

# Hitting the sweet spot of complexity: Reasons why the development of new custom-tailored models is still warranted and should be encouraged in aquatic sciences

Gregorio A. López Moreira M.,<sup>1,2\*</sup> Marco Toffolon,<sup>2</sup> Franz Höfker<sup>1</sup>

<sup>1</sup>Department of Ecohydrology and Biogeochemistry, Leibniz-Institute of Freshwater Ecology and Inland Fisheries (IGB), Müggelseedamm 310, 12587 Berlin, Germany; <sup>2</sup>Department of Civil, Environmental and Mechanical Engineering, University of Trento, Via Mesiano 77, 38123 Trento, Italy

## ABSTRACT

Process-based aquatic ecosystem models are increasingly being developed and used in freshwater ecology and other aquatic sciences, as they are powerful tools to gain a mechanistic understanding of ecological processes and inform policy and decision making in environmental management. Over the last decades, not only have these models increased considerably in number, but also in their degree of complexity, which can improve predictive capacity. Nevertheless, it is also because of the higher degree of complexity of many models of current widespread use, that not all the hypotheses and assumptions upon which they have been built are always met by the relatively simple experiments that characterise fundamental ecological research. This is true for both laboratory experiments and those carried out outdoors, under semi-controlled conditions. Examples of the latter are the mesocosms experiments through which several novel questions are nowadays being addressed. In this article, we present our views on why the development of new custom-tailored aquatic ecosystem models of varying degrees of complexity is still very much warranted and should, therefore, be encouraged despite arguments in favour of always increasing complexity and against the creation of new models that are largely based on previously published ones ('*reinventing the wheel*'). Deciding on the right complexity level should be linked to the biological organisation levels that are relevant to the specific research questions, and to how much knowledge on the subject is already available. Spatial and temporal scales are additional factors that a modeller should weigh in when deciding on the complexity of a model. To address these needs in the long term, the modelling community needs to grow. Training a new generation of model developers will not only benefit other scientists to better design future experiments but will also facilitate interdisciplinary research and teamwork, approaches such as ensemble modelling, as well as the communication of science to managers and many other stakeholders.

Corresponding author: [ga.lopez@igb-berlin.de](mailto:ga.lopez@igb-berlin.de) / [ga.lopez@unitn.it](mailto:ga.lopez@unitn.it)

Key words: mechanistic models; ecological modelling; lake ecosystems; limnology; reduced complexity; young modellers.

Authors' contribution: GALMM, wrote the paper based on literature review and modelling work carried out as part of his doctoral and postdoctoral research projects. This work was jointly supervised by MT and FH, who also helped conceive the article and contributed to its contents with comments and edits.

Funding: This work was partially carried out within the Science for Management of Rivers and their Tidal Systems Erasmus Mundus Joint Doctorate (SMART EMJD) programme, funded by the Education, Audiovisual and Culture Executive Agency (EACEA) of the European Commission.

Conflict of interest: The authors declare no conflicts of interest. The funding agency had no role in the writing of the manuscript or the decision to publish it. The contents of the article, therefore, reflect the views of the authors only.

Received: 20 May 2021.

Accepted: 11 September 2021.

This work is licensed under a Creative Commons Attribution Non-Commercial 4.0 License (CC BY-NC 4.0).

©Copyright: the Author(s), 2021

Licensee PAGEPress, Italy

*J. Limnol.*, 2021; 80(3):2035

DOI: 10.4081/jlimnol.2021.2035

## INTRODUCTION

Process-based models help scientists to better understand field and laboratory observations, push theory development further and advance fundamental knowledge in multiple disciplines. This is especially true in aquatic sciences, where the development of such models has been a prolific endeavour over the last few decades, both in number and diversity (Trolle *et al.*, 2012; Janssen *et al.*, 2015). Because these models provide a mechanistic basis for the interpretation of observed phenomena, they make it possible to reconstruct the past (hindcasting), helping us to fill knowledge gaps through simulation and explain the present conditions of a system as a result of its past evolution (path dependence). Based on the present, they also enable scientists to estimate the most likely immediate future (forecasting) and even attempt long term predictions.

Once calibrated and validated, process-based models allow for countless *in silico* experiments. With the only limit of available computing power, modellers can run simulations under numerous system set-ups, initial conditions and scenarios of one or multiple stressors applied in combination, in varying frequency and intensity over potentially very long periods. They can also test wide ranges of variation within the parameter space of the model, a feat that would be too expensive and impractical with real-life experiments, not to say impossible. By analysing model re-

sults, they may identify the factors and parameters that effectively drive the response of a system (those that the model output is most sensitive to) and through bifurcation analysis, for instance, they might also discover the existence of multiple equilibria, some of which could be qualitatively different alternative stable states.

An example of such a case is that of shallow lakes, which have become the archetype of alternative stable state theory in freshwater ecology. Classic shallow lake theory explains how these ecosystems may be found in either a macrophyte-dominated, clear water state, or a phytoplankton-dominated, turbid water state (Scheffer *et al.*, 1993). In each of these states, the aforementioned primary producers stabilise the system through positive feedback loops. Rooted macrophytes help to keep the system in a clear water state by reducing sediment resuspension and nutrient availability for phytoplankton. Phytoplankton reinforces the turbid water state by shading rooted macrophytes and hampering their development. For low nutrient concentrations, only the macrophyte-dominated state is possible, whereas phytoplankton invariably dominates under eutrophic and hypereutrophic conditions.

In between these two extremes, *i.e.*, within a range of intermediate nutrient concentrations, both states may be possible, but due to the aforementioned stabilising mechanisms, a perturbation is required for a shift from one stable state to the other to occur. For instance, the artificial removal of a significant amount of macrophytes may force a shift from the clear to the turbid water state (Kuiper *et al.*, 2017). But because macrophyte-dominated, clear water shallow lakes support a larger set of ecosystem services (Janssen *et al.*, 2020), shallow lake ecologists and managers mainly aim to achieve the opposite. Consequently, a great focus has been given over the years to finding effective strategies to restore the clear water state in lakes that have shifted to phytoplankton dominance as an undesirable consequence of eutrophication worldwide (Jeppesen *et al.*, 2007; Søndergaard *et al.*, 2017; López Moreira, *et al.*, 2018).

Due to the implications for policy and management, part of this effort has been put into determining the critical nutrient loading of different types of shallow lakes (Janse *et al.*, 2010), for which complex shallow lake ecosystem models have been developed, as is the case of PCLake (Janse, 1997). In time, multiple studies carried out using this model have proven the usefulness of highly complex ecosystem models to provide answers to very relevant questions, such as how future environmental conditions, *e.g.*, climate change, might affect shallow lake resilience and modify critical nutrient loads (Mooij *et al.*, 2007). The remarkable success of PCLake has motivated further developments of the model, like the addition of previously absent ecosystem components (Hölker *et al.*, 2015) and its expansion into PCLake+ (Janssen *et al.*, 2019), which

allows for the simulation of deep and stratified lakes. This ever-increasing degree of complexity that keeps improving PCLake's capacity to simulate real-world lakes makes it extremely useful for management-oriented applications. Yet complexity always comes at a cost, and that is the difficulty to disentangle the effects of a large number of factors on a large number of elements and processes. For this reason, minimal models are still necessary when basic knowledge about a particular ecological interaction is yet to be acquired or, for instance, to link species- to ecosystem-level responses (Mooij *et al.*, 2009).

In between these two extremes of the complexity spectrum, intermediate or reduced complexity models may be developed in a custom-tailored fashion, to answer very specific, novel questions for which a knowledge basis is still developing and has therefore yet to be incorporated into well-established, state-of-the-art complex models like PCLake+, MyLake, ALBM, the coupled GLM-AED2 and GOTM-WET, the Delft3D suite, among many others, extensive lists of which can be found in modelling literature (Janssen *et al.*, 2015; Mooij *et al.*, 2010). These novel questions could relate, for instance, to the potentially synergistic, non-linear effects of two or more combined stressors such as climate change (Jackson *et al.*, 2016), artificial light at night (Gaston *et al.*, 2015), browning (Senar *et al.*, 2021) and agricultural pollution (Havens *et al.*, 2011). The aim of a reduced complexity model could also be to simulate the response of a very particular setup that does not necessarily fit the hypotheses and assumptions under which ready-to-use models have been developed. A very good example of these exceptional cases is that of mesocosms. These increasingly popular experimental tools help ecologists investigate foreseeable future ecological dynamics, serving as a bridge between simple, laboratory microcosms, and the highly complex natural systems in which they are ultimately interested (Stewart *et al.*, 2013).

A challenge arises though when developing these models of intermediate complexity. That is, to include all (but only) the elements and processes that are strictly necessary for observed dynamics to be captured by the model, but without feeding it too many unverified hypotheses or assumptions about these elements and processes. This leads to a high degree of uncertainty about model results, even when they match the observations (which are usually based on a very limited, unique dataset).

In this paper, we present our views on what aspects need to be considered to hit the sweet spot of complexity. We also argue that despite sensible arguments in favour of increasing model complexity as much as possible (Hellweger, 2017) and against the development of new models largely following the concept of others, so as to not "*reinvent the wheel*" (Mooij *et al.*, 2010), enough reasons remain to encourage the continued development of

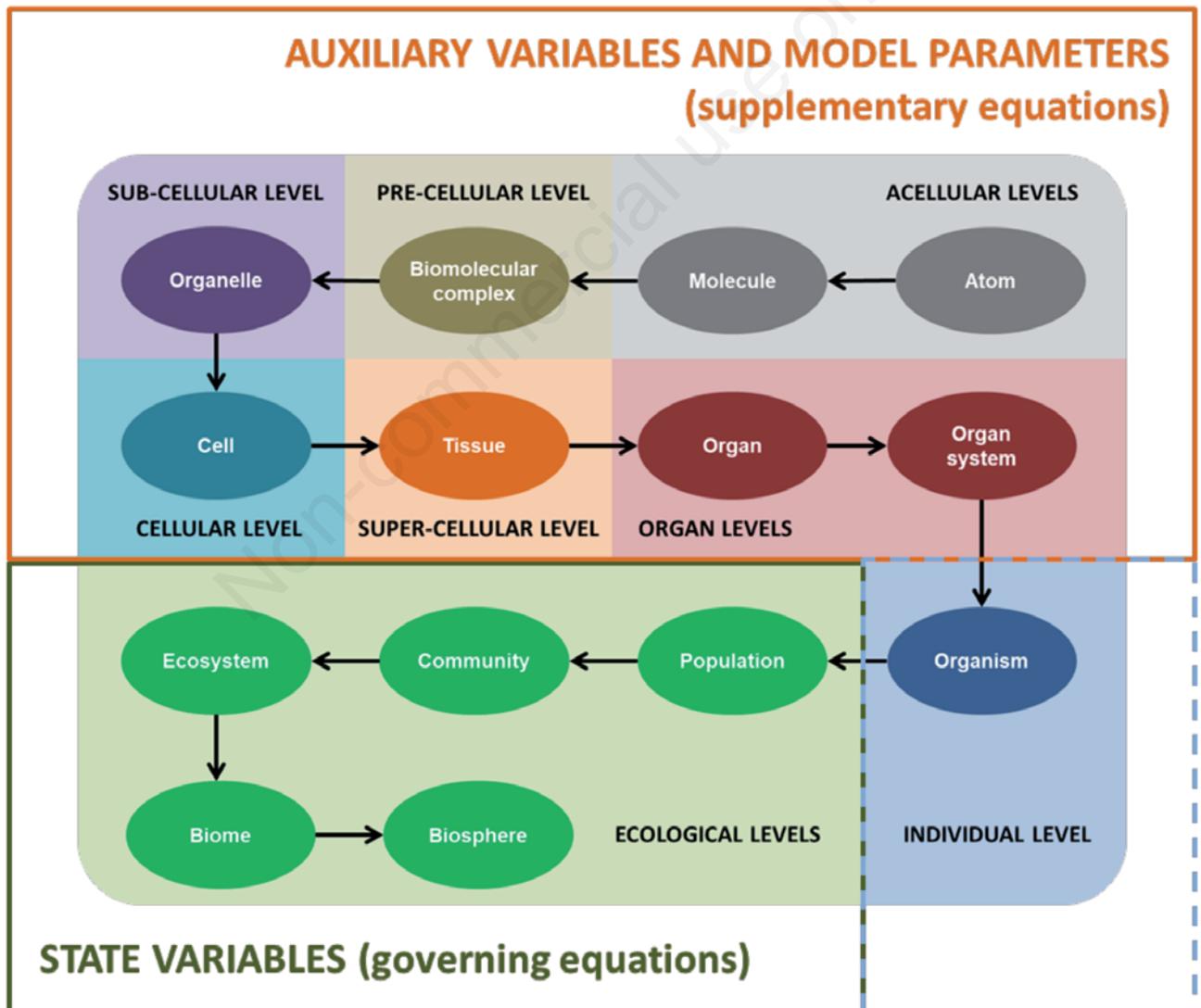
custom-tailored models of varying degrees of complexity in the field of freshwater ecology and other aquatic sciences. Finally, we present our vision on why this particular area is a fertile field for young researchers to gain both the technical and scientific knowledge required to become model developers, which would facilitate approaches such as ensemble modelling, interdisciplinary research and teamwork, as well as science communication.

### A HIERARCHY OF COMPLEXITY FOR A HIERARCHY OF BIOLOGICAL ORGANISATION

A first thing to keep in mind when developing a new model is that there is a nested hierarchy to the organisation of life (Fig. 1). Everything we observe at the ecosystem

scale is the macroscopic expression of processes happening at a series of progressively smaller ones, all the way down to the tiniest particles that constitute all matter. For example, the thermal (random) motion of atoms and molecules is responsible for the molecular diffusion process, through which the concentration of substances (*e.g.*, nutrients) in a stagnant fluid or laminar flows tends to homogenise over time (Fick's second law of diffusion). Still, in an aquatic ecosystem model, it would be a fool's errand to attempt to model the motion of every single atom or molecule. In practice, this random motion of particles is parameterised in the form of a background diffusion coefficient, as part of a much larger eddy diffusion coefficient that also accounts for turbulent motion.

Biological processes, however, do not obey such laws,



**Fig. 1.** Hierarchy of biological organisation and the usual mathematical structure of process-based models focusing on the ecological levels. The individual (organism) level may be either parameterised or modelled explicitly (as in individual-based population and community models).

and hence cannot hope to find a universal description in the form of simple mathematical expressions. Even our most cumbersome equations fall short of the task, as the realm of the living is characterised by randomness, variation, adaptation and evolution through natural selection. But as challenging as it may be to reproduce the complexity of life on a computer, we can always think of strategies to approximate our models to the real world.

When developing equations for processes occurring at the higher levels, a first step could be to look for inspiration in the equations that describe well-studied processes happening at the lower ones. A good example of such an approach would be that of the widely used Monod equation:

$$\mu = \mu_{max} \frac{S}{K_s + S} \quad (\text{eq. 1})$$

which describes the growth rate of microorganisms at the population level ( $\mu$ ) as a function of its maximum value ( $\mu_{max}$ ), the concentration of the limiting substrate for growth ( $S$ ), and a half-velocity constant ( $K_s$ ) (Monod, 1942). This empirical equation is the mathematical equivalent of a theoretical one, the Michaelis-Menten equation, a model of enzyme kinetics at the biomolecular level (Johnson and Goody, 2011). In the former, the population growth rate replaces the biochemical reaction rate of the latter, with all analogous terms having the same mathematical meaning.

Over the years, as research progresses, it may lead to a more detailed (and complex), mechanistic description of a process. In the case of microorganism growth, this happened when what is commonly known today as the ‘Droop model’ was proposed (Droop, 1983). By introducing the concepts of cell quota and the uptake rate of a nutrient, differentiating between the extra- and the intracellular concentration of the limiting nutrient, this model could more accurately describe the growth of microorganisms, in particular under environmental nutrient depletion. This is because considering an internal nutrient ‘pool’ allows for the model to sustain further growth even after external nutrients are fully consumed, a more realistic biological mechanism that the Monod equation cannot possibly reproduce (Sommer, 1991).

One might argue that examples like these make a strong case in favour of updating the biological equations of our models with the newer and more complex ones (Hellweger, 2015). In current modelling practice, however, despite the increased accuracy of more recent mathematical formulations of important physiological processes, simplicity still wins more often than not. But why is the Monod equation still so widely used in modern aquatic ecosystem models even though it no longer reflects the current state of the art of biological science? A possible answer is that for most present-day model ap-

plications, updating this equation does not significantly improve the goodness of fit at the levels of biological organisation relevant to the research questions being asked. This relates to the fact that in most ecological models, the state variables and their governing equations are all scaled up to the higher levels of biological organisation, and the processes occurring below the population level are introduced in the form of model parameters and auxiliary variables described by supplementary equations (Fig. 1).

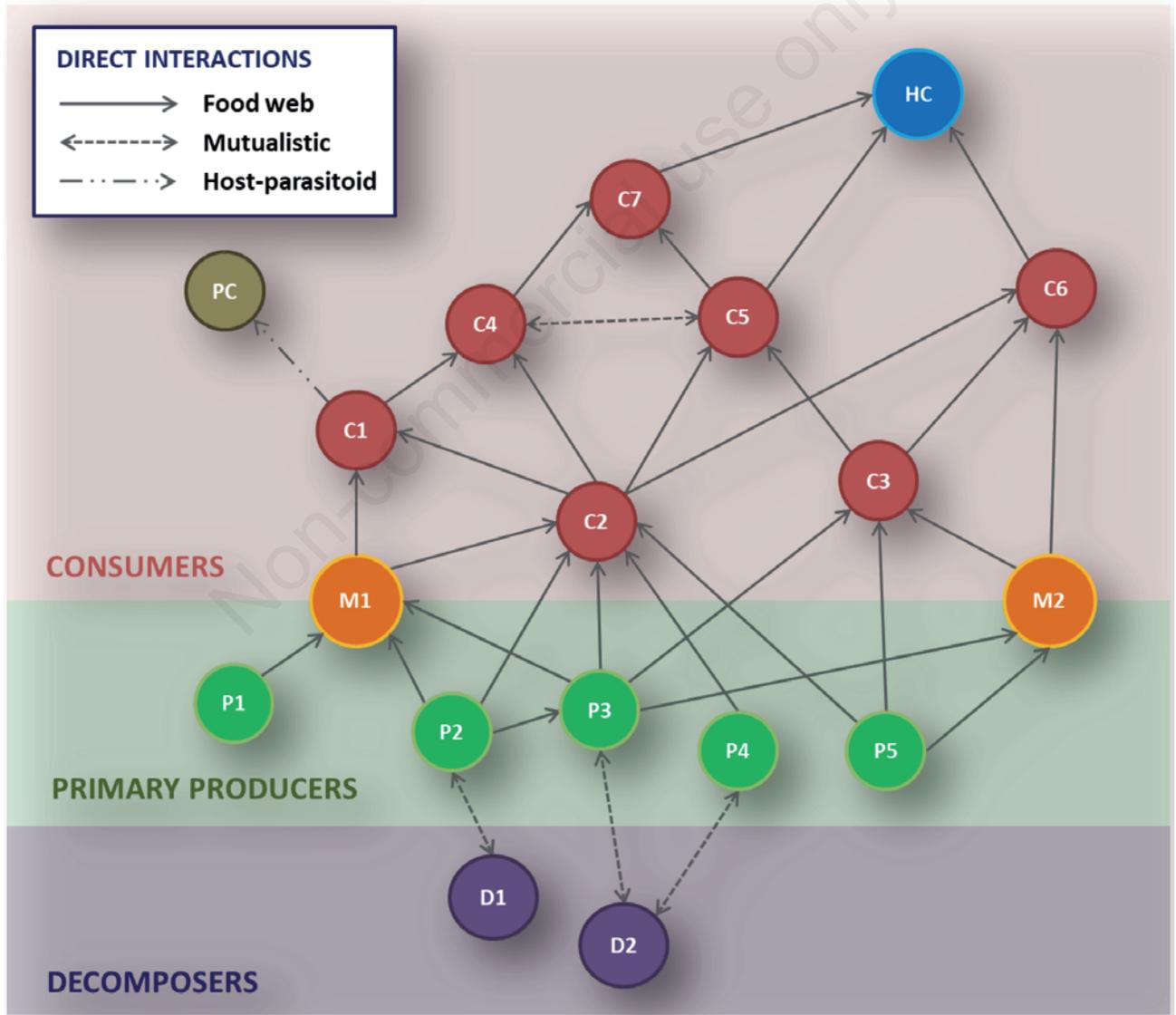
Even in the most complex lake models, the finest level of biological representation that is explicitly implemented is usually that of the ecological role (niche). In PCLake, for instance, planktivorous fish are implemented as a single element even though, among these, in reality, there are several species, each represented by numerous genetically and phenotypically different individuals that are either male or female, at a different stage of their life cycle and are ultimately the result of their particular life history. But because ecological roles are enough to address the type of questions PCLake was designed for (*i.e.*, those related to lake eutrophication), individual variation becomes unimportant and such a high level of detail is hence avoided through ‘bulk’ parameter values that enable the model to effectively reproduce the average dynamics of all relevant elements at the ecosystem level. This is, however, not always the case. When capturing individual differences between (groups of) members of the same species becomes necessary, for instance, to answer specific questions about their influence on emergent properties of the system at higher hierarchical levels, complex agent-based models are the appropriate choice (Van Nes *et al.*, 2005).

In between these two extremes is the case when the research question relates not to the variation among members of the same species, but the variation among different (groups of) species occupying the same niche. For example, when investigating the response of the phytoplankton assemblage (community adaptation) to sudden or gradual environmental change. To capture this, one would need to explicitly model at least different phytoplankton functional groups based on certain common characteristics (green algae, diatoms, cyanobacteria, etc.), or even the individual species belonging to each of these functional groups. This is the approach, for instance, of PROTECH (Reynolds *et al.*, 2001), a model where phytoplankton dynamics are not implemented ‘in bulk’, but simulated separately for each species. This is done based on a growing library of hydrodynamically and ecologically important traits (morphology, daily movement, nutrient requirements, vulnerability to grazing, nitrogen-fixation capacity, etc.) of more than a hundred individual genera and/or species, a feature that has allowed for many applications in recent years (Elliot, 2021).

Other examples of models that focus on interspecific interactions at the community level are ecological network models, which may be very complex, including different types of relationships between species in addition to those of classical food webs (*e.g.*, mutualism between decomposers and primary producers, or host-parasitoid relationships) (Fig. 2). These models can grow in complexity to include dozens of functional nodes (D'Alelio *et al.*, 2016), or be simplified interpretations of the food web (Wollrab *et al.*, 2012), depending on what is relevant to the research question. In this respect, they can also focus exclusively on the interaction between two species, being as simple as the Lotka-Volterra predator-prey equations

$$\begin{cases} \frac{dx}{dt} = (k_1 - k_2 y)x \\ \frac{dy}{dt} = (k_3 x - k_4)y \end{cases} \quad (\text{eq. 2})$$

where  $x$  and  $y$  are the numbers of prey and predators, respectively,  $t$  is time and the  $k_i$  are positive real coefficients. In the case of the planktonic community of an aquatic ecosystem, these models could very well describe the interactions between zooplankton and phytoplankton. By adding nutrient dynamics (an extra layer of complexity), one has what is known as an NPZ model (nutrients N, phytoplankton P, and zooplankton Z), the most basic rep-



**Fig. 2.** An ecological network model including three different types of direct interspecific interactions among the decomposers ( $D_i$ ), primary producers/autotrophs ( $P_i$ ), mixotrophs ( $M_i$ ) and consumers/heterotrophs ( $C_i$ ) of an ecosystem, including a parasitic consumer (PC) and the human consumer (HC) as top predator of the food web.

resentation of a pelagic ecosystem (Franks, 2002). The latter can be further extended to include detritus  $D$ , *i.e.*, an NPZD model (Edwards, 2001; Duquesne *et al.*, 2021). Hereafter, we present an example of such a model, simplified from those in Edwards (2001):

$$\begin{cases} \frac{dN}{dt} = -\frac{aN}{b+N}P + \frac{\beta\lambda P^2}{\mu^2 + P^2}Z + \phi D \\ \frac{dP}{dt} = \left(\frac{aN}{b+N} - r - s\right)P - \frac{\lambda P^2}{\mu^2 + P^2}Z \\ \frac{dZ}{dt} = \left(\frac{\alpha\lambda P^2}{\mu^2 + P^2} - \gamma\right)Z \\ \frac{dD}{dt} = rP + (1 - \alpha - \beta)\frac{\lambda P^2}{\mu^2 + P^2}Z - (\phi + \psi)D \end{cases}$$

(eq. 4) (eq. 5) (eq. 6) (eq. 7)

where the growth of phytoplankton ( $P$ ) is limited by the environmental concentration of nutrients ( $N$ ), formulated as in the Monod equation (eq. 1) with maximum fractional growth rate  $a$  and uptake half-saturation constant  $b$ ; and phytoplankton losses are due to respiration, sinking (at fractional rates  $r$  and  $s$ , respectively) and grazing by zooplankton ( $Z$ ) following a Holling Type III function (Holling, 1959), where  $\lambda$  is the maximum fractional grazing rate and  $\mu$  is the concentration of phytoplankton for which the effective fractional grazing rate is half of the maximum (*i.e.*,  $0.5\lambda$ ). In this model, the zooplankton population grows as it is fuelled by grazing at a fractional rate  $\alpha$  and decreases due to mortality at a fractional rate  $\gamma$ , with dead biomass sinking instantly out of the pelagic zone. Nutrients are replenished through remineralisation of detritus ( $D$ ), at fractional rate  $\phi$ , and from zooplankton excrements, a part of which is instantly remineralised at fractional rate  $\beta$ . The concentration of detritus increases as a result of phytoplankton respiration, zooplankton excretion (faecal pellets generated at fractional rate  $1 - \alpha - \beta$ ); and decreases due to remineralisation and sinking, the latter of which occurs at fractional rate  $\psi$ . Despite the higher complexity of eqs. 4-7, we recognise therein the same basic structure of eqs. 2-3, which may already give rise to nonlinear dynamics.

By expanding these simple ecosystem models beyond the pelagic, integrating them with other simple models that describe the dynamics of the remaining compartments, they can evolve into much more complex models like PCLake/PCLake+ (Janse, 1997; Janssen *et al.*, 2019). This is to say that, at every level of biological organisation, a system can be modelled following different approaches and at various degrees of complexity (Tab. 1) by progressively building upon the simple equations that describe its single elements, integrating them in a way that emergent properties of higher levels are also captured by the model.

## DIFFERENT MODELS FOR DIFFERENT TEMPORAL AND SPATIAL SCALES

Research has shown that models that use genetically fixed traits may not provide reasonable projections for community adaptation, because organisms can adapt to modest changes in environmental factors such as temperature or, especially in the case of photosynthetic organisms, irradiance. For phytoplankton, this can even happen over the time scale of decades (Irwin *et al.*, 2015). Hence, running a phytoplankton model over long periods under a moderately increasing temperature and/or decreasing light availability scenario, with a focus on community composition, would warrant the implementation of community acclimation and adaptation mechanisms acting at the species level, *i.e.*, resulting from genotypic and phenotypic variation, neither of which is normally considered in ready-to-use models (Anderson, 2005). For this reason, modelling phytoplankton functional groups might be strictly necessary but not necessarily sufficient to obtain an accurate prediction of community composition changes over the long term. In any case, because robust group-specific parameterisations that can support predictions in a wide array of spatiotemporal domains are still unavailable, complexity should only be increased gradually (Shimoda and Arhonditsis, 2016).

Also in this respect, eco-evolutionary dynamics have been recently cited as one of the challenges at the next generation of aquatic ecosystem models will need to tackle (Mooij *et al.*, 2019). This is because evolution and ecological adaptation mechanisms, *e.g.*, behavioural changes in response to stress, may happen at overlapping time scales, and the natural selection of some ecologically important phenotypes can happen fast enough to impact the outcome of ecological interactions ( Hairston *et al.*, 2005). Conversely, the integration of slow processes may also be needed when running simulations over long periods. For example, in coupled human-freshwater systems, in addition to human-induced climate change, cultural change and technological innovation happening at the time scale of decades may ultimately drive ecosystem trajectories in the long term (Ward *et al.*, 2019).

But in nature, adaptation happens not only over time but also across geographical gradients such as those related to climate. In the case of phytoplankton, for instance, cell size and shape are key traits under selection by the environment that link to phenotypical differences at the species level. To avoid the difficulties of modelling single individuals in models that would only require focusing on the ecosystem level, trait-based approaches are a suitable strategy to reduce complexity while retaining realism (Litchman *et al.*, 2013). Among these, spatially resolved aggregated trait-based models with adaptive traits (Peeters and Straille, 2018) could eventually enable modelling life

## Hitting the sweet spot of complexity

**Tab. 1.** The study of biological organisation levels in an aquatic context and examples of modelling approaches that are suitable to provide answers to questions relating to each level.

Level of biological organisation	Example/s of discipline/s	Example/s of study object/s	Model example/s	Reference/s
Atom	Atomic physics	Random walk of a single atom	A mechanical model of Brownian motion	Dürr <i>et al.</i> , 1981
Molecule	Molecular physics	Diffusion of oil molecules in water	A numerical oil spill transport model based on an fBm (fractional Brownian motion) particle tracking algorithm	Guo <i>et al.</i> , 2009
Biomolecular complex	Molecular biology	An enzyme-catalysed chemical reaction Denitrification in lake sediments	Michaelis-Menten equation of enzyme kinetics A nitrate-limited denitrification model based on Michaelis-Menten kinetics and the Arrhenius temperature dependence equation	Johnson <i>et al.</i> , 2011 Messer and Brezonik, 1983
Organelle	Plant cell biology	Photosynthesis within the chloroplast	Photosynthesis model at the chloroplast level	Bernacchi <i>et al.</i> , 2013
Cell	Microbiology	Microbial replication	Cell quota and nutrient uptake rate (Droop model)	Droop, 1983
Tissue	Plant histology	Growth of macrophyte surfaces	SAGA 1 model for the growth of leaves, stems and roots of the macrophyte <i>Potamogeton pectinatus</i> L. (sago pondweed)	Hootsmans, 1994
Organ	Fish anatomy and physiology	The liver of a fish	A model of the fatty acid content of the liver of a fish (regression model)	Turchini <i>et al.</i> , 2006
Organ system	Fish anatomy and physiology	The buoyancy regulation of <i>Gadus morhua</i> (Atlantic cod)	A bioenergetic buoyancy regulation model of fish species that have a swim bladder	Strand <i>et al.</i> , 2005
Organism	Reptile anatomy and physiology Invertebrate anatomy and physiology	A specimen of <i>Chelonia mydas</i> (green turtle) A specimen of <i>Chironomus riparius</i> (harlequin fly)	Somatic growth model (statistical model) Energy-based life cycle model	Bjorndal <i>et al.</i> , 2000 Péry <i>et al.</i> , 2005
Population	Population ecology	The population of a microorganism The population of <i>Rutilus rutilus</i> (common roach)	Monod equation as part of a logistic growth model A spatiotemporal individual-based fish model	Monod, 1942 Hölker and Breckling, 2005
Community	Community ecology	The food web of a given lake (or part of it)	Lotka-Volterra equations Food web models Ecological network models PROTECH, a phytoplankton community model Individual-based models Trait-based models	Liu and Chen, 2003 Wollrab <i>et al.</i> , 2012 D'Alelio <i>et al.</i> , 2016 Reynolds <i>et al.</i> , 2001; Elliot, 2021 Van Nes <i>et al.</i> , 2002 Litchman <i>et al.</i> , 2013
Ecosystem	Ecosystem ecology	A specific lake and its immediate surroundings	NPZ model NPZD model Vollenweider input-output eutrophication model PCLakePCLake+	Franks, 2002 Duchesne <i>et al.</i> , 2021 Vollenweider, 1968 Janse, 1997 Janssen <i>et al.</i> , 2019
Biome	Biome ecology	The marine biome	The Atlantis modelling framework for the marine biome	Fulton <i>et al.</i> , 2011
Biosphere	Earth Systems Science (ESS)	Earth's ecosphere	Madingley model, a Global Ecosystem Model (GEM)	Harfoot <i>et al.</i> , 2014

at the highest levels of organisations (the biome and the biosphere); an ambitious goal that some have already argued is time to pursue (Purves *et al.*, 2013).

---

### ADJUSTING COMPLEXITY LEVEL TO CURRENT KNOWLEDGE THROUGH MODEL CALIBRATION, VALIDATION AND UNCERTAINTY ANALYSIS

In practice, increasing model complexity does not necessarily improve model performance (Arhonditsis and Brett, 2004). For this to be the case, a good mechanistic understanding of the system and the problem under investigation is a prerequisite. One important reason why simplicity is so appealing in process-based modelling is precisely that results are more easily understood in terms of both external forcing and internal mechanisms. Because the equations of simpler models involve fewer unconstrained parameters, they may also be more easily subjected to sensitivity analyses (Arhonditsis and Brett, 2004) or even allow for, *e.g.*, bifurcation analysis (as in Mooij *et al.*, 2009). Moreover, the more numerous parameters of the more complex models require substantially bigger calibration efforts and pose a higher risk of overfitting, which decreases the predictive skill of the model, *i.e.*, its applicability beyond the calibration dataset. For this reason, unless extensive datasets are available to support model optimisation, complexity should only be increased gradually and be accompanied by objective assessments of model parameterisations (Flynn, 2005). In addition, the more complex models should always demonstrate a better performance than the simpler ones they intend to supersede (Anderson, 2005).

As a rule of thumb, if a model that includes additional elements and processes does not lead to a better fit between observed and simulated values, a model with fewer elements and processes should be preferred (a parsimony principle in modelling akin to ‘Occam’s razor’ in problem-solving philosophy). Because the latter would normally involve fewer parameters (and assumptions regarding their values), the risk of overfitting would be minimised. But because simplifications of reality also rely on a series of hypotheses regarding the importance (or lack thereof) of system elements and processes, the application of such a parsimony principle is not as straightforward in modelling (Hellweger, 2017). In some cases, an exception should be made in favour of higher complexity even when the more complex model does not perform any better (or performs even worse). One such case is when there is sufficient knowledge about the system to support the inclusion of additional elements and/or processes because they are known to be of critical importance.

An example of this is provided by the inclusion of

tube-dwelling invertebrates (chironomids) into the PCLake model structure (Hölker *et al.*, 2015). This study highlighted the important water filtration capacity of these benthic organisms and their role as stabilisers of the clear water state in shallow lakes. In the former, simpler version of PCLake, the filtration rate of chironomids was incorporated into that of pelagic zooplankton, which had likely been overestimated in previous applications of the model (Janse *et al.*, 2010). Studies like this clearly show other aspects that require attention when designing and implementing a new model (*e.g.*, the need for lower-level process validation), as well as when adding new elements or processes to an existing model, which is the need for recalibration. Without this, uncertainty about chosen parameter values can be very high, and so the applicability of the updated model to other case studies cannot be guaranteed until it has been extensively re-validated.

In this respect, a good practice that should be encouraged is not only to calibrate, but also to validate a model alongside all of its sub-models, *i.e.*, at all affected levels of biological organisation (as in Mintram *et al.*, 2020). With this aim, a framework for the hierarchical assessment of aquatic ecosystem models has been proposed very recently (Hipsey *et al.*, 2020). Referred to as CSPS (Concept, State, Process, System), it aims to guide future validation efforts at all levels of model design and implementation, starting from the conceptualisation stage (level 0). At this level, an *a priori* assessment is conducted of the applicability of the model to the specific case study and its suitability to answer the research questions over the range of conditions to be simulated. Additionally, sub-models are evaluated against the state of the art of relevant ecological theory. This is followed by the *a posteriori* assessment of how well available datasets compare with simulated values of state variables or their derived metrics (level 1), transfer functions, *i.e.*, mass and energy fluxes, and process rates (level 2). Finally, an evaluation is carried out of whether the emergent properties of the model at the ecosystem level (level 3) that are not necessarily predictable from its formulation, match those that have been observed in or theorised for the real system (*e.g.*, if it can mimic ecological succession or shifts between alternative stable states). The process is iterative, as failure to validate the model at any of the higher levels can lead to the reconsideration of the model at lower ones.

Within the CSPS framework, validation at levels 1 and 2 is largely based on an array of widely used metrics as well as less commonly applied techniques. Examples of these metrics are model bias, the Mean Absolute Error (MAE), the Root Mean Square Error (RMSE), the Nash-Sutcliffe Efficiency (NSE) and Spearman’s rank correlation coefficient. Additional strategies covered by the CSPS framework include the comparison of exceedance probabilities, data transformation methods such as the

Fast Fourier Transform (FFT) and wavelet analysis and the application of distribution, cross-correlation and autocorrelation functions. While quantitative approaches may always be supplemented by qualitative and semi-quantitative techniques like visual inspection of time series and contour plots, and the construction of spatial maps and Taylor diagrams, these become the main source of validation at level 3. At this level, judgement based on expert knowledge is fundamental considering that results will always be affected by a certain degree of uncertainty that is rarely quantified.

Although quantifying uncertainty is not a trivial task, it is a requirement to make models more useful and appealing in environmental management (Schuwirth *et al.*, 2019). In light of this, Bayesian inference methods have been getting more attention in recent years. These allow not only for a data-driven statistical calibration of model parameters, which is much more robust than traditional manual calibration (trial and error) but also for uncertainty analysis based on the posterior probability distributions of model outputs. They may also help to select between models of different complexity through well-established criteria like the Deviance Information Criterion (DIC), a generalisation of the Akaike Information Criterion (AIC) used in statistical model selection. Akin to the AIC, the DIC provides a measure of predictive error, penalising the goodness of fit of a model by its degree of complexity (determined by the number of parameters). Lastly, when several models exist, none of which can be rejected because they are equally good for the intended purpose, Bayesian Model Averaging (BMA) offers a means to assess uncertainty in model selection and facilitate ensemble modelling approaches (Hoeting *et al.*, 1999).

---

## NEW MODELS FOR NEW PROBLEMS AND NOVEL QUESTIONS

In principle, more knowledge about a system enables for more complexity, and more complexity, when justified, arguably increases the predictive capacity of a model (Hellweger, 2017). However, fundamental knowledge is still lacking when it comes to relatively novel environmental stressors. For instance, artificial light at night (Gaston *et al.*, 2015) challenges the common observation that biological activity is regulated by diurnal and seasonal cycles, particularly in aquatic systems (Perkin *et al.*, 2011), and may require a deep rethinking of the fundamental assumptions embedded in different models, as it may act at different spatial and temporal scales and multiple levels of biological organisation (genotypes, individuals, populations, communities and ecosystems). Explicitly including novel stressors in a complex ecosystem model is, therefore, not straightforward. Parameterising effects that have only been observed at the highest

biological organisation levels without first gaining a better understanding of stressor impacts at the lower ones (*e.g.*, critical behavioural changes at the individual level, or the inhibition of ecologically relevant biochemical reactions at the biomolecular level) would compromise the potential to extend the model beyond the specific calibration domain. The development of models tailored to individual experiments aiming to improve this knowledge base would help to gradually increase complexity, laying the foundation for their future incorporation into the more complex ecosystem models.

Mesocosm experiments are another example where the development of custom-tailored models is especially warranted. They are unique systems increasingly being used in current ecosystem research (Stewart *et al.*, 2013) that in general do not satisfy the hypotheses of ready-to-use ecosystem models. Firstly, their very particular geometries may introduce artefacts that require special consideration and an explicit treatment within a model. Examples of these are the growth of periphyton on the walls of the mesocosms (Chen *et al.*, 1997), shielding from wind and differential shading related to the presence of these walls and other surrounding structures. Secondly, the reduced scale of mesocosms in comparison with natural systems may invalidate the standard calibration of readily available models, because large-scale processes are inhibited in these intermediate environments (*e.g.*, the role of seiches on vertical diffusion), while other processes may be induced (*e.g.*, convective motions near the lateral boundaries due to temperature gradients between the interior and the exterior of the mesocosm). In flow-through mesocosms, the experimental scale has also been shown to result in altered temperature patterns and hydraulic conductivity, as well as changes in water quality parameters (*e.g.*, altered oxygen dynamics related to lower turbulence resulting from the reduced fetch, with consequences for phosphorus dynamics) (Ahn and Mitsch, 2002).

In addition to geometry and scale, a third reason why custom-tailored models are more suitable to accompany mesocosms experiments is that they normally address questions for which scientific understanding is only starting to develop. An example of this is the combined effects of multiple stressors. These effects are very challenging to capture with present models because not only can they be additive but also synergistic or antagonistic (Radinger *et al.*, 2016; Jackson *et al.*, 2016). In other words, they may affect the elements of a system not only directly, but also through indirect, cascading effects that have likely not been deemed important or even thought of when designing these models. Other questions that current aquatic ecosystem models are generally not able to address are those related to the toxicity of both naturally occurring chemicals such as allelopathic substances, and those of human origin such as pesticides and other micro-pollu-

tants (hormones, antibiotics, microplastics, etc.) (Pal *et al.*, 2016; Fischer *et al.*, 2016).

Lastly, one needs to be aware that complex models might also fail in capturing potentially important feedbacks, especially the ones that biological and ecological processes may have on the hydro-thermodynamics of the water body. Examples of these feedbacks have been highlighted by several studies, such as the effect of changing colour (*e.g.*, due to browning) and turbidity of water on the thermal stratification of the water column (Persson and Jones, 2008; Rinke *et al.*, 2010), the analogous effect of developing macrophytes (Herb and Stefan, 2004) or even the role of planktonic events in determining the mixing regime of some lakes (Shatwell *et al.*, 2016). Recent efforts have addressed the need for two-way communication (full coupling) between biogeochemical and hydrodynamic models that would enable them to capture some of these well-known feedbacks. For instance, through modern frameworks such as FABM (Bruggeman and Bolding, 2014), water quality and ecosystem models (*e.g.*, AED or PCLake), can be linked to a hydrodynamic model (*e.g.*, GOTM or GLM) to simulate both spatial and temporal changes in the light extinction coefficient (Hu *et al.*, 2016; Hipsey *et al.*, 2019). Still, other less-known or still poorly understood feedbacks have yet to find an implementation on widely used models, like the influence of phytoplankton exopolymeric substances on the rheological properties of the aquatic medium (Jenkinson and Sun, 2011; Jenkinson *et al.*, 2015).

---

## TRAINING A NEW GENERATION OF MODELLERS

It is our impression that the development of new models has been somewhat discouraged in recent years. The risk of having too many models with the same features and with insufficient effort in optimisation has been presented as a challenge to overcome within the aquatic ecosystem modelling community (Mooij *et al.*, 2010). Although reasonable arguments are offered as to why this should be avoided (*e.g.*, that it would be more efficient to just use existing models rather than always creating new ones), we believe it is also true that training a new generation of model developers (and not only model users) is of great value, as there are no better ways to learn new skills than to “learn by doing” and especially “learn by failing” (Tawfik *et al.*, 2015). Supporting this belief is that the need to turn aquatic ecosystem modelling into a community effort has already been well established (Trolle *et al.*, 2012; Janssen *et al.*, 2015), possibly through frameworks like the one offered by the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP, [www.isimip.org](http://www.isimip.org)). In this respect, young modellers have proven to be particularly keen on networking, suc-

cessfully overcoming recent challenges through, *e.g.*, virtual summits (Meyer and Zwart, 2020). Other clear examples of this are groups such as the “Young Modellers in Ecology” (YoMos, [www.yomos.org](http://www.yomos.org)) and the “Aquatic Ecosystem MOdelling Network – Junior” (AEMON-J, <https://github.com/aemon-j>). Among the latter, ensemble modelling has been gaining popularity (Moore *et al.*, 2021). This approach has been deemed ‘superior’ to the use of a single model, not only because the mean of all models can be a better fit to observed dynamics, but because ensemble modelling allows for a more effective quantification and communication of uncertainty (Trolle *et al.*, 2014).

To facilitate this, however, the modelling community should worry about increasing its ranks. It has been argued that the number of aquatic ecosystem modellers is already too low at present and that they are even rapidly going extinct in some disciplines like water quality engineering (Hellweger, 2017). It is perhaps this scarcity of trained modellers relative to other freshwater disciplines one of the main reasons why biological and ecological experiments are usually designed without the input of a scientist with modelling expertise: there are only so many modellers out there that it is practically impossible to always involve at least one. This disconnection between biologists and modellers has been harshly criticised, as it leads to a lot of wasted effort (Flynn, 2005): so much data has been produced over many years that are of little value for modelling, without measuring critical parameters that would allow for the development and application of a model (*e.g.*, measuring only total chlorophyll without collecting any data about carbon, nitrogen and phosphorus biomass). To correct this imbalance, models should be well established as a standard tool to gain knowledge in biological and ecological research, alongside statistical and laboratory-based molecular methods.

To achieve this, young researchers need to feel that it makes sense to invest time and effort in acquiring the basic knowledge and building the necessary skill set to become a model developer. This includes a range of mathematical methods for the discretisation of differential equations, the implementation of numerical methods to approximate their solutions, assessing numerical stability, to perform the calibration of model parameters, carry out sensitivity and uncertainty analysis, not to mention the necessary computer programming skills. And because the most effective strategy to learn something new is by simply doing it (and failing at it), we believe there is no better way to foster the new generation of model developers than to encourage them to develop, implement and run their own models. This will require the active contribution of educators, who should be convinced that curricula (also in ecology programmes) can profit from introductory modelling courses. Stimulating the curiosity of students and providing them with the conceptual and practical

tools to use and possibly develop models will increase the general awareness that models cannot be used as black boxes. Furthermore, advanced courses, international summer schools, workshops and virtual communities will be needed to promote cultural growth and the spirit of collaboration among modellers, as well as to facilitate the crucial issue of maintaining accessible and usable versions of open-source modelling software (Frassl *et al.*, 2019; Jansse *et al.*, 2015).

The challenge is to turn all this effort into more than just an academic exercise, but into meaningful contributions to the scientific community. An important area for this might be the design of new experiments that address the novel questions we described in the previous section. Because of the very specific set-ups and needs of these experiments (which do not usually fit the hypotheses of ready-to-use models), they provide a fertile ground for the development of new, custom-tailored models of reduced complexity that can better link empirical and theoretical knowledge while tackled by young modellers (*e.g.*, GPLake by Chang *et al.*, 2019).

Having a thriving community of modellers would bring about several benefits beyond the obvious scientific ones. Modellers regularly work in teams and need to interact with scientists specialised in many different disciplines. In doing so, they progressively gain a generalist understanding of environmental systems and inadvertently become interpreters and translators, facilitating integrative research processes as they help overcome some of the major difficulties of interdisciplinary research (Kragt *et al.*, 2013). This role can only become more important in the future, as research is increasingly being conducted in interdisciplinary teams (Wuchty *et al.*, 2007, a fact that has been well acknowledged by recent doctoral programmes in freshwater sciences (*e.g.*, Serlet *et al.*, 2020). Moreover, because models are essential in communicating science, modellers are the natural interface between pure scientists, on the one hand, and both decision-makers and natural resource managers, on the other one. For this purpose, the new generation of modellers should also be trained to communicate with stakeholders effectively, an ability that – we feel – was not so commonly fostered in previous generations. Finally, and coming back to the importance of encouraging custom-tailored model development, we must add that having a suite of models of different size, complexity and scope can be very effective as a multi-model approach to engaging stakeholders and modellers in complex environmental problems and better address management needs (Fulton *et al.*, 2015).

---

## SUMMARY AND CONCLUSIONS

Any given process-based model is only one of many possible interpretations of a particular system, at the levels

of biological organisation that are relevant to the questions being addressed by the modeller. Although the design and implementation of a model can be realised at varying degrees of complexity, the level of detail should match existing knowledge of the phenomena under investigation. When there is sufficient knowledge about the system, a higher complexity level can lead to an improved predictive capacity of the model. Nevertheless, simplicity should be preferred when process understanding is lacking, as is the case in many areas of ecological research, such as those where the feedbacks of biology on hydrothermodynamics might be crucial, or those dealing with the combined effects of multiple stressors (especially novel ones), for which the knowledge base is only starting to develop. When deciding on the complexity level of a new model, some questions that might help guide this decision are the following.

- Does a more detailed description of a process lead to better model performance? If the answer is no, and especially if uncertainty is high about whether the process is indeed important in the real system, simplicity should be preferred. This reduces the risk of overfitting and allows for an easier mechanistic interpretation of model results.
- Do the research questions require lower levels of biological organisation to be incorporated into the model? For instance, if the applied stressors are expected to act on a biomolecular level, this level should be described in as much detail as possible, or if their effects are expected to be seen at the community level (*e.g.*, through adaptation), multiple taxonomic groups or even specific species may need to be modelled separately to be able to capture this reorganisation.
- Are simulations going to be run over long periods that required the consideration of eco-evolutionary dynamics? If so, to capture adaptation and evolution, adaptive parameters that change over simulation run time should be implemented (instead of fixing parameter values from start to end).
- Will the spatial domain of the model span a range of biogeographical regions? If such large spatial scales are to be considered, taxonomical and phenotypical adaptation to climate gradients is likely going to play an important role, and so trait-based approaches might be more suitable.

For all this, we argue that the continued development of new, custom-tailored models of reduced complexity is still warranted and should be encouraged, particularly so because they provide an excellent entry point for young modellers. These model development exercises will serve a double purpose. Firstly, they will provide a fertile ground for young modellers to gain basic knowledge and build their skill sets, allowing them to become an active part of the aquatic ecosystem modelling community. Sec-

only, they will provide meaningful contributions to many other scientists who already recognise the value of mathematical modelling as a tool for biological and ecological research, but whose experiments do not normally fit the hypotheses and requirements of readily available models. Increasing the number of active modellers would help reverse their scarcity relative to other scientists in the field of freshwater ecology, finally allowing for a more regular and systematic involvement of model developers in the early stages of experimental design. We hope to see the number of young model developers grow quickly, in particular, to promote continuing education and training in the form of international modelling courses and workshops; and enable the community-based framework that will push the state of the art forward through the development of new open-source models and the further development of existing ones, the construction and sharing of global data and code repositories, libraries of models, equations and parameter values, wider adoption of ensemble modelling approaches, *etc.*

---

## REFERENCES

- Ahn C, Mitsch W J, 2002. Scaling considerations of mesocosm wetlands in simulating large created freshwater marshes. *Ecol. Engin.* 18:327-342.
- Anderson TR, 2005. Plankton functional type modelling: running before we can walk? *J Plankton Res* 27:1073-1081.
- Arhonditsis GB, Brett TM, 2004. Evaluation of the current state of mechanistic aquatic biogeochemical modelling. *Mar. Ecol. Progr. Ser.* 271:13-26.
- Bernacchi CJ, Bagley JE, Serbin SP, Ruiz-Vera UM, Rosenthal DM, VanLoocke A, 2013. Modelling C<sub>3</sub> photosynthesis from the chloroplast to the ecosystem. *Plant Cell Environ.* 36:1641-1657.
- Bjorndal KA, Bolten AB, Chaloupka MY, 2000. Green turtle somatic growth model: Evidence for density dependence. *Ecol. Appl.* 10:269-282.
- Bruggeman J, Bolding K, 2014. A general framework for aquatic biogeochemical models. *Environ. Modell. Softw.* 61:249-265.
- Chang M, Teurlinx S, DeAngelis DL, Janse JH, Troost TA, Van Wijk D, Mooij WM, Janssen ABG, 2019. A Generically Parameterized model of Lake eutrophication (GPLake) that links field-, lab- and model-based knowledge. *Sci. Total Environ.* 695:133887.
- Chen C-C, Petersen JE, Kemp WM, 1997. Spatial and temporal scaling of periphyton growth on walls of estuarine mesocosms. *Mar. Ecol. Progr. Ser.* 155:1-15.
- D'Alleio D, Libralato S, Wyatt T, Ribera d'Alcalà M, 2016. Ecological-network models link diversity, structure and function in the plankton food-web. *Sci. Rep.* 6:21806.
- Droop MR, 1983. 25 years of algal growth kinetics: A personal view. *Bot. Mar.* 26:99-112.
- Duquesne F, Vallaey V, Vidaurre PJ, Hanert E, 2021. A coupled ecohydrodynamic model to predict algal blooms in Lake Titicaca. *Ecol. Model.* 440:109418.
- Dürr D, Goldstein S, Lebowitz JL, 1981. A mechanical model of Brownian motion. *Commun. Math. Phys.* 78:507-530.
- Edwards AM, 2001. Adding detritus to a nutrient-phytoplankton-zooplankton model: a dynamical-systems approach. *J. Plankton Res.* 23:389-413.
- Elliot JA, 2021. Modelling lake phytoplankton communities: recent applications of the PROTECH model. *Hydrobiologia* 848:209-217.
- Fischer EK, Paglialonga L, Czech E, Tamminga M, 2016. Microplastic pollution in lakes and lake shoreline sediments – A case study on Lake Bolsena and Lake Chiusi (central Italy). *Environ. Pollut.* 213:648-657.
- Flynn KJ, 2005. Castles built on sand: dysfunctionality in plankton models and the inadequacy of dialogue between biologists and modellers. *J. Plankton Res.* 27:1205-1210.
- Franks PJ, 2002. NPZ models of plankton dynamics: Their construction, coupling to physics, and application. *J. Oceanogr.* 58:379-387.
- Frassl MA, Abell JM, Botelho DA, Cinque K, Gibbes BR, Jöhnk KD, Muraoka K, Robson B, Wolski M, Xiao M, Hamilton D, 2019. A short review of contemporary developments in aquatic ecosystem modelling of lakes and reservoirs. *Environ. Model. Softw.* 117:181-187.
- Fulton EA, Boschetti F, Sporcic M, Jones T, Little R, Dambacher JM, Gray R, Scott R, Gorton R, 2015. A multi-model approach to engaging stakeholder and modellers in complex environmental problems. *Environ. Sci. Policy* 48:44-56.
- Fulton EA, Link JS, Kaplan IC, Savina-Rolland M, Johnson P, Ainsworth C, et al., 2011. Lessons in modelling and management of marine ecosystems: the Atlantis experience. *Fish Fisher.* 12:171-188.
- Gaston KJ, Visser ME, Hölker F, 2015. The biological impacts of artificial light at night: The research challenge. *Philos Trans R Soc Lond B Biol Sci* 370:20140133.
- Guo WJ, Wang YX, Xie MX, Cui YJ, 2009. Modeling oil spill trajectory in coastal waters based on fractional Brownian motion. *Mar. Pollut. Bull.* 58:1339-1346.
- Hairston NG, Ellner SP, Geber MA, Yoshida T, Fox JA. 2005. Rapid evolution and the convergence of ecological and evolutionary time. *Ecol. Lett.* 8:1114-1127.
- Harfoot MB, Newbold T, Tittensor DP, Emmott S, Hutton J, Lyutsarev V, et al., 2014. Emergent global patterns of ecosystem structure and function from a mechanistic general ecosystem model. *PLoS Biol.* 12:e1001841.
- Havens KE, 1995. Secondary nitrogen limitation in a subtropical lake impacted by non-point source agricultural pollution. *Environ. Pollut.* 89:241-246.
- Hellweger FL, 2015. 100 years since Streeter and Phelps: It is time to update the biology in our water quality models. *Environ. Sci. Technol.* 49: 6372-6373.
- Hellweger FL, 2017. 75 years since Monod: It is time to increase the complexity of our predictive ecosystem models (opinion). *Ecol. Model.* 346:77-87.
- Herb WR, Stefan HG, 2004. Temperature stratification and mixing dynamics in a shallow lake with submersed macrophytes. *Lake Reserv. Manage.* 20:296-308.
- Hipsey MR, Bruce LC, Boon C, Busch B, Carey CC, Hamilton DP, et al., 2019. A General Lake Model (GLM 3.0) for linking with high-frequency sensor data from the Global Lake Ecological Observatory Network (GLEON). *Geosci. Model Dev.* 12:473-523.

- Hipsey MR, Gal G, Arhonditsis GB, Carey CC, Elliott JA, Frassl MA, et al., 2020. A system of metrics for the assessment and improvement of aquatic ecosystem models. *Environ. Model. Softw.* 128:104697.
- Hoeting JA, Madigan D, Raftery EA, Volinsky CT, 1999. Bayesian model averaging: A tutorial. *Stat. Sci.* 14:382-401.
- Hölker F, Breckling B, 2005. A spatiotemporal individual-based fish model to investigate emergent properties at the organismal and the population level. *Ecol. Model.* 186:406-426.
- Hölker F, Vanni MJ, Kuiper JJ, Meile C, Grossart H-P, Stief P, et al., 2015. Tube-dwelling invertebrates: tiny ecosystem engineers have large effects in lake ecosystems. *Ecol. Monogr.* 85:333-351.
- Holling CS, 1959. Some characteristics of simple types of predation and parasitism. *Can. Entomol.* 91:385-398.
- Hootsmans M, 1994. A growth analysis model for *Potamogeton pectinatus* L., p. 250-286. In: W. Van Vierssen, M. Hootsmans and J. Vermaat (Eds.), *Lake Veluwe, a macrophyte-dominated system under eutrophication stress*. Dordrecht: Springer.
- Hu F, Bolding K, Bruggeman J, Jeppesen E, Flindt MR, van Gerben L, et al., 2016. FABM-PCLake – linking aquatic ecology with hydrodynamics. *Geosci. Model Dev.* 9:2271-2278.
- Irwin AJ, Finkel ZV, Müller-Karger FE, Troccoli Ghinaglia L, 2015. Phytoplankton adapt to changing ocean environments. *P. Natl. Acad. Sci. USA* 112:5762-5766.
- Jackson MC, Loewen CJ, Vinebrook RD, Chimimba CT, 2016. Net effects of multiple stressors in freshwater ecosystems: a meta-analysis. *Global Change Biol.* 22:180-189.
- Janse JH, 1997. A model of nutrient dynamics in shallow lakes in relation to multiple stable states, p. 1-18. In: L. Kufel, A. Prejs and J. I. Rybak (Eds.), *Shallow Lakes '95*. Dordrecht: Springer.
- Janse J, Scheffer M, Lijklema L, Van Liere L, Sloot J, Mooij W, 2010. Estimating the critical phosphorus loading of shallow lakes with the ecosystem model PCLake: Sensitivity, calibration and uncertainty. *Ecol. Model.* 221: 654-665.
- Janssen AB, Arhonditsis GB, Beusen A, Bolding K, Bruce L, Bruggeman J, et al., 2015. Exploring, exploiting and evolving diversity of aquatic ecosystem models: a community perspective. *Aquat. Ecol.* 49:513-548.
- Janssen AB, Hilt S, Kosten S, de Klein JJ, Paerl HW, Van de Waal DB, 2020. Shifting states, shifting services: Linking regime shifts to changes in ecosystem services of shallow lakes. *Freshwater Biol.* 66:1-12.
- Janssen AB, Teurlincx S, Beusen AH, Huijbregts MA, Rost J, Schipper AM, et al., 2019. PCLake+: A process-based ecological model to assess the trophic state of stratified and non-stratified freshwater lakes worldwide. *Ecol. Model.* 396:23-32.
- Jenkinson IR, Sun J, 2011. A model of pycnocline thickness modified by the rheological properties of phytoplankton expolymeric substances. *J. Plankton Res.* 33:373-383.
- Jenkinson IR, Sun J, Seuront L, 2015. Thalassorheology, organic matter and plankton: towards a more viscous approach in plankton ecology. *J. Plankton Res.* 37:1100-1109.
- Jeppesen E, Meerhoff M, Jacobsen BA, Hansen RS, Søndergaard M, Jensen J, et al., 2007. Eutrophication in Lakes: Restoration of shallow lakes by nutrient control and biomanipulation—the successful strategy varies with lake size and climate. *Hydrobiologia* 581:269-285.
- Johnson KA, Goody RS, 2011. The original Michaelis constant: Translation of the 1913 Michaelis–Menten paper. *Biochemistry* 50:8264-8269.
- Kragt ME, Robson BJ, Macleod CJ, 2013. Modellers' roles in structuring integrative research projects. *Environ. Model. Softw.* 39:322-330.
- Kuiper JJ, Verhofstad MJ, Louwers EL, Bakker ES, Brederveld RJ, van Gerven LP, et al., 2017. Mowing submerged macrophytes in shallow lakes with alternative stable states: Battling the good guys? *Environ. Manage.* 59:619-634.
- Litchman E, Ohman MD, Kjørboe T, 2013. Trait-based approaches to zooplankton communities. *J. Plankton Res.* 35:473-484.
- Liu X, Chen L, 2003. Complex dynamics of Holling type II Lotka–Volterra predator–prey system with impulsive perturbations on the predator. *Chaos Solitons Fract.* 16:311-320.
- López Moreira M.G.A, Hinegk L, Salvadore A, Zolezzi G, Hölker F, Monte Domecq RA, et al., 2018. Eutrophication, research and management history of the shallow Ypacaraí Lake (Paraguay). *Sustainability* 10:2426.
- Messer JJ, Brezonik PL, 1984. Laboratory evaluation of kinetic parameters for lake sediment denitrification models. *Ecol. Model.* 21:277-286.
- Meyer MF, Zwart JA, 2020. Virtual summit: Incorporating data science and open science in aquatic research. *Limnol. Oceanogr. Bull.* 29:144-146.
- Mintram KS, Maynard SK, Brown AR, Boyd R, Johnston AS, Thorbek P, Tyler CR, 2020. Applying a mechanistic model to predict interacting effects of chemical exposure and food availability on fish populations. *Aquat. Toxicol.* 224:105483.
- Monod J, 1942. [Recherches sur la croissance des cultures bactériennes]. [Book in French]. Hermann et cie, Paris.
- Mooij MW, van Wijk D, Beusen AH, Brederveld RJ, Chang M, Cobben MMP, et al., 2019. Modeling water quality in the Anthropocene: directions for the next-generation of aquatic ecosystem models. *Curr. Opin. Environ. Sustain.* 36:85-95.
- Mooij MW, De Senerpont Domis LN, Janse JH, 2009. Linking species- and ecosystem-level impacts of climate change in lakes with a complex and a minimal model. *Ecol. Model.* 220:3011-3020.
- Mooij MW, Janse JH, De Senerpont Domis LN, Hülsmann S, Ibelings BW, 2007. Predicting the effect of climate change on temperate shallow lakes with the ecosystem model PCLake. *Hydrobiologia* 584:443-454.
- Mooij MW, Trolle D, Jeppesen E, Arhonditsis G, Belolipetsky PV, Chitamwebwa DB, et al., 2010. Challenges and opportunities for integrating lake ecosystem models. *Aquat. Ecol.* 44:633-667.
- Moore TN, Mesman JP, Ladwig R, Feldbauer J, Olsson F, Pilla RM, et al., 2021. LakeEnsemblR: An R package that facilitates ensemble modelling of lakes. *Environ. Model. Softw.* 143:105101.
- Pal A, Yew-Hoong Gin K, Yu-ChenLin A, Reinhard M, 2010. Impacts of emerging organic contaminants on freshwater resources: Review of recent occurrences, sources, fate and effects. *Sci. Total Environ.* 408:6062-6069.
- Peeters F, Straile D, 2018. Trait selection and co-existence of phytoplankton in partially mixed systems: Trait based mod-

- elling and potential of an aggregated approach. *PLoS One* 13:e0194076.
- Perkin EK, Hölker F, Richardson JS, Sadler JP, Wolter C, Tockner K, 2011. The influence of artificial light on stream and riparian ecosystems: questions, challenges, and perspectives. *Ecosphere* 2:1-16.
- Persson I, Jones ID, 2008. The effect of water colour on lake hydrodynamics: a modelling study. *Freshwater Biol.* 53:2345-2355.
- Péry AR, Mons R, Garric J, 2005. Modelling of the life cycle of Chironomus species using an energy-based model. *Chemosphere* 59: 247-253.
- Purves D, Scharlemann JP, Harfoot M, Newbold T, Tittensor DP, Hutton J, Emmott S, 2013. Time to model all life on earth. *Nature* 493:295-297.
- Radinger J, Hölker F, Horký P, Slavík O, Dendoncker N, Wolter C, 2016. Synergistic and antagonistic interactions of future land use and climate change on river fish assemblages. *Global Change Biol.* 22:1505-1522.
- Reynolds CS, Irish AE, Elliot JA, 2001. The ecological basis for simulating phytoplankton responses to environmental change (PROTECH). *Ecol. Model.* 140:271-291.
- Rinke K, Yeates P, Rothhaupt K-O, 2010. A simulation study of the feedback of phytoplankton on thermal structure via light extinction. *Freshwater Biol.* 55:1674-1693.
- Schäfer RB, von der Ohe PC, Kühne R, Schüürmann G, Liess M, 2011. Occurrence and toxicity of 331 organic pollutants in large rivers of north Germany over a decade (1994 to 2004). *Environ. Sci. Technol.* 45:6167-6174.
- Scheffer M, Hosper SH, Meijer ML, Moss B, Jeppesen E, 1993. Alternative equilibria in shallow lakes. *Trends Ecol. Evol.* 8:275-279.
- Schuwirth N, Borgwardt F, Domisch S, Friedrichs M, Katzwinkel M, Kneis D, et al., 2019. How to make ecological models useful for environmental management. *Ecol. Model.* 411:108784.
- Senar OE, Creed IF, Trick CE, 2021. Lake browning may fuel phytoplankton biomass and trigger shifts in phytoplankton communities in temperate lakes. *Aquat. Sci.* 83:21.
- Serlet AJ, López Moreira M. GA, Zolezzi G, Wharton G, Hölker F, et al., 2020. SMART Research: Toward interdisciplinary river science in Europe. *Front. Environ. Sci.* 8:00063.
- Shatwell T, Adrian R, Kirillin G, 2016. Planktonic events may cause polymictic-dimictic regime shifts in temperate lakes. *Sci. Rep.* 6:24361.
- Shimoda Y, Arhonditsis GB, 2016. Phytoplankton functional type modelling: Running before we can walk? A critical evaluation of the current state of knowledge. *Ecol. Model.* 320:9-43.
- Sommer U, 1991. A comparison of the Droop and the Monod models of nutrient limited growth applied to natural populations of phytoplankton. *Funct. Ecol.* 5:535-544.
- Søndergaard M, Lauridsen TL, Johansson LS, Jeppesen E, 2017. repeated fish removal to restore lakes: Case study of Lake Væng, Denmark - Two biomanipulations during 30 years of monitoring. *Water* 9:43.
- Stewart RIA, Dossena M, Bohan DA, Jeppesen E, Kordas RL, Ledger ME, et al., 2013. Chapter Two - Mesocosm experiments as a tool for ecological climate-change research, p. 71-181. In: G. Woodward and EJ O'Green (eds.), *Advances in ecological research*. Academic Press.
- Strand E, Jørgensen C, Huse G, 2005. Modelling buoyancy regulation in fishes with swimbladders: bioenergetics and behaviour. *Ecol. Model.* 185:309-327.
- Tawfik AA, Rong H, Choi I, 2015. Failing to learn: towards a unified design approach for failure-based learning. *Educ. Technol. Res. Dev.* 63:975-994.
- Trolle D, Elliott JA, Mooij WM, Janse JH, Bolding K, Hamilton DP, Jeppesen E, 2014. Advancing projections of phytoplankton responses to climate change through ensemble modelling. *Environ. Model. Softw.* 61:371-379.
- Trolle D, Hamilton DP, Hipsey MR, Bolding K, Bruggeman J, Mooij WM, et al., 2012. A community-based framework for aquatic ecosystem models. *Hydrobiologia* 683:25-34.
- Turchini GM, Francis DS, De Silva SS, 2006. Modification of tissue fatty acid composition in Murray cod (*Maccullochella peelii peelii*, Mitchell) resulting from a shift from vegetable oil diets to a fish oil diet. *Aquacult. Res.* 37:570-585.
- Van Nes EH, Lammens EH, Scheffer M, 2002. PISCATOR, an individual-based model to analyze the dynamics of lake fish communities. *Ecol. Model.* 152:261-278.
- Vollenweider RA, 1968. Scientific fundamentals of the eutrophication of lakes and flowing waters, with particular reference to phosphorus and nitrogen as factors in eutrophication. OECD Technical Report DAS/CS1/68.27. Paris: Organisation for Economic Co-operation and Development.
- Ward NK, Fitchett L, Hart JA, Shu L, Stachelek J, Weng W, et al., 2019. Integrating fast and slow processes is essential for simulating human-freshwater interactions. *Ambio* 48:1169-1182.
- Wollrab S, Diehl S, De Roos AM, 2012. Simple rules describe bottom-up and top-down control in food webs with alternative energy pathways. *Ecol. Lett.* 15:935-946.
- Wuchty S, Jones BF, Uzzi B, 2007. The increasing dominance of teams in production of knowledge. *Science* 316:1036-1039.