from the plankton experts.

This is an important step in building a digital twin platform because such platforms are unlikely to be built as bespoke products targeting only specific scenarios. As a digital twin is a model, the user can do things that cannot be done with real organisms. For example, while it is not possible to feed a zooplankton with dead yet motile prey, to test the performance of just the prey handling feature without feedback processes affecting prey nutrient status, such tests can be undertaken with a model. For microbes using dissolved nutrients, this is akin to growing the axenic organism in a flask or in a chemostat. One could also test behaviour in different steady-state or perturbation situations, checking how stable isotopes flow through the system with a temporal resolution of minutes, or conduct the equivalent of instantaneous population fraction with 100% efficiency.

To build an ecosystem-facing plankton digital twin one could construct a single all-encompassing product, or more likely build it as a collection of autecology descriptions. Either way, it is important to verify that each component (here, each plankton functional type) does what it should. This requires the building and testing of a description for each plankton component. Here, contributors were asked to describe such a conceptual structure for a chosen planktonic scenario, considering what it would contain, how the parts would interact, and also how they would want to interact with the model of this particular plankton in a digital twin setting.

2. Method

Contributors were requested to provide a description of a plankton digital twin for a scenario of interest. They could do this in whatever form they wished using the supplied guidance, which is reproduced here (as **Appendix 1**) as it was delivered to the contributors.

The returned descriptions were categorised with respect to the scenario type, plankton types, input abiotic, and output expectations. There were no restrictions to the form of the responses given and hence typically the numbers across options in the following figures sum to more than the 32 total respondents.

To preserve the intellectual property of those descriptions, the returns themselves are not presented.

3. Results

In total 32 replies were made. The identity of the contributors, and the broad subject of their returns as key words, are given in **Appendix 2**. The returns are summarised graphically in **Figs. 1–7**.

Most interest was expressed for laboratory or mesocosm scenarios, and with short duration simulations (**Fig. 1**). The longest request period was in support of studies of evolution (years) but most were for weeks, or a season at most. Some requested a chemostat simulation facility in addition to batch-type or natural scenarios. Some saw the attraction of a rewind facility to more readily explore 'what-if?' scenarios; it is possible that many others did not fully appreciate the attraction of such a tool.

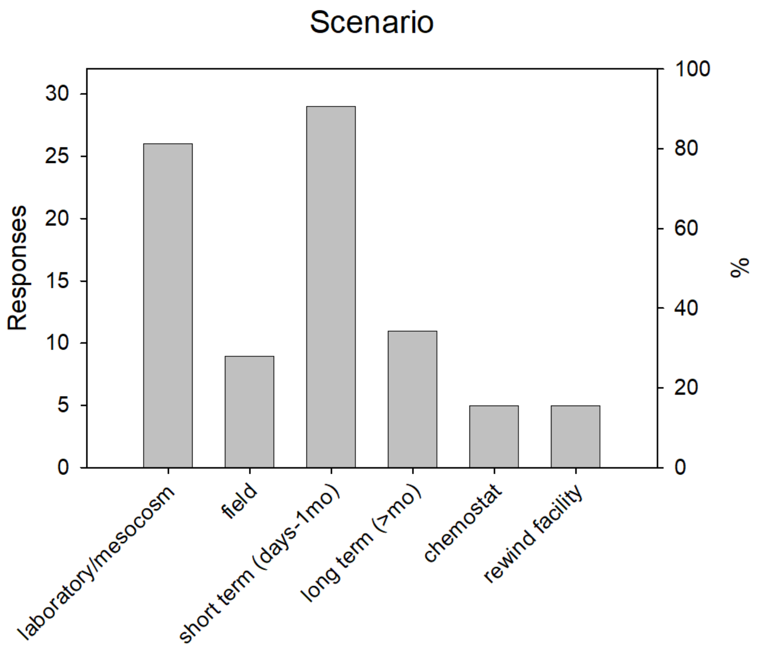

Fig. 1 Target scenario. The rewind facility allows a model to be wound back in time, adjustments made and the run resumed.

With respect to required plankton types, most interest was shown for phytoplankton, but more showed an interest in multi-PFT scenarios involving competitors or host-virus or prey-predator combinations (**Fig. 2**). Those concepts that were not multi-PFT were intended for exploring facets of phototrophy.

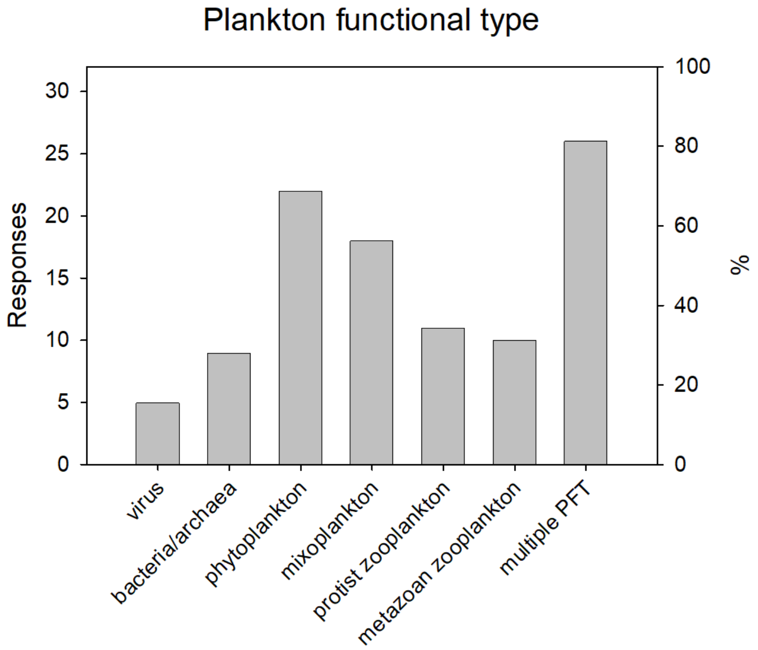

Fig.2 Plankton functional types. 'Phytoplankton' is defined as in Flynn et al. (2019), as photoautotrophs that lack the phagotrophic abilities of mixoplankton; this thus includes cyanobacteria, autotrophic flagellates, and diatoms.

The most important abiotic requirements were for temperature and light (both PFD and light:dark cycle), but interest in pH-related features, including pCO_2 also ranked highly (**Fig. 3**). Interest in depth or pressure descriptions were related to ontogeny or diel vertical migration aspects. Interests in salinity and O_2 were also expressed.

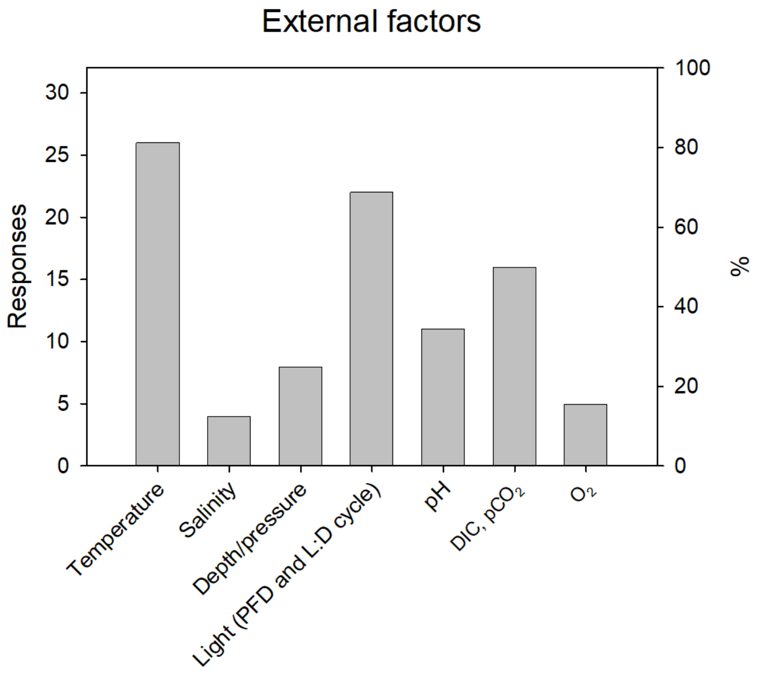

Fig. 3 External factor input requirements.

Nutrient interests were dominated by DIN and DIP, but there were also considerable interests in dissolved organics (DOC or forms of DOM) and also in Fe, with some in Si (**Fig. 4**).

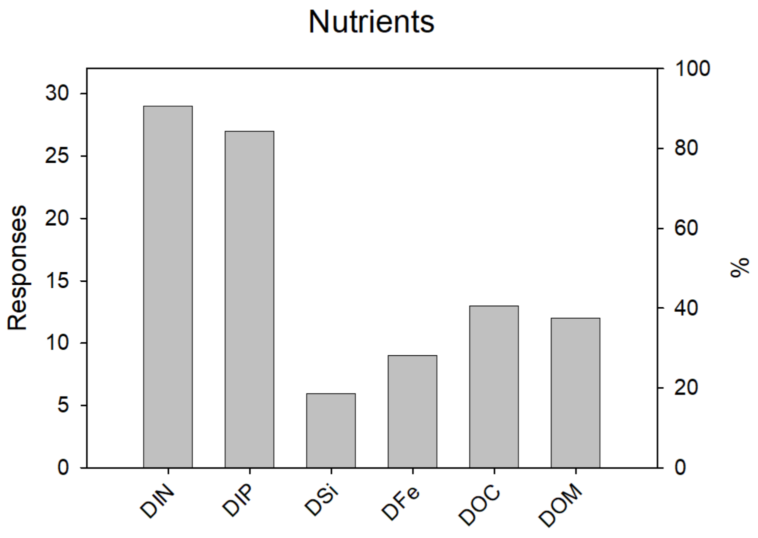

Fig. 4 Nutrient input requirements.

Interest in abiotic factors was dominated by inorganic nutrients and also organic nutrients (**Fig. 5**). Various forms of debris (dead organisms, faeces etc.) were of common interest. Not only was pH (H^+) of interest as an input factor (**Fig. 3**) but it was also of interest (actually more interest) as a variable fluctuating with biological activity (**Fig. 5**). Likewise, changes in O_2 were of interest to a good proportion of the contributors.

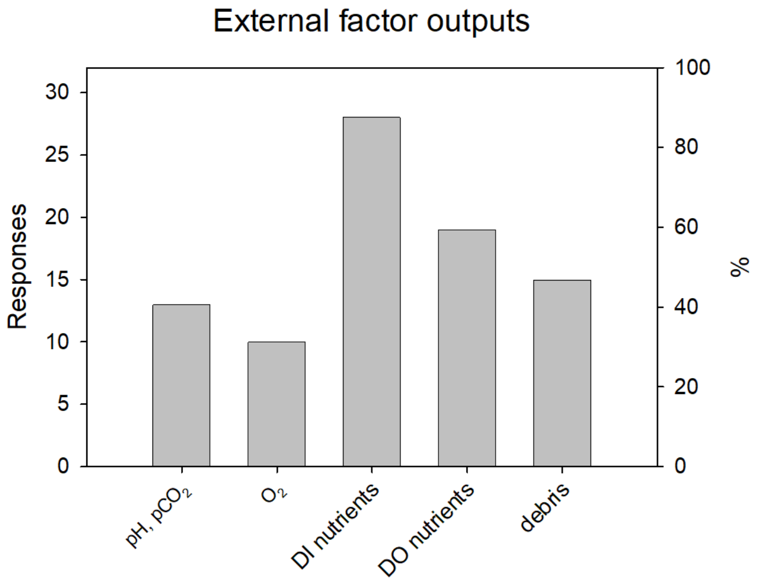

Fig. 5 Factors external to the plankton required as outputs.

All expressed a requirement for biomass as an output, and also for that biomass to be described with reference to variable stoichiometry (**Fig. 6**). Numeric abundance (and thence when considered with biomass, also organism size and allometric factors) was expected as an output. Chlorophyll was widely expected as an output, with some also requiring descriptions of other pigment types. A third were interested in allelopaths and/or toxins. For some, fatty acids were also of interest.

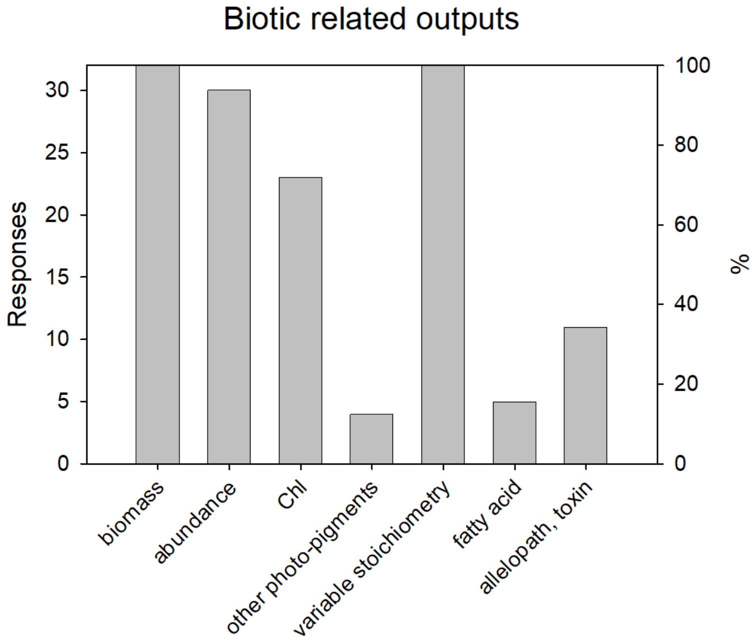

Fig. 6 Required biotic related outputs.

Most contributors required extensive physiological outputs from the model, including respiration and nutrient regeneration rates, as well as gross photosynthesis and growth (**Fig. 7**). Some were interested in detailed aspects of photosynthesis. A third output required by contributors was for variable host or prey selectivity (changing with nutrient status or with the presence of other materials). Ontogeny was of interest not only for zooplankton but also for different stages of protist life cycles. There were lesser interests in interactions with microplastics and water-depth-related behaviours. A few expected explicit information to emerge linking plankton activity to ‘omic signatures.

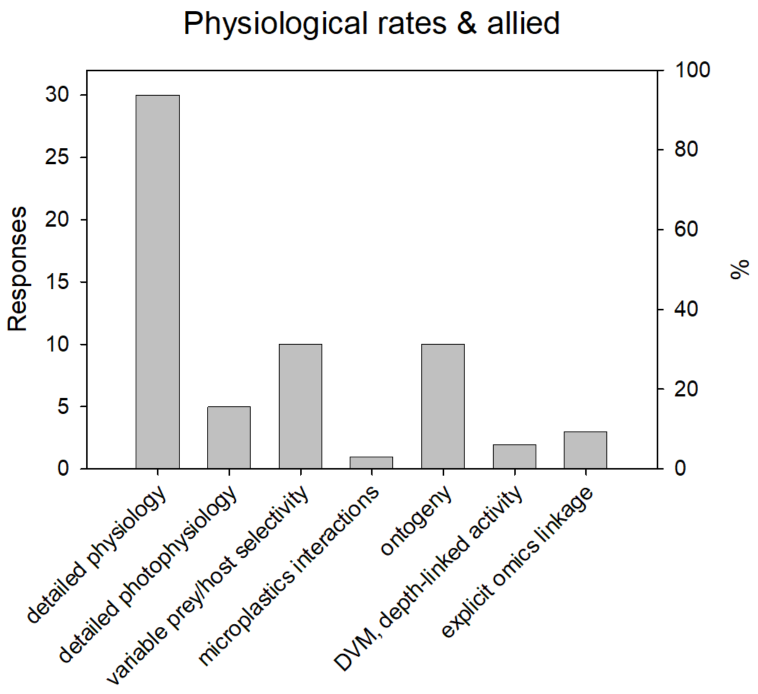

Fig. 7 Physiological rates and allied required outputs.

Output requirements comprised time-based and also x-y scatter graph combinations. No explicit mention was made for other outputs types (such as a dump of all data for subsequent analysis), although that may have been expected by default.

4. Discussion

4.1 Overview

The following discussion builds not only on the analysis shown above but also on other aspects of the responses from contributors to the task described in Methods which are not shown here to preserve intellectual property rights (many of those contributions were provided to the level of detail expected of a grant-application). In discussing the results it is worth emphasising that, like most activities in this project (Flynn 2024), contributors worked alone with no, or little, interaction and thus there was little scope for developing group-think or synergy. Next steps in development would, however, benefit from multi-collaborator engagements.

The most notable feature of the analysis of expectations was an interest in multi-stressor descriptions (**Figs. 3,4**). The complexity of those interactions were considerable, and typically far exceeded those seen in any logistically plausible multi-factorial empirical experiment design. In addition, feedbacks and interactions beyond the usual nutrient recycling were expected, notably with inclusion of changes in H^+ (pH) concurrent with CO_2 fixation and O_2 production/respiration. Temperature also featured commonly, with lesser interest in salinity. Turbulence was notable by its absence as a requested input, especially given the perceived role of turbulence in Margalef's mandala (Margalef 1978) with its effects on encounter rates for virus-host (Murray & Jackson 1992) and predator-prey (Rothschild & Osborn 1988). Interactions of biological items and light climate variations (Brunet et al. 2003, 2008) were also lacking. It is possible that such absences reflect topics being less in vogue at present or, for the mandala, that this is recognised to fail to account for species-level understanding (Glibert 2016). Likewise, the common interest expressed by contributors in pH, temperature etc, likely reflect current interests in those factors as multi-stressors. In reality, all these factors are of importance (more or less so, depending on the scenario), and the constructor of plankton digital twins likely needs to balance such interests in configuring their platforms. There is also a possibility that by building and operating a plankton digital twin with a capacity for exploring multiple stressors beyond the logistic constraints set by empirical experiment design, plankton researchers will discover new important avenues for study.

Of the nutrients, interest in dissolved organics was notable (**Fig. 4**). That interest was not just in a single 'DOM' category, but in multiple forms of DOC and DOM. Similarly, interests in bacteria were invariably directed at multiple types of bacteria, not for inclusion of a single 'bacteria box'. These interests are also reflected in another output from the project (Flynn et al. 2024a), and come not only from those specialising in bacteria but from others with more general and overarching interests in how plankton communities link together.

With the exception of specialised interests in photophysiology, most contributors included multiple levels of plankton functional/trophic types in their descriptions (**Fig. 2**). Within those descriptions, changing physiologies and behaviours were expected to be emergent, relating to the underwater light climate (Brunet et al. 2013; Polimene et al. 2014), nutrient status of prey (Li et al. 2000; Maselli et al. 2022) or of host, and/or the presence of other plankton types. This reveals an interest in behavioural features such as active prey switching and explains in large measure also the dominant expectation that variable stoichiometry would be described by the model, as well as changes in organism size. Although most contributors have specialised interests, they are alive to the needs to consider the states of organisms in trophic positions around their core interests. While in empirical studies, conducting experiments with multiple organisms and multi-stressor situations (e.g., predator-prey), can readily become extremely challenging, if not intractable, such investigations are potentially quick and easy with a digital twin.

4.2 Implementation in Digital Twins

Set against the performance of traditional plankton models, and accepting the recognised utility of simple models directed at specific tasks (Flynn et al. 2024a), the analysis demonstrates the very considerable increase in plankton model capabilities required to meet expectations by empirical plankton researchers for digital twins in comparison with typical (extant) plankton models. It is perhaps convenient then, from a computational aspect, that most interest was expressed in platforms for short-duration scenarios, typically of a few weeks to a month (**Fig. 1**). These scenarios could also largely be met with physical descriptions that are 0D (i.e. totally mixed, like a flask or simple mesocosm) or 1D (i.e., including an explicit depth dimension). This aligns well with the projected development pattern of plankton digital twins (Flynn et al. 2022b).

Below we consider individual facets of some of the required features of Plankton Digital Twins.

4.2.1 Plankton descriptions

Viruses are very rarely included in plankton models (Mateus 2017). Empiricists, however, often see the importance of viruses in trophic interactions, not only in 'killing the winner' (Thingstad 2000; Chen et al. 2021), or at least constraining the potential winner (Flynn et al. 2022a). There are no specific challenges in implementing viruses into models other than those of describing the biodiversity of the host options; the challenge rests more with obtaining the empirical details. The same applies to the role of parasites (Flynn et al. 2024a).

Responses to the task set out here (**Methods**), and also to other components of this project (Flynn 2024, Flynn et al. 2024a, 2024b), flag the importance of separating bacteria into at least two functional groups, which display differential activities (growth and respiration rates, GGE etc.) on various DOM components. Challenges here rest equally between the modelling and empirical sides of the topic. That DOM is so poorly characterised in a systematic objective fashion in empirical studies is a major hindrance in modelling, not least because each sub-component of DOM demands another state variable and allied numeric data for rates of production, consumption and concentration (Anderson et al. 2015), and also for temperature, pH and oxygen. That empiricists are not unhappy with the prospect of models failing (Flynn et al. 2024a), because it is all part of the learning process, provides an impetus for continuing with various DOM descriptions in models rather than awaiting for chemical oceanography and plankton ecophysiology to provide definitive, or at least better defined, data for different forms of DOM.

Of all the plankton types, photoautotrophic phytoplankton are best described in extant models; most others being either ignored (virus, bacteria), miscategorised (mixoplankton), or otherwise poorly described (all zooplankton). Even with phytoplankton, however, there are serious challenges (Flynn et al. 2024a). The requirement to consider not only irradiance, but also the light:dark cycle periodicity of its delivery (Paasche 1968; Durbin 1974) is considered as important by contributors. Newer models can more readily resolve this challenge through not only describing Chl:C but also the changes in the capacity of the enzyme components controlling the maximum rate of photosynthesis, P_{max} (Flynn & Mitra 2023b). Interest in more detailed aspects of photochemistry, the need to describe pigments other than just Chl (Brunet & Lavaud 2010; Polimene et al. 2012) and the consequential changes in the water column light and organism action spectra, need further attention, as would descriptions of rapid photoacclimation in diatoms (Dimier et al. 2007; Dimier et al. 2009; Brunet et al. 2014; Silkin et al. 2021). Models describing short-term changes in photosynthesis have been of interest for the microalgal biotechnology sector for some time (e.g., Costache et al. 2013; Brindley et al. 2016), associated with optimising biomass and biochemical production (Pistelli et al. 2023). For plankton ecology, and actually also for biotechnology, longer-term non-steady-state illumination patterns are of greater importance, as they integrate total cell physiology and life cycle dynamics.

Rather little mention was made by contributors about differential usage of forms of dissolved N and P (Flynn & Butler 1986; Orefice et al. 2019). However, that interest was perhaps shrouded by interests expressed in the roles of osmotrophy (use of DOM) and allelopathy (**Figs. 4,5**). Challenges with DOM are mentioned above, but they are similar to those for allelopathic interactions; again we lack good chemical understanding and quantification (but see Meyer et al. 2022). The challenge for modelling per se is perhaps not as great as it is for empirical science, but both are very significant. The interest expressed in allelopathy will doubtless support greater interactions between empirical and simulation plankton science.

Interest in mixoplankton was significant (**Fig. 2**). While there have been various descriptions of mixoplankton used in models that do not accord with how real organisms work (Mitra et al. 2023), there are also models available that do describe the activity of most types of mixoplankton (e.g., Leles et al. 2021; Flynn & Mitra 2023a). Again, then, it is not the modelling side that is constraining progress but more the empirical science. The availability of a suitable digital twin platform for researchers to experiment with, to help design effective empirical studies, would likely be of benefit here.

Zooplankton have long been known to represent a problem for modelling (Mitra et al. 2014) and there are various problems with the way that most plankton models handle even the basics of their activity (Flynn et al. 2024a). The problem starts in separating protist and metazoan zooplankton types; many models do not even explicitly recognise those groups. In the broadest sense zooplankton encompass all types of strictly heterotrophic plankton that prey on other plankton. The spectrum of predators ranges from heterotrophic nanoflagellates less than 20 µm in size that consume bacteria and small phytoplankton, to microzooplankton sized (~20 to 200 µm) ciliates, dinoflagellates and rhizaria that consume larger prey, and finally metazooplankton, primarily copepods and krills but also to gelatinous species of metre dimensions. All categories of zooplankton include species that are raptorial or suspension feeders. Very few plankton models pay anything beyond lip-service to these matters.

With respect to protozooplankton, several of the contributed digital twin models recognized that evidence from field and laboratory studies has demonstrated the significance of protozooplankton grazing on primary producers (Sherr & Sherr 2002; Calbet & Landry 2004). Phagotrophic protists are capable of consuming all size classes of phytoplankton, from bacterial sized cyanobacteria to chain forming diatoms. The dilution assay (Landry & Hassett 1982; Calbet & Landry 2004) has quantified the role of microzooplankton, including nanoflagellates, ciliates, and heterotrophic dinoflagellates, as major activities in marine food webs. There is established evidence that heterotrophic dinoflagellates can be as important as copepods as consumers of diatom production (Saito et al. 2006; Sherr & Sherr 2007). In addition to their role as predators, phagotrophic protists are vital in nutrient recycling and are prey for metazoans (Stoecker & Pierson 2019).

For metazooplankton, interests expressed by contributors in digital twins of metazooplankton demand enhanced descriptions of behaviour (migration, prey selectivity) and physiology (digestion, fatty acid deposition/use, ontogeny). Individually, none of these present a challenge for modelling (e.g., Carlotti & Nival 1992; Gentleman et al. 2008). However, a major challenge is to mechanistically understand (and represent in models) how zooplankton simultaneously influence the fate of multiple elements (i.e., ecosystem stoichiometry and the phase [particulate vs dissolved] within which the elements sit), and how these impacts change in response to the direct (e.g., warming) and indirect (changes in food quantity and quality) effects of future climate change.

The importance of ontogeny in plankton models has been questioned by Clerc et al. (2021) and clearly has important implications in some applications (Flynn & Irigoien 2009; Cripps et al. 2014; Anderson et al. 2022). The extent to which ontogeny warrants descriptions of every individual life stage is unclear, but would in any case depend on the particular zooplankton and the simulation scenario. As is often the case, it is more the lack of empirical data and information that constrains progress; that is especially so when considering the effects of different prey types and the physiological status (and hence biochemical composition) of those prey for different ontogenic stages growing under different physical environments (temperature, pH, pressure etc.). While that concern absolutely applies to under-represented groups such as gelatinous zooplankton and chaetognaths, it applies even to the most studied groups including protist-zooplankton and copepods. As noted above, just attempting to construct plankton digital twins brings such short comings into sharp relief. Elsewhere in this project (Flynn et al. 2024a), contributors stated a preference for imperfect quasi digital twins with which to explore concepts and help enhance understanding, rather than waiting for impossible perfection; learning from failure is considered as beneficial.

An interest in coupling ‘omics to the digital twin was mentioned by several contributors (**Fig. 7**). This linkage is problematic in the extreme (Flynn et al. 2022) because of the inability of omics to provide rate measurements (e.g., Caron et al. 2017; Strzepek et al. 2022). One route through which omics can interface with digital twins is through descriptions of (de)repression of traits (Flynn & Mitra 2023b). The more usual way to interface omics with models is through steady-state flux balance analyses (Orth et al. 2010), but such modelling is not compatible with dynamic modelling as required for ecological research. It would only be appropriate for plankton digital twins operating in a chemostat setting.

4.2.2 Abiotic setting

Temperature was the most requested abiotic feature (**Fig. 3**). While also a not-infrequent feature of extant plankton models, it is also clear from responses elsewhere in this project (Flynn et al. 2024a) that including temperature is far more nuanced than is commonly appreciated by both modellers and empiricists (Atkinson et al. 2003; Simonds et al. 2010; Wang et al. 2019).

pH and aspects of DIC chemistry and alkalinity were also of interest to many (**Fig. 3**). There are two parts of this topic that are problematic. One is the responses of the organisms to changes in pH, H^+, pCO_2 etc.; that can be resolved through appropriate experimentation (e.g., Cripps et al. 2014, 2016; Flynn et al. 2015). More problematic, perhaps, is the modelling of the interactivity between plankton activity and pH and pCO_2; for example, net phototrophy raises pH which has a profound (negative) impact on plankton growth. While this bidirectional interaction has been described in models running at full ocean salinity (Flynn et al. 2015), the carbonate chemistry interaction is more complex at lower salinities down to freshwaters (Mook & Koene 1975; Cai & Wang 1998) which hinders such a rigorous mathematical description of the equilibria and its bidirectional interaction with biological activity.

Light is a fundamental requirement for most of the aspirational digital twin models, described ideally as spectral irradiance within stated light:dark cycles. Typically, empirical studies do not measure the light spectrum (nor the action spectrum for the phototrophs; Brunet et al. 2014; Orefice et al. 2016), and indeed only light as photon flux density at the surface of culture vessels is usually reported in empirical studies. Relating light to what is actually happening within the plankton community is all too often poorly measured (Brunet et al. 2013; Brunet et al. 2007, 2008) While there are models that define changes in the light spectrum with water depth (Anderson 1993), these are of little use unless they are matched with corresponding action spectra and photosynthesis rates (including DOC leakage) from the phototrophs.

Culture settings are easily configured in models; continuous flow, or partial harvesting of suspensions with instantaneous and 100% efficient fractionation of plankton types to replicate studies such as dilution experiments are conceptually simple to model. In addition, models can provide instantaneous reporting of individual rate processes, and even of isotope fractionations (Flynn et al. 2018).

4.2.3 Input/output interface

The simplest graphic user interface (GUI) required to satisfy the requirements of the concepts described by contributors would use input tables and/or sliders to select for choices of plankton types, trait expression, inoculum values, nutrient and other abiotic features. Analogies could be drawn with setting up characters in a video game. Scientists, however, and especially given the complexity of the multi-stressors described by the contributors to this work, would likely be more comfortable using a more comprehensive input system. That could perhaps allow users to have more control over detailed facets of individual plankton ‘players’, and to be able to explore the autecology of them before committing them to an *in silico* playground for the full digital twin scenario.

Contributors to this project described outputs exclusively in the form of time and x,y scatter plots. In reality, it may be expected that a fair degree of post-simulation analysis would be desired. That would require an ability to download the model output; at the most basic level that would be to a spreadsheet.

5. Concluding comments

From the above, the general expectation for a plankton digital twin, at least for first-generation efforts, comprises an *in silico* (i.e. computer based, simulation) laboratory in which a range of simulated plankton types would be described to high levels of detail with respect to their physiology and behaviour, placed within an abiotically detailed but dimensionally simple (0D or 1D) setting for short-period incubations.

The greatest challenge in constructing plankton digital twins is perhaps knowing where to stop. The contributions to this project (Flynn et al. 2024b) indicate that significant progress can be made by constructing a flexible product, perhaps with a maximum of 10 plankton functional types, each of which can be tailored to best equate to the desired trait descriptions. Traditional plankton models are extremely simple, and lack the scope for such tailoring. Alternative structures, however, involving extensive feedback controls that enable descriptions of (de)repression and behavioural traits (Flynn & Mitra 2023b), may provide a suitable starting point. An alternative would be to configure a multitude of very simple descriptions. The problem there is whether such base models could even approach the expectations of plankton empiricists for each plankton type; the responses in Flynn et al. (2024a) suggest not. Either way, producing a complete, comprehensive and useful digital twin platform to meet aspirations will be challenging. Inevitably it will also be an iterative process involving empiricists and computer scientists. Plankton science can only benefit from such tasks, not least because it brings into focus what we do and do not understand. In addition, there will always be a place for simple models in making advances in overarching concepts.

References

Anderson, T. R. (1993). A spectrally averaged model of light penetration and photosynthesis. Limnology and Oceanography, 38(7), 1403-1419.

Anderson, T. R., Christian, J. R., & Flynn, K. J. (2015). Modeling DOM biogeochemistry. In *Biogeochemistry of marine dissolved organic matter* (pp. 635-667). Academic Press.

Anderson, T. R., Hessen, D. O., Gentleman, W. C., Yool, A., & Mayor, D. J. (2022). Quantifying the roles of food intake and stored lipid for growth and development throughout the life cycle of a high-latitude copepod, and consequences for ocean carbon sequestration. *Frontiers in Marine Science*, *9*, 928209.

Atkinson D, Ciotti BJ, Montagnes DJS (2003) Protists decrease in size linearly with temperature: ca. 2.5% °C–1. Proc. R. Soc. Lond. B.2702605–2611. http://doi.org/10.1098/rspb.2003.2538

Brindley, C., Jiménez-Ruíz, N., Acién, F. G., & Fernández-Sevilla, J. M. (2016). Light regime optimization in photobioreactors using a dynamic photosynthesis model. Algal Research, 16, 399-408.

Brunet, C., & Lavaud, J. (2010). Can the xanthophyll cycle help extract the essence of the microalgal functional response to a variable light environment?. *Journal of Plankton Research*, *32*(12), 1609-1617.

Brunet C., Casotti R., Aronne B. and Vantrepotte V. (2003). Measured photophysiological parameters used as tools to estimate vertical water movements in the coastal Mediterranean. J. Plankton Research, 25(11): 1413-1425. doi: 10.1029/2002JC001541

Brunet, C., Casotti, R., Vantrepotte, V., & Conversano, F. (2007). Vertical variability and diel dynamics of picophytoplankton in the Strait of Sicily, Mediterranean Sea, in summer. *Marine Ecology Progress Series*, *346*, 15-26.

Brunet C., Casotti R. and Vantrepotte V. (2008). Phytoplankton diel and vertical variability in photobiological responses at a coastal station in the Mediterranean Sea. J. Plankton Res., 30: 645-654. Doi: 10.1093/plankt/fbn028

Brunet C., Conversano F., Margiotta F., Dimier C., Polimene L., Tramontano F. and Saggiomo V. (2013). The role of light and photophysiological properties on phytoplankton succession during the spring bloom in the NorthWestern Mediterranean sea. Advances in Oceanography and Limnology, 4: 1-19. doi : 10.1080/19475721.2013.779937

Brunet, C., Chandrasekaran, R., Barra, L., Giovagnetti, V., Corato, F., & Ruban, A. V. (2014). Spectral radiation dependent photoprotective mechanism in the diatom *Pseudo-nitzschia multistriata*. *PLoS One*, *9*(1), e87015.

Cai, W. J., & Wang, Y. (1998). The chemistry, fluxes, and sources of carbon dioxide in the estuarine waters of the Satilla and Altamaha Rivers, Georgia. Limnology and Oceanography, 43(4), 657-668.

Calbet, A. , & Landry, M. R. , (2004). Phytoplankton growth, microzooplankton grazing, and carbon cycling in marine systems. Limnology and Oceanography, 49 (1), 51-57, doi: 10.4319/lo.2004.49.1.0051.

Carlotti, F., & Nival, P. (1992). Model of copepod growth and development: moulting and mortality in relation to physiological processes during an individual moult cycle. Marine Ecology Progress Series, 84, 219-233.

Caron, D. A., Alexander, H., Allen, A. E., Archibald, J. M., Armbrust, E. V., Bachy, C., ... & Worden, A. Z. (2017). Probing the evolution, ecology and physiology of marine protists using transcriptomics. *Nature Reviews Microbiology*, *15*(1), 6-20.

Chen, X., Weinbauer, M. G., Jiao, N., & Zhang, R. (2021). Revisiting marine lytic and lysogenic virus-host interactions: Kill-the-Winner and Piggyback-the-Winner. *Science Bulletin*, *66*(9), 871-874.

Clerc, C., Aumont, O., & Bopp, L. (2021). Should we account for mesozooplankton reproduction and ontogenetic growth in biogeochemical modeling?. *Theoretical Ecology*, *14*(4), 589-609.

Costache, T. A., Acien Fernandez, F. G., Morales, M. M., Fernández-Sevilla, J. M., Stamatin, I., & Molina, E. (2013). Comprehensive model of microalgae photosynthesis rate as a function of culture conditions in photobioreactors. *Applied microbiology and biotechnology*, *97*, 7627-7637.

Cripps, G., Lindeque, P., & Flynn, K. (2014). Parental exposure to elevated pCO_2 influences the reproductive success of copepods. *Journal of plankton research*, *36*(5), 1165-1174.

Cripps, G., Flynn, K. J., & Lindeque, P. K. (2016). Ocean acidification affects the phyto-zoo plankton trophic transfer efficiency. *PloS One*, *11*(4), e0151739.

Dimier, C., Corato, F., Tramontano, F., & Brunet, C. (2007). Photoprotection and xanthophyll-cycle activity in three marine diatoms 1. *Journal of Phycology*, *43*(5), 937-947.

Dimier, C., Giovanni, S., Ferdinando, T., & Brunet, C. (2009). Comparative ecophysiology of the xanthophyll cycle in six marine phytoplanktonic species. *Protist*, *160*(3), 397-411.

Durbin, E. G. (1974). Studies on the autecology of the marine diatom *Thalassiosira nordenskiöldii* Cleve. 1. The influence of daylength, light intensity, and temperature on growth 1. *Journal of Phycology*, *10*(2), 220-225.

Flynn, K. J. (2024) *Simulating Plankton – getting it right in the era of Digital Twins of The Ocean; project introduction and executive discussion.* Zenodo https://doi.org/10.5281/zenodo.10953377

Flynn, K. J., & Butler, I. (1986). Nitrogen sources for the growth of marine microalgae: role of dissolved free amino acids. *Marine Ecology Progress Series*, 34, 281-304.

Flynn, K. J., & Mitra, A. (2023a). Feeding in mixoplankton enhances phototrophy increasing bloom-induced pH changes with ocean acidification. *Journal of Plankton Research*, *45*(4), 636-651.

Flynn, K. J. & Mitra, A. (2023b). DRAMA – a cybernetic approach for Plankton Digital Twins. Zenodo. https://doi.org/10.5281/zenodo.7848329

Flynn KJ, Mitra A, Bode A (2018) Toward a mechanistic understanding of trophic structure: inferences from simulating stable isotope ratios. Marine Biology 165:147 https://doi.org/10.1007/s00227-018-3405-0

Flynn, K. J., Mitra, A., Anestis, K., Anschütz, A. A.. Calbet, A., Ferreira, G. D. …. & Traboni, C. (2019). Mixotrophic protists and a new paradigm for marine ecology: where does plankton research go now? Journal of Plankton Research, 41,375-391.

Flynn, K. J., Sherr, E., Stoecker, D. K., Thingstad, T. F., Atkinson, A., Beardall, J., Boersma, M., Bowler, C., Calbet, A., Dam, H., Glibert, P. M., Hansen, P. J., Jin, P., Lomas, M. W., Lønborg, C., Menden-Deuer, S., Mock, T., Needham, D., Robinson, C., Rokitta, S., Rost, B., Saiz, E., Schmidt, K., Thiele, S. & Våge, S. (2024a). *Simulating Plankton – getting it right in the era of Digital Twins of The Ocean; building and challenging perceptions.* Zenodo https://doi.org/10.5281/zenodo.10952555

Flynn, K. J., Sherr, E., Atkinson, A., Lønborg, C., Mulholland, M., Rokitta, S., Rosk, B., Schmidt, K., Thiele, S. & Thingstad, T. F. (2024b). *Simulating Plankton - getting it right in the era of Digital Twins of The Ocean; simplicity vs complexity*. Zenodo https://doi.org/10.5281/zenodo.10953197

Flynn, K. J., & Irigoien, X. (2009). Aldehyde-induced insidious effects cannot be considered as a diatom defence mechanism against copepods. *Marine Ecology Progress Series*, *377*, 79-89.

Flynn, K. J., Clark, D. R., Mitra, A., Fabian, H., Hansen, P. J., Glibert, P. M., ... & Brownlee, C. (2015). Ocean acidification with (de)eutrophication will alter future phytoplankton growth and succession. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1804), 20142604.

Flynn, K. J., Mitra, A., Wilson, W. H., Kimmance, S. A., Clark, D. R., Pelusi, A., & Polimene, L. (2022a). 'Boom-and-busted' dynamics of phytoplankton–virus interactions explain the paradox of the plankton. *New Phytologist*, *234*(3), 990-1002.

Flynn, K. J., Torres, R., Irigoien, X., & Blackford, J. C. (2022b). Plankton digital twins—a new research tool. *Journal of Plankton Research*, *44*(6), 805-805.

Gentleman, W. C., Neuheimer, A. B., & Campbell, R. G. (2008). Modelling copepod development: current limitations and a new realistic approach. ICES Journal of Marine Science, 65(3), 399-413.

Glibert, P. M. (2016). Margalef revisited: a new phytoplankton mandala incorporating twelve dimensions, including nutritional physiology. *Harmful Algae*, *55*, 25-30.

Landry, M. R. & Hassett, R. P. (1982). Estimating the grazing impact of marine micro-zooplankton. Marine Biology 67, 283-288.. 10.1007/BF00397668

Leles, S. G., Bruggeman, J., Polimene, L., Blackford, J., Flynn, K. J., & Mitra, A. (2021). Differences in physiology explain succession of mixoplankton functional types and affect carbon fluxes in temperate seas. Progress in Oceanography, 190, 102481.

Li, A., Stoecker, D. K., & Coats, D. W. (2000). Mixotrophy in *Gyrodinium galatheanum* (Dinophyceae): grazing responses to light intensity and inorganic nutrients. Journal of Phycology, 36, 33-45

Margalef, R. (1978). Life-forms of phytoplankton as survival alternatives in an unstable environment. *Oceanologica acta*, *1*(4), 493-509.

Maselli, M., Van de Waal, D. B., & Hansen, P. J. (2022). Impacts of inorganic nutrients on the physiology of a mixoplanktonic ciliate and its cryptophyte prey. Oecologia, 199, 41-52.

Mateus, M. D. (2017). Bridging the gap between knowing and modeling viruses in marine systems—an upcoming frontier. *Frontiers in Marine Science*, *3*, 284.

Meyer, N., Rydzyk, A., & Pohnert, G. (2022). Pronounced uptake and metabolism of organic substrates by diatoms revealed by pulse-labeling metabolomics. *Frontiers in Marine Science*, *9*, 821167.

Mitra, A., Castellani, C., Gentleman, W. C., Jónasdóttir, S. H., Flynn, K. J., Bode, A., ... & John, M. S. (2014). Bridging the gap between marine biogeochemical and fisheries sciences; configuring the zooplankton link. *Progress in Oceanography*, *129*, 176-199.

Mitra, A., Flynn, K. J., Stoecker, D. K., & Raven, J. A. (2023). Trait trade-offs in phagotrophic microalgae: the mixoplankton conundrum. *European Journal of Phycology*, 1-20.

Mook, W. G., & Koene, B. K. S. (1975). Chemistry of dissolved inorganic carbon in estuarine and coastal brackish waters. *Estuarine and Coastal Marine Science*, 3(3), 325-336.

Murray, A. G., & Jackson, G. A. (1992). Viral dynamics: a model of the effects of size, shape, motion and abundance of single-celled planktonic organisms and other particles. *Marine Ecology Progress Series*, 103-116.

Orefice, I., Chandrasekaran, R., Smerilli, A., Corato, F., Caruso, T., Casillo, A., ... & Brunet, C. (2016). Light-induced changes in the photosynthetic physiology and biochemistry in the diatom Skeletonema marinoi. *Algal research*, *17*, 1-13.

Orefice, I., Musella, M., Smerilli, A., Sansone, C., Chandrasekaran, R., Corato, F., & Brunet, C. (2019). Role of nutrient concentrations and water movement on diatom's productivity in culture. *Scientific reports*, *9*(1), 1479.

Orth, J. D., Thiele, I., & Palsson, B. Ø. (2010). What is flux balance analysis?. *Nature biotechnology*, *28*(3), 245-248.

Paasche, E. (1968). Marine plankton algae grown with light-dark cycles. II. *Ditylum brightwellii* and *Nitzschia turgidula*. *Physiologia plantarum*, *21*(1), 66-77.

Pistelli, L., Del Mondo, A., Smerilli, A., Corato, F., Sansone, C., & Brunet, C. (2023). Biotechnological response curve of the cyanobacterium Spirulina subsalsa to light energy gradient. *Biotechnology for Biofuels and Bioproducts*, *16*(1), 28.

Polimene, L., Brunet, C., Allen, J. I., Butenschön, M., White, D. A., & Llewellyn, C. A. (2012). Modelling xanthophyll photoprotective activity in phytoplankton. *Journal of plankton research*, *34*(3), 196-207.

Polimene L., Brunet C., Butenschon M., Martinez-Vicente V., Widdicombe C., Torres R. and Allen J.I. (2014). Modelling a light driven phytoplankton succession. Journal of Plankton Research, 36: 214-229. doi: 10.1093/plankt/fbt085

Rothschild, B. J., & Osborn, T. R. (1988). Small-scale turbulence and plankton contact rates. *Journal of plankton Research*, *10*(3), 465-474.

Saito, H., Ota, T., Suzuki, K., Nishioka, J., & Tsuda, A. (2006). Role of heterotrophic dinoflagellate Gyrodinium sp. in the fate of an iron induced diatom bloom' Geophysical Research Letters, 33, L09602, doi:10.1029/2005GL025366.

Sherr, E.B., & Sherr, B. F. (2002). Significant of predation by protists in aquatic microbial food webs. Antonie Van Leewenhoek International Journal of General and Molecular Microbiology 81, 293-308.

Sherr, E. B, & Sherr, B. F. (2007). Heterotrophic dinoflagellates: A significant component of microzooplankton biomass and major grazers of diatoms in the sea. Marine Ecology-progress Series, 352, 187-197, DOI:10.3354/meps07161

Silkin V, Fedorov A, Flynn KJ, Paramonov L, Pautova L (2021) Protoplasmic streaming of chloroplasts enables rapid photoacclimation in large diatoms. Journal of Plankton Research (2021) 43: 831–845. https://doi.org/10.1093/plankt/fbab071

Simonds, S., Grover, J. P., & Chrzanowski, T. H. (2010). Element content of *Ochromonas danica*: a replicated chemostat study controlling the growth rate and temperature. *FEMS microbiology ecology*, *74*(2), 346-352.

Stoecker, D., & Pierson, J. (2019). Predation on protozoa: its importance to zooplankton revisited. Journal of Plankton Research, 41, 367-373.

Strzepek, R. F., Nunn, B. L., Bach, L. T., Berges, J. A., Young, E. B., & Boyd, P. W. (2022). The ongoing need for rates: can physiology and omics come together to co-design the measurements needed to understand complex ocean biogeochemistry?. *Journal of Plankton Research*, *44*(4), 485-495.

Thingstad, T. F. (2000). Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnology and Oceanography*, *45*(6), 1320-1328.

Wang, Q., Lyu, Z., Omar, S., Cornell, S., Yang, Z., & Montagnes, D. J. (2019). Predicting temperature impacts on aquatic productivity: Questioning the metabolic theory of ecology's “canonical” activation energies. *Limnology and Oceanography*, 64, 1172-1185.

Appendix 1 Task Guidance

Other than a renumbering of Figures, the following is as provided to the contributors.

a) Scenario

First, **name your subject**! Please make the subject at least as detailed as a subset of a larger grouping (e.g, for phytoplankton as '*small chain-forming diatom*', or for zooplankton as '*HNF*'). This is to ensure that the conceptual framework has a level of detail linked to reality which is useful. You can go all the way to a named species if you wish.

If your organism has life stages that you wish the model to described (e.g., egg, nauplii, copepodite, adult; cyst, vegetative stage; etc.), then you may need to consider some of the following tasks separately for each stage, as appropriate.

For a model describing a virus or parasite you may likely need to also describe the host organism in a dynamic setting. For a mixoplankton or zooplankton, you can assume the supply of inert prey, perhaps of different sizes and stoichiometries, but you do not need to worry about modelling the dynamics of prey growth.

Please then add a **single short sentence describing what this model could be used to describe** – what could you explore with the model? For example: virus-host dynamics in sediment-laden water; bacterial transformation of DOM of different labilities; light-nutrient ecophysiology of *Emiliania* including DMSP synthesis and release; growth of *Calanus* in response to changes in diet.

Identifying the role is important, as it affects what you need to consider later on. For example, if your topic is within a scenario of short term incubations, do you really need to worry about aspects of changes in ontogeny (as distinct from having the option to run the model with different stages)?

If you wish, add an illustration; please include a scale bar if that is appropriate.

b) State variables

List the *state variables*; these describe the core building blocks of the organism model itself. State variables have a history to them; in a real world situation you could collect them, or an analogue, on a filter or in a bottle for analysis. Check the main project description for more information.

Please include whole organism C and N in your list, as these are core requirements for plankton models. Then add others as you require (e.g., numeric abundance, P, Fe, PUFA, toxin, Chl, Rubisco, acquired photosystems, metabolite pool, gut, feeding vacuole, etc.). In practice, often these are estimated by proxy (most obviously C from biovolume) – the emphasis here is on including parameters that you see as being important for describing the dynamics of plankton growth.

State variables for plankton model typically have units of water volume (either L^{-1} or m^{-3}); this allows the partial transfer (exchange) of materials between adjacent blocks of water. Organism size is thus given by calculation (e.g., (µgC L^{-1})/(nos. L^{-1}) = µgC individual^{-1}); you do not need a state variable for 'size'.

If your plankton type has multiple life stages you may, consistent with your interests, also need 'age', or some other feature (such as a critical size), to provide an ontogenic clock.

Now please rank these state variables as a **prioritised list** (where 1 is essential). Just as in a real experiment you are limited by pragmatism; here, computational loads may limit what you can actually do. You will realise very quickly as you work through the following that as the state variable count increases, so the complexity of the interactions multiplies!

If you are of the opinion that it is not plausible to have a model for this plankton with anything less than a certain set if state variables, you can rank those essential state variables all collectively as '#1'. Please do not list more than 4 items as essential and thus ranked as #1; think, in your experiments, how many plankton parameters do you always measure?

Now, please **state the units for each state variable**; typically state variables have units of amount m^{-3}, or L^{-1}. Please use the same volumetric unit throughout. For example, in terms of N, biomass, inorganic and organic nutrients would all have the same unit, here as for e.g., mgN m^{-3} (which as a numeric value is the same as µgN L^{-1}).

c) Resources: acquisition preference and regulation

Each of the elemental (and other chemical components if applicable, such as fatty acids) state variables in your plankton have to be obtained from external resources.

List the resources your plankton needs.

Again, **prioritise this list** to reflect what is most important to the organisms consistent with your interests . If not central to your interests, do you need to include vitamins, for example?

For each of these resources, again, **state the units**; organics will contain C, perhaps with N and P, for example. In some instances (e.g. for dissolved free amino or nucleic acids) you may be able to state the N:C and P:C as constants so you only need a single state variable for the resource (e.g., µgC L^{-1}). However, if the resource stoichiometry may vary over time, then you need state variables for each component (e.g., you could have 'DOM' described as a single entity, with a variable C:N:P, but you then need DOM described by three state variables, as µgC L^{-1}, µgN L^{-1}, µgP L^{-1}).

For each resource, **describe the form of the response curve relating abundance of the resource to its acquisition rate**. Most will either be a rectangular hyperbola (dissolved nutrient), or linear with encounter, but factors such as motility, turbulence, organism size, production of mucus, colony formation, etc. will affect these relationships, so **mention those as appropriate**. It may help to check the Project Introduction document.

What are the units on those response curve axes? – do you have those units available to you from the state variables? For example, if your grazing function relates to numeric abundance, do you have 'abundance' of your prey, and/or of your predator declared as a state variable?

As appropriate, please provide information on the **size scales** of different resource types (e.g., Dalton size for macromolecules supporting bacteria osmotrophy that can be directly transported into the cell vs the size of substrates that require extracellular digestion; µm prey size for predators).

Try to **assign a priority to resource preference**; if there is insufficient of a particular resource to satisfy demands, which other potential resources does the organism try and exploit next? (e.g. arginine ≥ ammonium > nitrate > N_2-fixation; DIP ≥ DOP; cryptophyte > large dinoflagellate > cyanobacteria; narrow optimal prey size > exploit larger prey > exploit smaller prey).

Try to **relate the priority of resource preference to some feature of the state variables**, such as N-resource transport to the organisms N:C, prey selection to satiation (e.g., perhaps to the metabolite pool:C or to gut satiation), and then try and configure a suite of curves akin to **Fig.A1**.

This only needs to be your best guess at what may be controlling trait expression! The critical thing is to try and relate features of resource acquisition to priority of their expression linked to some physiological feature that the model describes.

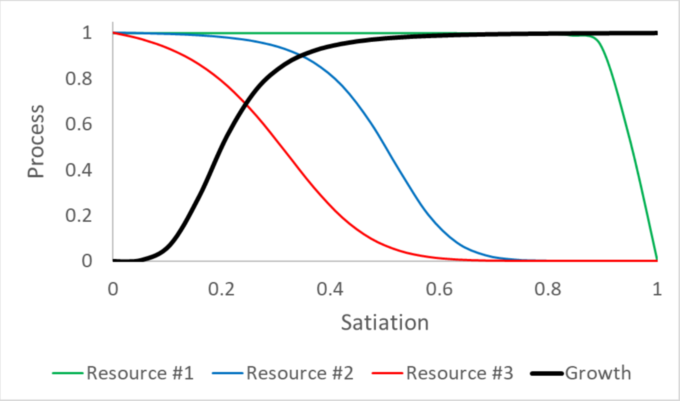

Fig.A1. Hypothetical series of resource acquisition controls related to the nutrient status of an organism. Satiation (e.g., related to the organism N-status or the size of the metabolite pool, or gut content), where 1 is optimal and 0 is starved, controls resource acquisition from 3 alternative sources and also supports growth of the organism. Here, resource #1 is always exploitable (its use is only repressed when satiation is maximum) and supports the maximum growth potential; the pathways to exploit #2 and #3 are only de-repressed as the organism becomes increasingly stressed (starved). Here, growth at high rates using #3 is not possible; this route is more of a survival mechanism. 'Preference' between resources as shown here would be #1>#2>#3. It is possible that the acquisition route for a resource would be de-repressed early as satiation declines but alone such a resource may be unable to support a high growth rate; an example is the use of the amino acid histidine which may be taken up via the arginine-transporter (which may always be 'on' as arginine is an excellent source of N and an amino acid that is readily exploited) but histidine is an amino acid that is slow to metabolise.

c.1) Multi-stress interactions

Consider how limitations due to the lack of each resource may impact the acquisition of others. For example, a lack of Fe affects phototrophy and the synthesis of enzymes for the assimilations of nitrate and N_2. Light and/or P limitation may affect N-acquisition, and also cell size. You may be able to summarise such events in a table like that shown in **Fig.A2**.

Limiting resource	Primary response	Potential secondary response
Light	[Chl:C & P_{max}]↑, mixotrophy↑	N:C↑, P:C↑, Fe:C↑, diatom Si:C↑, C:cell↓
N	N:C↓, use of alternates ↑, e.g. (NO_3, DON, N_2) mixotrophy↑	P:C↑, diatom Si:C↑, [Chl:C & P_{max}]↓, C:cell↓
P	P:C↓ use of DOP ↑, mixotrophy↑	C:cell↑, diatom Si:C↑, N:C↓, [Chl:C P_{max}]↓
Fe	Fe:C↓, siderophores ↑, mixotrophy↑	P:C↑, diatom Si:C↑, [Chl:C P_{max}]↓, C:cell↓
Si (diatoms)	Si:C↓, thin frustules, short setae	C:cell↑, P:C↑, [Chl:C P_{max}]↓

Fig.A2 Suggested primary and secondary stress responses for phytoplankton. The primary responses typically include immediate stoichiometric changes and patterns of (de)repression control through which the organism attempts to compensate for the stress (such as 'turning on' other nutrient acquisition pathways). Mixotrophy for phytoplankton is enabled through osmotrophy, the use of dissolved organic nutrients. The primary responses are those typically measured in experiments; ‘side effects', termed here as 'secondary’ responses, can have important implications for the drawdown of other nutrients and for predator-prey interactions via stoichiometric ecology.

If you cannot make a reasoned guess as to the form of any relationship controlling resource acquisition (as in **Fig.A1**), at least try and consider if the feedbacks are likely to be positive or negative (up vs down arrows in **Fig.A2**).

c.3) Behaviour and gross physiology

List what **facets of behaviour are exhibited by your chosen plankton**; growth and ontogeny (etc.) are considered in the next section.

These may include:

- changes in motility with turbulence, with satiation, or the presence of predators or of toxic prey, diel vertical migration;
- changes in physiology (detailed to specific processes, or to the overall growth rate potential) with T, light:dark periodicity, pH, salinity, sediment load, etc.

Again **prioritise this list** as to what is most important to the organisms, that you expect the digital twin to describe, as consistent with your interests. Of course, if the aim of your model was not to explore sediment load, for example, than even if this feature may be important under other circumstances, you do not need to list it here.

Can any of these behaviours be related to the physiological state of the plankton as may be deduced from the state variables? For example, a low cellular N:C might drive a dinoflagellate to migrate down to acquire ammonium or to feed.

c.4) Growth and development

Organism growth is controlled by the availability of resources inside them (which, in turn, is affected by what is outside).

List what features of **internal resource availability**, consistent with your interests, control the growth of your plankton. This could relate to organism stoichiometry (as C:N:P, as used in quota models), the relative size of the metabolite pool, to the abundance of a specific macromolecule such as an essential fatty acid, or to the balance of diet (quality and quantity related to the nutrient status of ingested prey). Some of these features may have already been listed with ‘Behaviour and gross physiology’; list them again here as required.

Growth also involves changes to the average cell-size for microbes, metazoan size at ontogenic stage transitions, maturity, or entry to resting stages. List **life-cycle changes** that you expect the model to describe what stimulates them and what affects these have. For example, protist cell size may vary with N or P limitations.

The consequence of mortality is an important output: what promotes death, is it age-related, UV light, exhaustion of Si-nutrient, what? In a starvation situation, consider how the organism responds (other than altering its resource acquisition priorities, which you will have already considered) – some die very quickly, while others can last for a long time (especially if engaging in intra-guild cannibalism).

List the abiotic factors that affect these processes? Factors to consider may include temperature (T), light (wavelength, PFD, light:dark periodicity), O_2, pH, salinity.

Again, prioritise your lists.

As before, if you can sketch response curves describing any of these interactions, please do so.

c.5) Release of materials (including losses, regenerations, leakages etc.)

What an organism releases is important to the ecology of the system in which it grows. Ultimately, materials that are released may also come around and after transformations constitute resources.

Please **list the materials released during growth and activity** of the plankton in the model. These may include dissolved inorganics (nitrite, nitrate, ammonium, phosphate), organics (DOC, DFAA, toxins, allelopaths, mucus, partial products of extracellular digestion, etc.), particulate structures (scales, diatom valves or skeleton moults), micro-faecal pellets (from protists), faecal pellets, aborted or vacated prey traps, etc.

For each of these releases, again as usual, please state the units (C,N,P, L^{-1}, m^{-3} etc.).

Please prioritise your list of released materials for inclusion in the model.

For each release type, please try to provide a **response curve relating the release to a physiological or abiotic feature**. This may include the relative acquisitions of different resources (e.g., enhanced DOC or toxin release when phytoplankton exhaust nutrients), the level of food satiation, temperature, light, salinity. Releases may relate to sudden changes in conditions (e.g., DMSP with salinity).

d) The Digital Twin Interface

Having completed the design of the plankton model, you need to consider how you might control it, and what you want to visualise. Please remind yourself of the use for which you intended your model before proceeding.

The design and usage of the interface is what defines a digital twin as compared to a regular simulation model. The latter is only available for experts in modelling to use (often only the originators of the computer code), while the digital twin (which may use the same simulation engine to provide the actual mathematical and other computational activities) is designed to be usable by anyone with an interest to do so.

d.1) Input control

Some of the features you have described will be relatively fixed, such as those configuring fundamental features of the organism type; these would be controlled via a configuration file. Others, however, you would wish to 'play' with in real time, by pressing buttons, moving sliders, or entering a few numbers.

Please **provide a prioritised list of those features you would wish to adjust in real time**. This would likely include resource types and abundance (addition of nutrients, prey or hosts, etc.), and certain abiotic features (temperature, pH, light, light:dark cycle periodicity). It may also include changing the order of resource preference (perhaps by changing the format of response curves such as those in **Fig.1**).

Would being able to operate the model in a chemostat mode be useful, with an adjustable dilution rate? Or would batch mode be sufficient, and if so over what typical duration (days, weeks)? Or both?

Would a 'rewind' option be useful, where you could take your simulation back to part way through the time course, change some parameters and then re-start it?

d.2) Output visualisation

While a full data download would be possible, invariably it is useful to watch certain key features change over time while the simulation is running. These visualisation (at least in initial digital twin presentations) will be provided as time or x,y plots.

Typically the major elements (C, N, P etc.) are plotted in all their different forms (e.g., biomass-C, prey-C, dissolved inorganic-C, organic-C, etc); these may also be plotted together with the sum-total for that element, which should remain constant in a batch scenario to check it all adds-up. But **which other data types would you most wish to visualise in real time**? For example, N:C, Chl:C, abundance, encounter rates, etc. all as time plots; or growth rate vs prey abundance, growth vs N:C quota etc. as x,y plots.

Please provide a prioritised list.

Appendix 2

Table A2.1 Summary of plankton digital twin descriptions by contributors. ‘1’ indicates a requirement.

		Scenario						Plankton Function Type								Abiotics							Nutrients						Abiotic outputs					Biotic-related outputs							Physiology rates and allied						
Surname	**key words**	laboratory/mesocosm	field	short term (days-1mo)	long term (>mo)	chemostat	rewind facility	virus	bacteria/archaea	phytoplankton (inc cyano)	mixoplankton	protist zooplankton	metazoan zooplankton	multiple PFT	detailed physiology for PFTs	T	S	Depth/pressure	Light (PFD and L:D cycle)	pH	DIC, pCO2	O2	DIN	DIP	DISi	DFe	DOC	DOM	pH, pCO2	O2	DI nutrients	DO nutrients	debris	biomass	numbers	Chl	other photo-pigments	variable stoichiometry	fatty acid	allelopath, toxin	detailed physiology	detailed photophysiology	variable prey/host selectivity	microplastics interactions	ontogeny	DVM, depth-linked activity	explicit omics linkage
Atkinson A	Diatom-Krill-Fe		1	1						1			1	1	1								1			1					1		1	1	1			1			1						
Beardall J	Cfix-pH-T-pCO2	1		1						1					1	1			1	1	1		1	1		1			1	1	1			1		1	1	1			1	1					
Berges JA	Phyto/Mixo death	1		1		1	1			1	1			1	1								1	1	1	1	1	1				1	1	1	1	1		1		1	1						
Boersma M	Diatom-copepod-ctenophore	1	1	1	1	1				1			1	1	1	1			1	1	1		1	1			1				1	1	1	1	1			1	1	1	1		1		1		
Brunet C	Diatom-nondiatom-Cfix-PFD	1		1						1					1				1											1				1	1	1	1	1			1	1					
Calbet A	Protozoan-mixo-copepod	1		1						1	1	1	1	1	1	1			1		1		1	1						1	1		1	1	1	1		1			1		1		1		
Caron D	Foram-symbiont-lifeCycle	1		L			1				1	1		1	1	1		1	1				1	1							1	1		1	1	1		1			1				1	1	
Dam H	Copepod-dino-toxin	1		1							1		1	1	1	1			1		1		1	1					1		1	1		1	1	1		1		1	1		1		1		1
Glibert P	Mixo-HAB-competition		1	1	1						1				1	1	1	1	1				1	1							1			1	1	1		1		1	1						
Hansen PJ	Mixo-predPrey-pH	1	1	1	1				1	1	1	1		1	1	1	1	1	1	1	1		1	1	1				1		1		1	1	1	1		1		1	1		1				
Jin P	Phyto-diatom-cocco-mixo-Cfix	1		1						1	1			1	1	1			1		1		1	1					1		1			1	1	1	1	1			1	1					
Lomas M	Phyto-DIP	1		1						1					1								1	1							1			1		1		1			1						
Lønborg C	bacteria-DOM-growth efficiency	1		1			1		1						1	1				1	1	1	1	1			1	1	1	1	1	1		1	1			1			1						
Mayor D	copepod-pred/prey-stoichiometry	1		1						1			1	1	1	1		1		1	1	1	1	1					1	1	1		1	1	1			1	1	1	1		1	1	1	1	
Meyer B	Krill-diatom-FA		1	1	1					1			1	1	1	1		1															1	1	1			1	1		1				1		
Millette N	Mixo-bloomDynamics		1	1	1				1		1		1	1	1	1			1				1	1							1			1	1	1		1	1		1						
Mock T	Diatom-bacteria-evolution	1		1	1				1	1				1	1	1			1				1			1	1	1			1	1		1	1	1		1		1	1						1
Mulholland M	Diazotroph-symbiont-diatom		1	1	1					1	1			1	1	1		1	1		1		1	1			1	1	1		1	1		1	1	1		1		1	1						1
Poulton,A	Emiliania-lifeCycle-virus	1	1	1	1			1			1			1	1	1			1	1	1		1	1		1		1	1				1	1	1	1		1			1				1		
Robinson C	bacteria-DOM	1		1	1				1			1		1	1	1		1		1	1	1					1	1	1	1	1	1	1	1	1			1			1						
Rokitta S	phyto-Cfix-photoInhibition-DOC	1		1	1					1					1	1			1	1	1		1	1	1	1	1		1		1	1		1	1	1		1			1						
Rost B	diatom-phyto-competition	1		1		1	1			1	1			1	1	1			1	1	1	1	1	1	1	1			1	1	1	1		1	1	1	1	1		1	1	1					
Saiz E	Copepod-phyo-mixo	1		1						1	1	1	1	1	1	1			1		1		1	1						1	1		1	1	1	1		1			1		1		1		
Scanlan D	Synechococcus-Prochlorococcus-virus	1		1				1		1				1	1	1		1	1	1	1		1	1		1	1	1	1	1	1	1		1	1	1		1	1		1	1					
Schmidt K	pico/nano phyto-mixo competition	1		L						1	1			1	1	1							1	1		1	1	1			1	1		1	1	1		1		1	1						
Sherr E	Gyrodinium-diatom-bloom-copepod	1	1	1						1	1		1	1	1	1			1				1	1	1						1	1	1	1	1	1		1			1		1		1		
Stoecker DK	ciliate-predPrey	1		1						1	1	1		1	1	1	1		1	1	1		1	1	1		1		1		1	1		1	1	1		1		1	1		1				
Svensen C	zooplankton-copepod-predPrey			1						1	1	1	1	1	1	1	1		1				1	1							1	1	1	1	1	1		1			1				1		
Thiele S	virus-bacteria-archaea-competition	1		1		1	1	1	1	1		1		1	1	1						1	1	1			1	1		1	1	1	1	1	1			1									
Thingstad TF	virus-bacteria-competition	1		1	1	1		1	1			1		1	1	1							1	1			1	1			1	1		1	1			1					1				
Unrein F	mixo-uZ-competition	1		1					1		1	1		1	1				1				1	1				1			1	1	1	1	1	1		1			1		1				
Våge S	bacteria-competition-virus-predPrey	1		S				1	1		1	1		1	1				1				1	1			1	1			1	1	1	1	1			1			1						