Ocean biogeochemical modelling

Katja Fennel[1] ✉, *Jann Paul Mattern*[2], *Scott C. Doney*[3], *Laurent Bopp*[4], *Andrew M. Moore*[2], *Bin Wang*[1] *and Liuqian Yu*[5]

Abstract | Ocean biogeochemical models describe the ocean's circulation, physical properties, biogeochemical properties and their transformations using coupled differential equations. Numerically approximating these equations enables simulation of the dynamic evolution of the ocean state in realistic global or regional spatial domains, across time spans from years to centuries. This Primer explains the process of model construction and the main characteristics, advantages and drawbacks of different model types, from the simplest nutrient–phytoplankton–zooplankton–detritus model to the complex biogeochemical models used in Earth system modelling and climate prediction. Commonly used metrics for model-data comparison are described, alongside a discussion of how models can be informed by observations via parameter optimization or state estimation, the two main methods of data assimilation. Examples illustrate how these models are used for various practical applications, ranging from carbon accounting, ocean acidification, ocean deoxygenation and fisheries to observing system design. Access points are provided, enabling readers to engage in biogeochemical modelling through practical code examples and a comprehensive list of publicly available models and observational data sets. Recommendations are given for best practices in model archiving. Lastly, current limitations and anticipated future developments and challenges of the models are discussed.

[1]*Department of Oceanography, Dalhousie University, Halifax, Nova Scotia, Canada.*

[2]*Ocean Sciences Department, University of California Santa Cruz, Santa Cruz, CA, USA.*

[3]*Department of Environmental Sciences, University of Virginia, Charlottesville, VA, USA.*

[4]*Laboratoire de Météorologie Dynamique, Institut Pierre-Simon Laplace, CNRS, Ecole normale supérieure, Paris Sciences Lettres Université, Paris, France.*

[5]*Earth, Ocean and Atmospheric Sciences Thrust, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China.*

✉*e-mail: katja.fennel@dal.ca*

https://doi.org/10.1038/s43586-022-00154-2

Ocean biogeochemical models (OBMs) are spatially explicit models, consisting of a component that describes the ocean's temperature and salinity distributions and circulation — including wind and density-driven currents, and wind, convective and eddy-driven mixing — and a component that describes the transformations of biogeochemical constituents contained in seawater. The biogeochemical constituents are typically nutrients, functional plankton groups, non-living organic matter, dissolved gases and parameters of the inorganic carbon system (FIG. 1). Both components consist of numerical codes approximating systems of partial differential equations. OBMs can be regional or global in terms of their geographical scope. They can themselves be a component of a larger model, for example in Earth system models (ESMs), where an OBM is coupled to a model of the atmosphere and land biosphere, or self-contained where information from the atmosphere and land is imposed. An OBM is typically run forward in time, starting from a defined initial condition. The model simulates the evolution of its state variables subject to external forcing, such as wind, atmospheric variables — air temperature and partial pressure of carbon dioxide (pCO_2) — riverine nutrient and freshwater inputs, and solar radiation. OBMs can be run as hindcasts, describing past conditions; as nowcasts, aiming to describe the current state of the ocean; or as forecasts or projections, intended to inform about possible future ocean conditions.

OBMs emerged in the 1990s as a common tool to address the needs of two distinct communities with different scientific objectives. One community was interested in plankton ecology and sought to explain and predict seasonal phytoplankton dynamics with the help of marine plankton models. The roots of these models go back to Gordon Riley[1] with further developments in the 1980s and 1990s (REFS.[2–4]). When computers became more widely available, these models were coupled to three-dimensional models of ocean circulation. A regional model of the North Atlantic[5] was probably the first three-dimensional ocean circulation model with explicit representation of plankton dynamics. The other community was interested in the role of the ocean as a sink of anthropogenic carbon. Building on concepts established by Roger Revelle[6], early ocean carbon cycle models did not include an explicit representation of plankton[7–9]. In the seminal work by Ernst Maier-Reimer[10,11], models of carbon cycling and plankton dynamics were combined and integrated into global ocean circulation models. This type of model is now a widely used tool for ocean ecologists and biogeochemists and has evolved to include diverse functional plankton groups and multiple distinct elemental cycles.

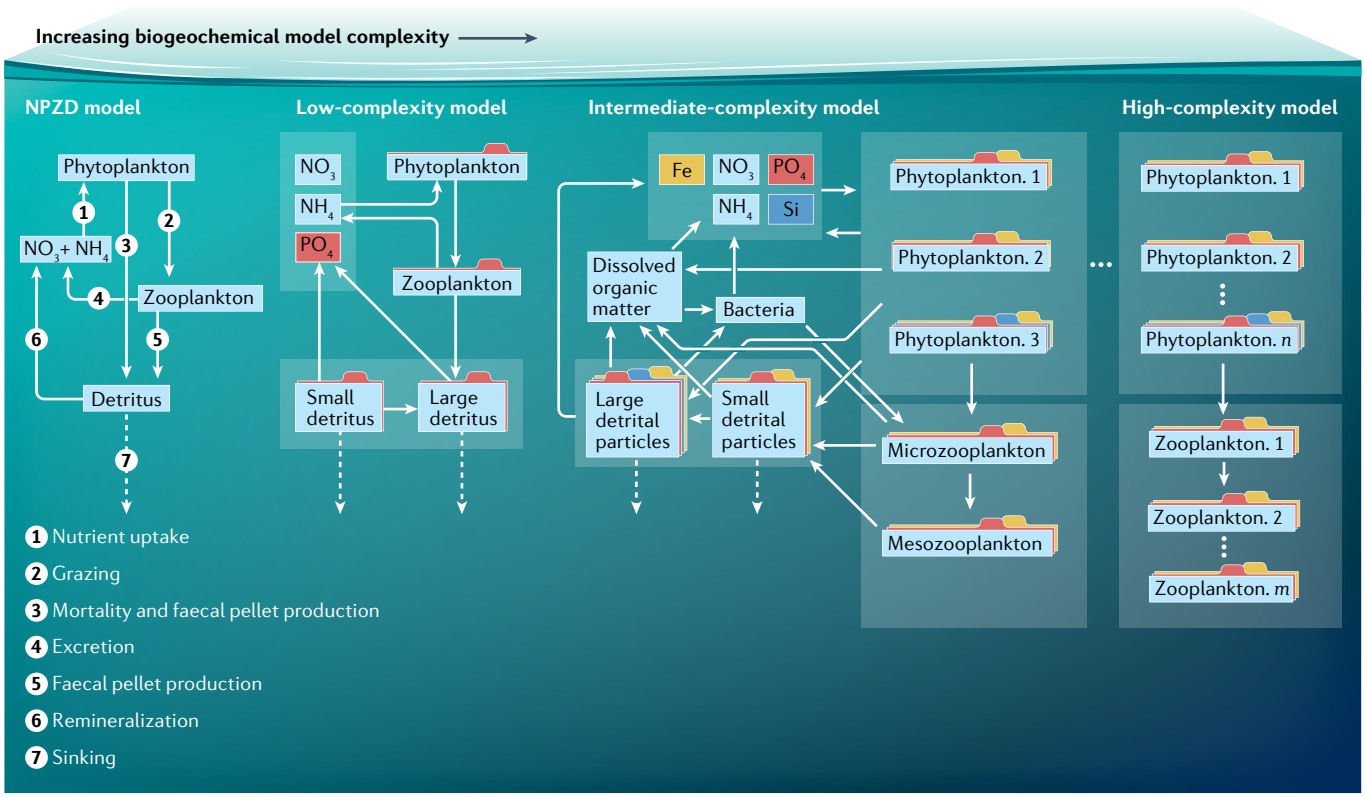


Fig. 1 | **State variables and biogeochemical transformations across a range of OBMs.** Ocean biogeochemical models (OBMs) increasing in complexity from left to right. The simplest, the nutrient–phytoplankton–zooplankton–detritus (NPZD) model, includes four state variables and one nutrient currency, often nitrogen. In the NPZD example, all transformations are indicated by labelled arrows. A typical low-complexity model includes several nutrients and nutrient currencies. Keeping track of multiple nutrient currencies requires multiple state variables for some functional groups and particulate pools (coloured tabs). In the low to high-complexity examples, only the general transformation direction is indicated for simplicity. In practice there are many more transformations between state variables, each represented by a parameterization that requires at least one, but typically more, biogeochemical model parameter. Chlorophyll is also omitted, although many models have a chlorophyll state variable for each phytoplankton group to account for photoacclimation. With greater complexity, there is an increase in plankton groups and organic matter pools with distinctions by particle size, dissolved organic matter and bacteria. For the high-complexity model, only functional phytoplankton and zooplankton groups are schematically represented because the inorganic and non-living organic pools are similar to the intermediate-complexity model. Squares represent state variables, solid arrows represent selected transformations between state variables and dashed arrows illustrate vertical sinking of particles.

Functional plankton groups
Groups of planktonic organisms that share similar traits, for example size, biogeochemical function or elemental requirements. These groups are defined to simplify the diversity of planktonic communities while capturing their essential biogeochemical functions in ocean biogeochemical models.

Initial condition
The complete set of state variables at one instant in time. Model integration starts from an initial condition.

State variables
A set of variables that fully characterize a model's dynamical state such that its future behaviour can be calculated, provided any external inputs are known. Each variable that belongs to this set is a state variable.

This Primer describes the process of OBM construction; reviews and illustrates methods and metrics for evaluating models against observations; and introduces approaches for combining models and observations. The latter methods are collectively referred to as data assimilation and include optimization of the model parameters and model state. Model parameters include plankton growth and grazing rates, rates of organic matter sinking and remineralization. The model state involves derivation of the most likely ocean state, given a mechanistic understanding of the system from model equations and available observations. Several important applications of OBMs are described to illustrate their breadth and utility. Additionally, best practices for archiving model codes, outputs and conducting intercomparisons are recommended. The Primer concludes by discussing the current limitations of OBMs and their applicability, plus anticipated new developments and challenges.

Experimentation

Model construction

Biogeochemical equations. The biogeochemical dynamics at the core of an OBM are commonly cast as a system of coupled partial differential equations. These equations describe the rate of change of state variables, C, that represent the concentrations of nutrients, the biomass of functional plankton groups and more[12,13]. The common form of these equations is:

$$\frac{\partial C}{\partial t} = physics + bgc\_sms \quad (1)$$

where *physics* includes all advective and dispersive transport processes affecting the concentration of C, and *bgc_sms* contains all local sources and sinks due to biogeochemical transformations, air–sea gas exchange, atmospheric deposition, sediment–water exchange, river input and any transport not arising from ocean circulation, such as vertical sinking of organic matter. Biogeochemical state variables are specified as a concentration of some element, often nitrogen. Other elements such as carbon, phosphorus, silicon or iron can also be included using either fixed or variable elemental stoichiometry to relate the different state variables.

One of the lowest-order, but complete, biogeochemical models is the nutrient–phytoplankton–zooplankton–detritus (NPZD) model. The NPZD model describes the concentration of the four variables in a homogeneously

External forcing
All prescribed inputs that are needed to determine the evolution of a model's state and are not calculated internally by the model.

Projections
Simulations into the future that go significantly beyond the timescale for which models have demonstrated predictive or forecast skill, such as Earth system model (ESM) simulations to the end of the current century or longer.

Model parameters
Constants that are usually specified at the beginning of model integration and determine the dynamical behaviour of the model.

A priori knowledge
Assumptions about ocean processes, represented by the equations of an ocean model and its parameters and initial and boundary conditions, that are available before data assimilation is applied.

mixed volume or box. It is obtained by neglecting the physical terms to effectively create a zero-dimensional box model. Consequently, the equations are simplified to ordinary differential equations (dC/dt) and include only four state variables. Assuming a closed system, the terms on the right-hand side of the equation reflect transformations between the state variables:

$$\frac{dN}{dt} = -uptake + remineralization \quad (2)$$

$$\frac{dP}{dt} = uptake - grazing \quad (3)$$

$$\frac{dZ}{dt} = grazing - excretion \quad (4)$$

$$\frac{dD}{dt} = excretion - remineralization \quad (5)$$

Note that the four coupled NPZD equations above are mass conserving, where loss from one variable to another is balanced by a corresponding gain in the latter. Also, the four equations make a coupled system of equations because the terms on the right-hand side are dependent on multiple state variables.

The next step in modelling the NPZD system is to specify the functional form and parameters for each of the biogeochemical transformations, referred to as parameterizations. They are defined using conceptual understanding, also referred to as a priori knowledge, from laboratory experiments, field studies and biological theory[14]. For example, the grazing term of zooplankton consuming phytoplankton is a function of P, Z and, perhaps, temperature, T, namely $grazing = f(P, Z, T, \ldots)$. Below are three example grazing parameterizations:

$$grazing = gZ\frac{P}{K_P} \quad (6)$$

$$grazing = gZ\frac{P}{P + K_P} \quad (7)$$

$$grazing = gZ\frac{P^2}{P^2 + K_P^2} \quad (8)$$

where g (day^{-1}) is a rate parameter and K_P (same units as P) is a saturation parameter. All three parameterizations are common in ecological modelling for capturing consumer–resource interactions and reflect distinctly, often subtlety different, biological dynamics[15]. The first parameterization assumes that the grazing rate increases linearly with the prey concentration (P). By contrast, the latter two assume the grazing rate saturates at high concentrations of P, meaning $P \gg K_P$. The P^2 terms in the third parameterization result in reduced grazing at very low P concentrations. The choice of functional form for the grazing parameterization is typically based on theoretical arguments and considerations about the numerical model stability. The parameters for the grazing parameterization can be determined by dilution experiments for some zooplankton species, where phytoplankton loss rates are measured across a range of prey dilution levels. However, it should be recognized that single species experiments in the laboratory do not necessarily translate to diverse natural communities. Similar decisions on the functional forms and parameter values have to be made for all other parameterizations in the NPZD model and initial concentrations of all variables have to be prescribed. An example of a complete, vertically resolved NPZD model is provided in BOX 1.

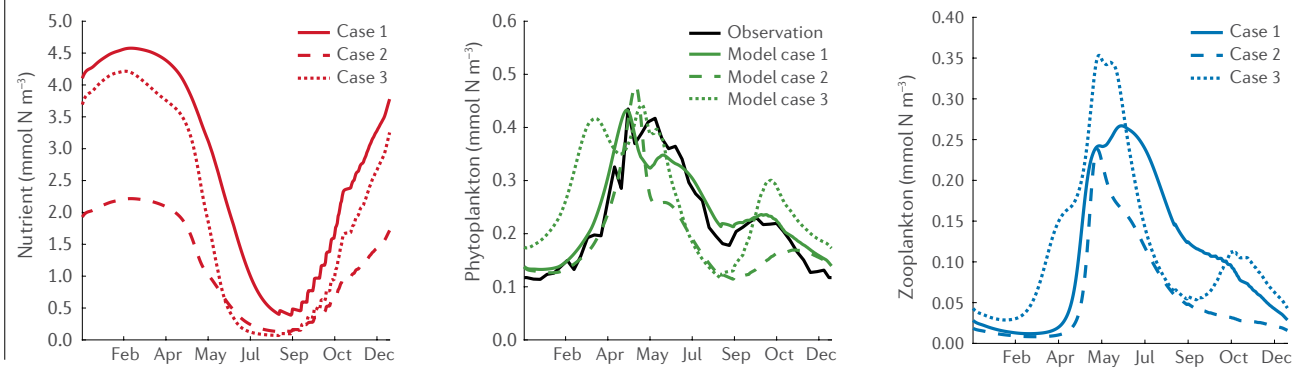
Most OBMs in current use are extensions of the basic NPZD framework but have more complex biogeochemical model components. Additional state variables include multiple nutrients — nitrate, ammonium, phosphate, silicate and dissolved iron — multiple phytoplankton and zooplankton functional groups, dissolved gases — for example, oxygen and dissolved inorganic carbon — and related properties such as alkalinity. Multiple nutrients are needed to address spatial and temporal switches between limiting nutrients and unique requirements by some functional phytoplankton groups, such as diatoms. Multiple plankton variables are included to account for the biogeochemically distinct roles played by different size classes and functional groups[16]. For example, diatoms have a unique requirement for silicate and contribute significantly to biological carbon export. Coccolithophores produce calcium carbonate ($CaCO_3$) as shells and affect vertical carbonate transport and remineralization at depth. Diazotrophs fix gaseous nitrogen, turning it into bioavailable forms. Furthermore, chlorophyll concentration, the most abundant biological measurement in the ocean, is non-linearly related to phytoplankton biomass. The chlorophyll to biomass ratio can vary by an order of magnitude due to photoacclimation[17,18]. Many, but not all, biogeochemical models account for variations in the ratio of phytoplankton biomass to chlorophyll using a parameterization of photoacclimation[19]. Including the inorganic carbon cycle is crucial for any OBM used for climate studies[20]. This requires inclusion of state variables for dissolved inorganic carbon and alkalinity, unless alkalinity can be inferred from other state variables, typically salinity. Knowledge of these two properties enables calculation of other carbonate system properties, including pCO_2, which is required to parameterize air–sea gas exchange, and pH, which is of considerable interest given concerns about ongoing ocean acidification. Another common state variable in OBMs is oxygen because of its relevance for climate and ecosystem health. Oxygen minimum zones in the open ocean are sites of trace gas production and nitrate loss via denitrification[21]. Low oxygen concentrations (hypoxia) or a complete absence (anoxia) have deleterious impacts on ecosystems.

Although virtually all OBMs for biogeochemical and climate studies follow the approach of defining a moderate number of functional groups, there are alternative approaches. One family of models initializes simulations with dozens to hundreds or more phytoplankton state variables, using either a randomly chosen or specifically crafted size structure and physiological parameters. The model then allows competition within the simulation

Box 1 | **NPZD model example**

Simple MATLAB code of the one-dimensional nutrient–phytoplankton–zooplankton–detritus (NPZD) model described in REF.[202] is available on GitHub[201]. The model represents a station in the subpolar North Atlantic Ocean. The default simulation yields a good agreement between the simulated phytoplankton concentration and satellite observations. After running the default, increasing and decreasing the initial nitrate concentration, maximum phytoplankton growth rate, maximum zooplankton grazing rate and latitude enable exploration of the model's dependencies on these parameters.

The figure shows the simulated surface concentrations of nitrate, phytoplankton and zooplankton variables in the second year of the simulation for the default parameter set (case 1), for a 50% decrease in the initial nitrate concentration (case 2) and for a doubling of the maximum phytoplankton growth rate (case 3). In case 1, the surface phytoplankton concentrations agree with satellite-based observations. The decrease in initial nitrate in case 2 has only a small effect on the timing and amplitude of the spring phytoplankton bloom, but leads to much smaller phytoplankton and zooplankton concentrations in summer and autumn than case 1. The doubling of the maximum phytoplankton growth rate in case 3 leads to a much earlier spring bloom initiation, a larger autumn phytoplankton bloom and larger spring and autumn peaks in zooplankton than for case 1.



to sub-select regional and seasonal plankton communities[22,23]. Another family of models uses allometric relationships to represent a continuum of plankton size classes to simulate grazing relationships and distinct trophic interactions in different marine ecosystems[24,25]. These two approaches move in the direction of representing more of the complexity inherent in natural plankton communities. Others have moved in the direction of less complexity by drastically reducing the number of biogeochemical state variables to four[26].

Model uncertainties enter the OBM biogeochemical equations from several sources. The models have many parameters which are not well known or easily quantifiable[27]. Even parameters that can be determined experimentally may not effectively represent real-world communities in the field. Furthermore, model parameters are not independent for coupled differential equations, and system-level uncertainties can arise because of dynamical interactions between state variables. Parameter optimization aims to address this issue but depends critically on the availability of a broad suite of observations. More challenging uncertainties arise from the choice of model structure and model parameterizations. Coupling of biogeochemical equations with ocean circulation results in additional sources of errors. Careful validation of OBMs to evaluate whether they are fit for purpose, be it for a specific scientific question or an applied purpose, is an integral part of model development and application.

Parameter optimization
The determination of the most likely values of poorly known model parameters based on the agreement of model output with observations.

Coupling with ocean circulation. In an OBM, the transformations between biogeochemical state variables are connected to their advective and dispersive transport arising from ocean circulation, by partial differential equations of the general form given by Eq. 1. This equation can be rewritten as follows for each state variable C:

$$\frac{\partial C}{\partial t} = -\mathbf{u}\cdot\nabla_3 C + \nabla_2\cdot k_H\nabla_2 C + \frac{\partial}{\partial z}\left(k_V\frac{\partial C}{\partial z}\right) + bgc\_sms \quad (9)$$

where the first term on the right-hand side represents the advective transport of constituent C ($\mathbf{u}$ is the fluid velocity vector), the second and third terms represent dispersion in the horizontal and vertical directions, respectively, and the last term refers to the biogeochemical sources and sinks of C. The parameters k_H and k_V are the horizontal and vertical dispersion coefficients and $\nabla_3 = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$ and $\nabla_2 = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y})$ are three-dimensional and two-dimensional operators. The combination of the first three terms on the right-hand side is referred to simply as *physics* in Eq. 1. As physical transport processes operate in all three spatial directions, Eq. 9 is three-dimensional in space and includes partial derivatives with respect to time, t, and the three spatial dimensions, x, y and z. In addition to an equation of this form for each biogeochemical variable, an OBM includes partial differential equations for the physical state variables, including temperature, salinity and velocity, as well as parameterizations for horizontal and vertical dispersion coefficients, which can vary in space and time. For detailed descriptions of the physical model equations, see REFS.[13,28].

Except for a few highly idealized cases — for example, when considering only one spatial dimension or a circular or rectangular two-dimensional domain with homogeneous initial conditions and constant forcing — the solution to these equations cannot be obtained

analytically and must be approximated numerically. Commonly, the equations are discretized in time, using finite time steps, Δt, and space, on a three-dimensional grid representing the model domain, with the help of finite differences. In the finite difference methods, the derivatives in the differential equations are replaced by finite difference approximations, for instance $\partial C/\partial t$ and $\partial C/\partial x$ become $\Delta C/\Delta t$ and $\Delta C/\Delta x$, respectively. This results in a system of prognostic equations, which include only basic arithmetic operations on defined quantities that can be carried out on a computer. There are subtle issues and several options when defining spatial grids, the finite difference discretization of equations on the grids and time stepping[29,30]. This explains the large diversity of finite difference-based ocean circulation models in current use. Alternative approaches are finite element and finite volume methods, which both allow for the creation of unstructured, curved, varying resolution grids, especially suitable for models with complex coastlines or bathymetry. Owing to the subtleties of discretizing differential equations on unstructured grids, these methods are an area of active research[31] and, to date, only a few OBMs have been coupled to finite element[32] and finite volume[33] circulation models.

Finite difference approximations are not exact solutions to the OBM equations; they only approximate the solutions. The accuracy of these approximations depends on the chosen difference scheme, the size of the time step, Δt, and the spatial resolution, Δx, Δy and Δz. It is generally desirable to use the finest spatial resolution possible with the available computational resources and the longest time step that keeps the model numerically stable. Although computing power has increased over the past two decades, the computational cost of running realistic OBMs is so demanding that trade-offs between domain size, resolution and integration time must always be considered. When doubling the horizontal resolution Δx and Δy in a finite difference model, the maximum allowable time step Δt is shortened to a quarter of its previous value, leading to a factor of 16 increase in overall computation time.

ESMs have a spatial resolution on the order of 1° (~100 km; FIG. 2) and are typically integrated for several hundred years. Given their spatial resolution, the models are unable to capture a range of important bathymetric and circulation features, such as continental shelf edges, mesoscale eddies and currents, and river plumes. Regional models have a finer spatial resolution (FIG. 2) on the order of single to tens of kilometres, but they have much shorter integration times (months to decades). The models' pros and cons can be illustrated using the north-west North Atlantic Ocean as an example. The broad, passive-margin shelf in this region, located at the confluence of two large-scale current systems, the Gulf Stream and the Labrador Current, supports economically and culturally important fisheries that are particularly vulnerable to warming and ocean deoxygenation[34,35]. A defining circulation feature in this region is the shelf-break current, a branch of the Labrador Current system, which effectively isolates shelf water from adjacent open-ocean water leading to distinct properties and long residence times[36]. Owing to their low resolution, global models typically lack the shelf-break current and cannot reproduce these features[37]. As a result, they do not well reproduce biogeochemical properties in this region[38,39]. Recent efforts to increase a global OBM resolution to the level of a regional model showed that properties are simulated much more realistically when the shelf-break current is properly represented[40]. However, the computational effort was so large that its integration time was limited to 100 years. Only a highly simplified three-variable model based on REF.[26] was included, and the model cannot be run routinely given present computational resources. The drawback of a high-resolution, regional model is that its integration time is limited to decades. Additionally, atmospheric and larger-scale ocean forcing must be specified instead of evolving internally as is the case for ESMs.

Running an OBM involves integration forward in time from defined initial conditions for each state variable. It is subject to external forcing and boundary conditions at the model domain edges. Some initial conditions, such as temperature and salinity distributions, must be prescribed, whereas the model may start from rest, with initial velocities at zero. External forcing of the ocean circulation component includes solar radiation; air–sea fluxes of momentum, for example wind forcing, heat and freshwater, such as precipitation minus evaporation, sea-ice formation and melt; and freshwater inputs from rivers. Examples of boundary conditions include the stipulation that fluid flow cannot be normal or at a right angle to the coast and may be required to vanish at the coast — the no slip boundary condition. Regional models typically have lateral boundaries that do not coincide with coastlines, referred to as open boundaries. Flow and associated transport of seawater constituents across these open boundaries must be specified for regional models, which is one of their drawbacks.

For the ocean circulation component, initial and boundary conditions must be specified for the biogeochemical state variables. The distributions of nutrients, dissolved gases, alkalinity and long-lived organic pools, for instance long-lived dissolved organic matter, should be prescribed as accurately as possible for initial and open boundary conditions. Pools with fast turnover times, such as plankton groups and reactive detrital pools, can be set to small positive numbers and will adjust quickly during model spin up. Additional boundary conditions include nutrient and organic matter concentrations in river inputs, the mole fractions of gases in the atmosphere, atmospheric deposition and exchange fluxes across the sediment–water interface.

Ocean biogeochemical processes can also influence and feed back on ocean physical dynamics. For example, several model studies investigated the influence of phytoplankton chlorophyll on the penetration depth of solar radiation. This highlighted its effect on the vertical profile of surface heating and upper ocean stratification[41]. More indirect influences are possible in coupled ESMs via changes in the ocean's production and emission of radiatively active gases, including dimethyl sulfide (DMS)[42].

Integration time
The simulated length of model integration. It varies from months to decades in regional models and hundreds of years in Earth system models (ESMs).

Spin up
The initial period of a model simulation during which the model adjusts from its initial state to a new state according to the internal model dynamics and subject to external forcing. The spin up period ranges from a few months or years for regional models to one or a few hundred years for global models.

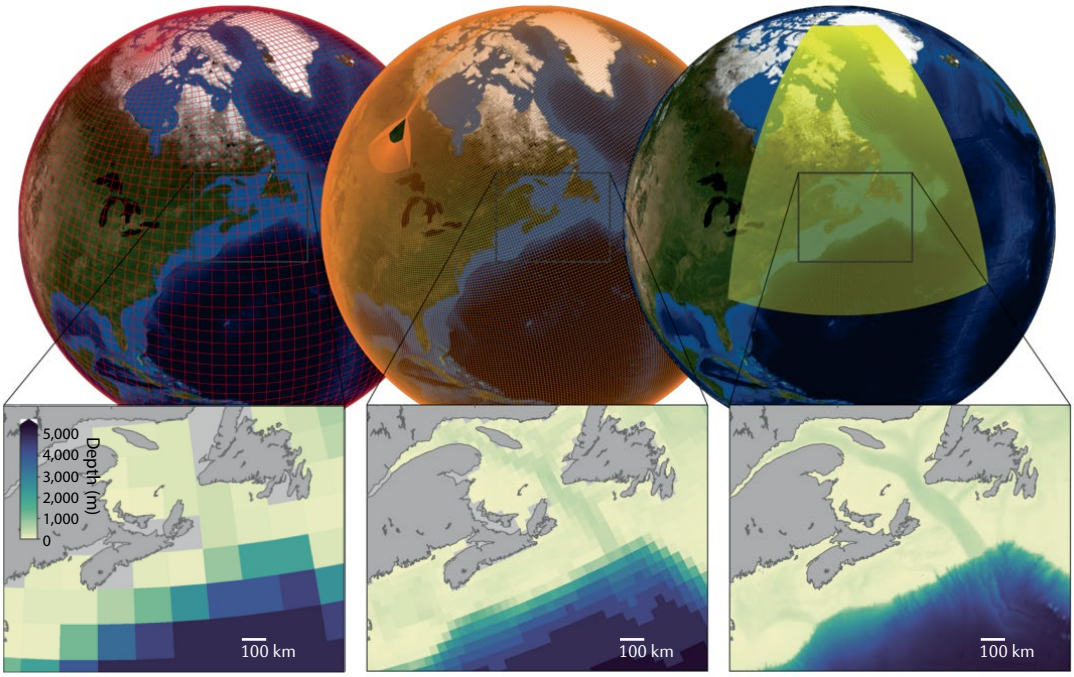


Fig. 2 | **Typical horizontal resolutions and bathymetries in global and regional models.** The left and middle globes show the grids of two global models that are part of the Coupled Model Intercomparison Project's 6th Assessment Round (CMIP6): the IPSL-CM5A-LR grid with a horizontal resolution of about 160 km in the corresponding inset (left) and the MPI-ESM1-2-HR grid with a resolution of about 30 km in the inset (middle). The right globe and inset show the domain extent and horizontal resolution of a regional ocean biogeochemical model (OBM) for the north-west North Atlantic and Labrador Sea, with a resolution of about 5 km in the inset.

State estimation
A method to obtain the optimal model state by combining the information contained in the model equations and the available observations.

Variational methods
Methods aimed at obtaining the best fit, in a least-squares sense, between model and observations by minimizing a cost function. These can be applied to parameter and state estimation problems.

Sequential methods
The model state, and, sometimes, its parameters are updated through an alternating sequence of forecast steps when the model is integrated forward in time, and update or analysis steps when the model state and, if applicable, the parameters are updated using observations.

Cost function
A measure of the misfit between observations and their model counterparts in a least-squares sense.

Combining models and observations

Data assimilation, the process of statistically combining models and observations, has the overarching goal of achieving the best possible representation of past, current or future ocean states. Data assimilation methods combine the a priori knowledge of the ocean state and its processes that is contained in an OBM with observations. Two applications for OBMs are parameter optimization and state estimation. Parameter optimization is aimed at addressing systematic biases in models that arise from inaccurate parameter values and initial or boundary conditions. State estimation typically assumes the model is unbiased and aims to correct random errors, such as deviations between the observed and simulated ocean state due to stochastic processes, for example the ocean's eddy field or larger-scale variations such as the North Atlantic Oscillation and El Niño-Southern Oscillation. Both applications can be realized through variational methods or sequential methods (BOX 2). Typically, OBM implementation includes parameter optimization initially to remove biases within the model, potentially followed by state estimation to minimize random errors.

Parameter optimization. Variational data assimilation derives from the mathematical field calculus of variations, which uses small variations in the inputs of functions to find their minima or maxima. A long-standing OBM application is parameter optimization[43–45], where poorly known model parameters are varied systematically to minimize the misfit between observations and their model equivalents across the whole integration period. This results in better agreement between the model and observations (FIG. 3). The misfit is measured by the cost function, typically of the form:

$$J(\mathrm{p}) = \frac{1}{N}\sum_{i=1}^{N} w_i (y_i - \hat{y}_i(\mathrm{p}))^2 \text{ or}$$
$$J(\mathbf{p}) = (\mathbf{y} - \mathbf{H}\mathbf{x}(\mathbf{p}))^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}(\mathbf{p})) \quad (10)$$

where $\mathbf{p}$ is a vector of the parameters to be optimized, also referred to as the control vector; $\mathbf{y}$ is a vector of the available observations; $\hat{\mathbf{y}} = \mathbf{H}\boldsymbol{x}(\mathrm{p})$ is a vector of the model equivalents to these observations, obtained by mapping the model state $\mathbf{x}(\mathbf{p})$ onto the observation vector $\mathbf{y}$ using the linear operator $\mathbf{H}$; and w_i and $\mathbf{R}$ contain the weights that each observation contributes to the cost function. Typically, $\mathbf{R}$ is assumed to be diagonal, where the weights are inverses of the variances or based on the observation error for each observation type. $\mathbf{R}$ is considered to be the covariance matrix of deviations between model and observations, referred to as the observation error covariance matrix. In practice, assumptions about these weights must be made.

Solving this minimization problem yields the optimal parameters, which can be obtained with iterative gradient descent methods, for example conjugate gradient search[13] or stochastic approaches such as simulated annealing[43] or evolutionary algorithms[46–48]. Parameter optimization is widely applied to OBMs because they typically have many poorly known, difficult to determine and application-specific parameters that govern

Control vector
A vector containing all of the values to be optimized during data assimilation. It can include model parameters, the full model state, a subset thereof or a combination of both.

Optimal parameters
The results from parameter optimization; the parameter values that minimize the cost function in a parameter optimization problem.

the biogeochemical transformations. The method aims to extract information about these biogeochemical transformations, which is available in observations, to inform or constrain the poorly known parameters.

In practice, the success of parameter optimization depends on whether the available observations contain enough information to constrain the parameters to be optimized[49]. For example, chlorophyll observations are often the most abundant observation type. Relating chlorophyll to phytoplankton biomass is not straightforward because the chlorophyll to biomass ratio is variable and often not well known. Even if the chlorophyll to biomass ratio was known, chlorophyll observations are useful for informing phytoplankton-related parameters, such as the total phytoplankton growth rate, but may not contain much information about grazing, remineralization and species-specific growth rates. As a result, those variables would remain poorly constrained after optimization using only chlorophyll or chlorophyll and nutrient observations[44]. Because observational data sets are often limited in terms of the biogeochemical properties available, this is a common problem for OBMs, referred to as the underdetermination problem[50]. A closely related challenge is that of interdependent, correlated or non-unique parameters, which arises when different combinations of parameters yield the same result. For example, a reduction in plankton mortality and an increase in plankton growth may give the same change in biomass. The available observations may not provide enough information to distinguish between multiple plausible combinations without further information. An example of underdetermined and interdependent parameters is given in FIG. 4. An a posteriori error analysis provides insight into interrelated and poorly constrained parameters for a specific optimization problem[44,49].

Parameter optimization is routinely used in biogeochemical modelling[51,52]. It is often an integral step in model development and, if applied systematically for different model structures, can guide model construction[53,54]. Parameter optimization and state estimation are sometimes combined[55,56].

Parameter optimization worked example. Python and MATLAB codes for parameter optimization of a zero-dimensional (single box) NPZD model are available on GitHub[57]. These code examples perform twin experiments using the stochastic ensemble Kalman filter (SEnKF). The default setting (FIG. 4) illustrates the effect of both interdependent and underdetermined parameters. When only phytoplankton observations are available, there is a tight interdependence between phytoplankton growth and mortality rates. For example, low growth and low mortality rates give a similar fit to high growth and high mortality. Likewise, phytoplankton observations contain little information about the nutrient remineralization rate, which is not improved by assimilation in this example and remains underdetermined. Using different combinations of observation types and parameter sets in the assimilation example allows these issues to be explored further.

In the default example, three parameters are estimated as follows. First, ten synthetic phytoplankton observations are generated from a model simulation with known parameters, labelled 'true' in FIG. 4a. The prior estimate of the parameters (FIG. 4b) results in a large spread of the forecast ensemble in state space. In FIG. 4c,d, true and forecast ensemble parameters are shown as black plus signs and blue dots, respectively, overlaying the misfit between model and synthetic observations. The forecast and analysis parameter ensemble means are shown as a large blue dot and green square, respectively, in FIG. 4c,d. Assimilating the data moves the mean parameter estimate closer to the true values in parameter space for the phytoplankton parameters (FIG. 4c) but farther from the true values for the nutrient remineralization rate (FIG. 4d). The analysis ensemble, shown in FIG. 4e, envelops the observations more tightly than the forecast ensemble.

Box 2 | General data assimilation machinery

In data assimilation, an optimization problem is solved where the initial control vector — also referred to as the background or initial guess — is updated to minimize, in a least-squares sense, the misfit between observations and their model equivalents. In some cases, a combination of the misfit between the initial and updated control vector and the misfit between observations and their model equivalents is minimized.

The solution to the optimization problem can be written as:

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}[\mathbf{y} - H(\mathbf{x}_b)] \quad (13)$$

where $\mathbf{x}_a$ is the optimized control vector; $\mathbf{x}_b$ is the initial control vector; $\mathbf{K}$ is the Kalman gain matrix; $\mathbf{y}$ is a vector containing the observations to be assimilated; and H is a non-linear operator containing the ocean model, which maps the initial control vector onto the observations. This equation applies to both parameter and state estimation but is more intuitive for the latter. In state estimation, $[\mathbf{y} - H(\mathbf{x}_b)]$ represents a vector of observation-model misfits that, when multiplied by the matrix $\mathbf{K}$, is projected onto the model state and yields the increments needed to obtain the optimal ocean state, $\mathbf{x}_a$.

The optimal solution $\mathbf{x}_a$ can be obtained by calculating or approximating $\mathbf{K}$, or by solving an equivalent minimization problem without explicit evaluation of $\mathbf{K}$. The true control vector, which represents the desired solution, is denoted by $\mathbf{x}_t$. Then, $\mathbf{x}_b - \mathbf{x}_t$ represents the deviations between the initial control vector and the truth, referred to as background errors, and $\mathbf{B}$ their covariance matrix. In the case of a linear model, H is also linear and denoted by $\mathbf{H}$. Observation errors are $\mathbf{y} - \mathbf{H}\mathbf{x}_t$ and their covariance matrix given by $\mathbf{R}$. Assuming unbiased background and observation errors and zero cross-correlation between these errors, the algebraic form of the gain matrix that provides the optimal analysis $\mathbf{x}_a$ is:

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} \quad (14)$$

Alternatively, $\mathbf{x}_a$ can be obtained by minimizing the cost function:

$$J(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + (\mathbf{y} - \mathbf{H}\mathbf{x})^T\mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}) \quad (15)$$

Equation 15 follows Bayes' theorem in that J is the argument of the Gaussian conditional probability function of $\mathbf{x}$ given $\mathbf{y}$. The first and second terms on the right-hand side of Eq. 15 are the arguments of the Gaussian distribution functions for errors in the background and errors in the observations, respectively. Identifying the $\mathbf{x}$ value that minimizes Eq. 15 is equivalent to maximizing the conditional probability $P(\mathbf{x}|\mathbf{y})$. Assuming a linear model is not strictly necessary to obtain $\mathbf{x}_a$ by minimization of the cost function (Eq. 15). For non-linear models, the term $\mathbf{H}\mathbf{x}$ can be replaced by $H(\mathbf{x})$.

Methods have been developed to obtain, or at least approximate, these solutions for realistic models with large control vectors. Ocean biogeochemical models (OBMs) are highly non-linear, have large state vectors and are computationally expensive. They also violate some of the underlying assumptions, such as Gaussian error distributions. Furthermore, the background and observation error distributions are not well known and not necessarily unbiased. Sequential methods estimate the gain matrix $\mathbf{K}$, whereas variational methods avoid explicit calculation of $\mathbf{K}$ and instead minimize $J(\mathbf{x})$, which is equivalent to maximizing $P(\mathbf{x}|\mathbf{y})$.

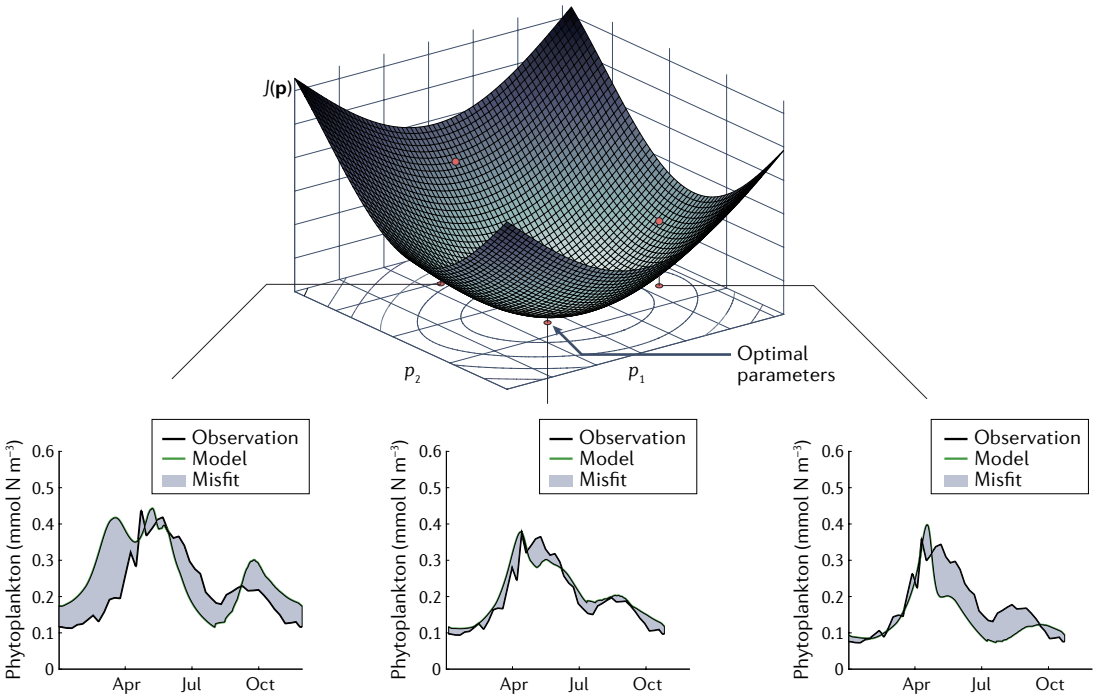


Fig. 3 | **Representation of a two-dimensional cost function.** The cost function $J(\mathbf{p})$ measures the misfit, indicated as the grey area in the three insets, between observations and their model equivalents in parameter space. Optimal parameters correspond to the minimum of the cost function and produce the best fit between the model and observations. **p**, control vector.

State estimation. State estimation is where a model's state variables are modified to reduce the discrepancy between the model and observations. It is typically applied sequentially by alternating forecast steps — where the model runs forward for a defined time window, usually a few days — followed by update or analysis steps, where newly available observations are used to update the model's state (FIG. 5). One of the most widely used and robust sequential data assimilation techniques is the ensemble Kalman filter (EnKF)[58]. The EnKF, its precursors and many variants apply Eq. 13 in BOX 2.

The EnKF is based on the Kalman–Bucy Filter[59], which yields the best possible estimate in a least-squares sense, when the model is linear and the distributions of the model state and observations are fully characterized by the mean and covariances[60]. The Kalman–Bucy Filter sequentially projects the model state, its mean and covariances forward in time using the, ideally linear, model. This is followed by a Bayesian update of the model state, its mean and covariances, informed by the newly available observations. However, application of the Kalman–Bucy Filter to OBMs is complicated by the models' non-linearity and the large state vector size, typically in the order of 10^8 or larger, which would require storing and modifying a prohibitively large covariance matrix. These issues led to development of the extended Kalman filter (ExtKF)[61] for non-linear models, and extensions thereof, such as the singular evolutive extended Kalman filter (SEEK)[62], which reduces the computational requirements for evolving the covariance matrix. However, propagation of the covariance matrix still requires linearization of the model, which can lead to bad approximations for highly non-linear models.

The EnKF removes the requirement for a linearized model by simulating mean and covariance directly with the help of a model ensemble. The underlying idea is that a model's probability distribution can be approximated by a finite model ensemble, which then allows relatively efficient calculation of the forecast error covariance. In the forecast step, ensemble members are propagated forward by the non-linear model. In the update step, the ensemble of forecasted states is used to compute the statistics required to perform the data assimilation update. More specifically, the covariance matrix **B** is approximated with the help of the ensemble of model states $\{\mathbf{x}^i\}$ as:

$$\mathbf{B} \approx < (\mathbf{x}_f - < \mathbf{x}_f >)(\mathbf{x}_f - < \mathbf{x}_f >)^T > \quad (11)$$

where $< \ldots >$ denotes the average over the ensemble[58]. The analysis step is:

$$\mathbf{x}_a{}^i = \mathbf{x}_f{}^i + \mathbf{K}[\mathbf{y}^i - \mathbf{H}\mathbf{x}_f{}^i] \quad (12)$$

for each ensemble member, where **K** is calculated as in BOX 2 using the ensemble approximation of **B**. The EnKF has been widely applied to OBMs and has many variants that differ in the way the update step is performed. One variant is the SEnKF[58,63], where an ensemble of observations is drawn from an assumed distribution of observations and used to update each ensemble member. Another variant is the deterministic EnKF (DEnKF)[64–67], where the mean and covariance of the ensemble are computed and updated using the new observations, and a new ensemble is drawn from the updated distribution.

A posteriori error
An estimate of the error in the solution of an optimization problem given the observations and numerical solution technique applied.

Least-squares
A measure of misfit between observations and the model equivalents of those observations that sums the squared distances between them.

During the EnKF forecast step, each ensemble member is integrated forward. This means the computational effort of running one realization of the OBM is multiplied by the number of ensemble members (FIG. 5). Computational constraints limit the possible size of the ensemble to between tens of and a few hundred members. The distribution of the high-dimensional state vector is under-sampled by ensembles that are computationally feasible. Two techniques, covariance localization and covariance inflation, are used to reduce the negative effects of this under-sampling. Localization decreases the impact of distant covariance estimates, thus reducing the effect of spurious long-distance correlations in the ensemble. Inflation artificially increases the ensemble covariances to counteract low covariance estimates due to small ensembles[68].

Analogues to the underdetermination and parameter interdependency problems also exist in the context of state estimation. OBMs have many biogeochemical state variables, most of which are not directly observed. State estimation is multivariate, meaning unobserved variables can be informed by observations of related variables through relationships expressed in the covariance matrix. However, many elements of the state vector may not be well informed by the available observations and the estimation problem is underdetermined. Furthermore, if an increase of one state variable, for example phytoplankton, is dictated by observations and achieved by different adjustments to the model's biogeochemical transformations, such as increasing nutrient supply or zooplankton grazing, additional observation types would be necessary to conclusively inform which update is correct[69,70]. Formal analysis of the impact of individual observation types can be useful in this context[71,72].

Although the EnKF is the most common sequential data assimilation technique, variational approaches are being applied to OBMs[73]. The three-dimensional variational (3D-Var)[74–76] and four-dimensional variational (4D-Var)[69,77] approaches include the sequence of forecast and analysis steps, where the analysis step uses the variational method (FIG. 5). In the 3D-Var approach, the observation operator H is assumed to be time independent, thought to be appropriate for short forecast windows. In the 4D-Var approach, H is time dependent and includes the non-linear forecast model, although it is common to linearize the problem[78]. Particle filters, such as sequential importance resampling[79], are promising alternative methods that do not rely on the assumption of Gaussian error distributions. They have been used for state and parameter estimation of OBMs[80] but are not yet widely used.

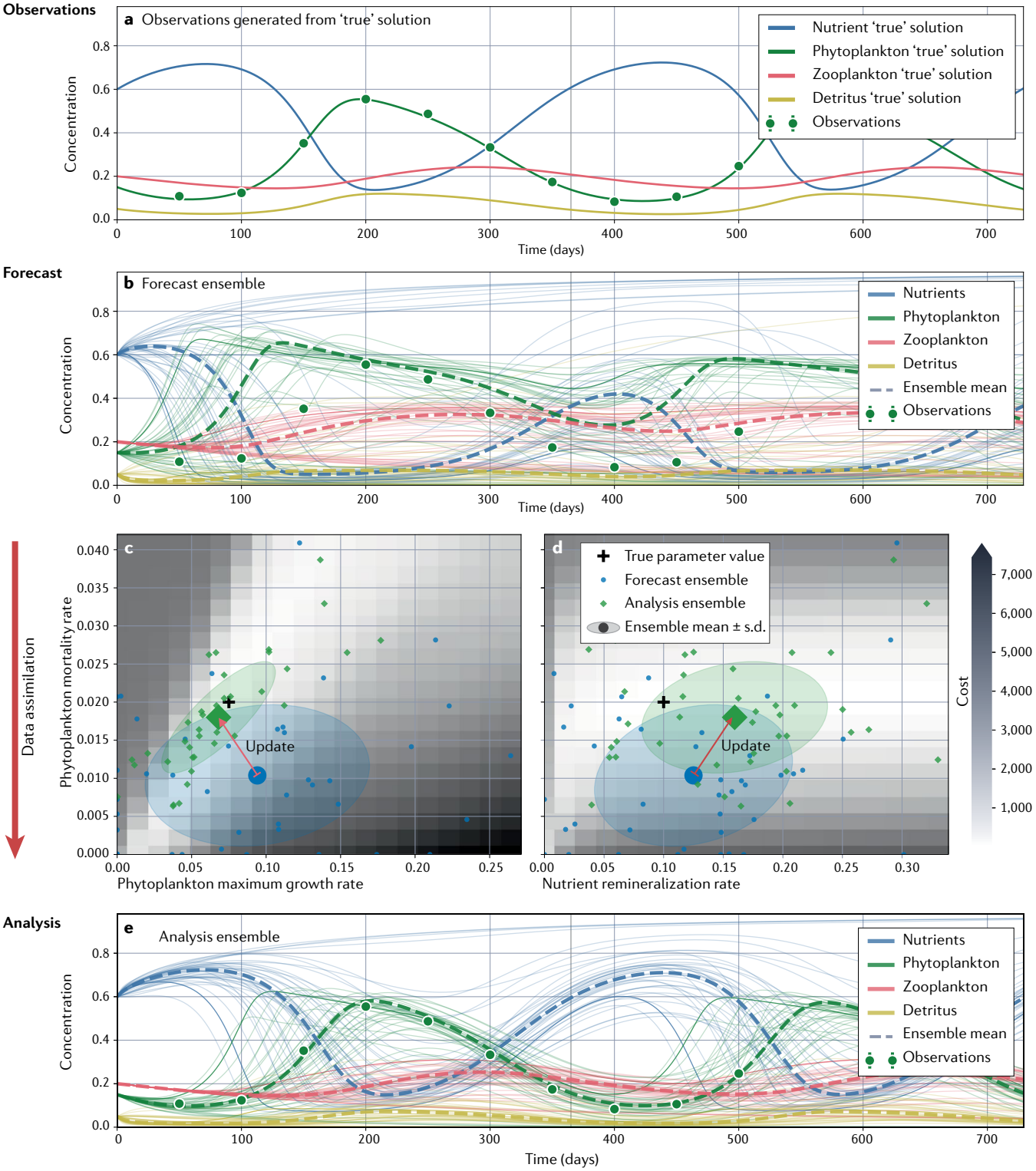
State estimation worked example. MATLAB and Fortran codes for ensemble-based state estimation suitable for three-dimensional OBMs are available on GitHub[81]. The example is set up as an identical twin experiment for an idealized three-dimensional OBM, using ROMS and the DEnKF as in REF.[64]. The model domain is an idealized north–south channel with periodic boundary conditions on the northern and southern boundaries and narrow shelves along the eastern and western edges. Wind forcing results in intermitted upwelling on the western side of the domain with upward transport of nutrients and stimulation of phytoplankton growth. In the free run, several biological parameters and the wind forcing are altered, resulting in a delayed upwelling and weaker phytoplankton response compared with the baseline run, which we consider the 'truth'. Sea surface height, surface temperature, surface chlorophyll and profiles of temperature and nitrate concentration are sampled from the 'truth' simulation and assimilated into the perturbed free run to obtain the data-assimilative simulation. The evolution of the mean surface chlorophyll concentrations (FIG. 6a) illustrates the differences between the 'truth' run (black), the perturbed free run (blue) and the sequential updates from forecast to analysis during update steps (red). The ensemble spread increases during most forecast steps. During assimilation steps the ensemble spread is reduced. The impact on surface and vertical distributions (FIG. 6b) of chlorophyll concentrations on day 16 illustrates the discrepancies between the 'truth' and free runs, and the improvement due to assimilation in the analysis.

Results

Model evaluation

Whether an OBM has value as a research tool depends on how accurately it represents the processes relevant to the scientific question being addressed. Evaluating a model's performance is an integral part of model analysis. It relies on comparing the model output with observations. Often, this occurs as an iterative loop, where the evaluation of a hindcast simulation is followed by model refinements, such as increasing model resolution, improving parameterizations or changing the model structure, followed by a new hindcast and evaluation[82]. The three most commonly used statistical metrics for model evaluation are the root mean-square error (RMSE), the bias and the correlation coefficient (BOX 3). All three are calculated by directly relating observations to their model counterparts. They are also all relative measures, without an objective criterion that indicates which range of values is acceptable or unacceptable. These metrics can be calculated using spatial and temporal averaging, temporal averaging only or spatial averaging only[83]. Specialized graphics have been devised to effectively represent some of these metrics for a large number of different models or different hindcasts from the same model, including the Taylor diagram[84] and the target diagram[85]. Two metrics with built-in criteria on the acceptability of a model's performance are Z-scores, which consider variability within the observational data set, and the model efficiency or model skill, which quantifies whether the model outperforms an observational climatology (BOX 3).

No single metric provides a complete evaluation of a model's predictive power. Instead, multiple, complementary metrics should be used in concert[86]. A model may provide accurate estimates for some variables, locations or times, but perform poorly for others[87]. Space, time and a breadth of variable types should be considered in any comprehensive model assessment. Furthermore, there may be aspects of a system that the model cannot



reasonably be expected to reproduce. For example, an OBM without state estimation cannot exactly reproduce stochastic aspects of the system, such as the exact timing and location of elevated chlorophyll due to mesoscale eddies. However, the magnitude, shape and frequency of eddy-induced chlorophyll enhancements are expected to be represented well. Specialized metrics that account for mismatches in space and time can be used for this purpose[88].

In OBMs with data assimilation, the need for model assessment expands further, to include evaluation of state or parameter estimates. This becomes more difficult, as these estimates already contain information from the observations that were assimilated. One general strategy is to only use observations in the evaluation that were not assimilated and can be considered independent[86]. In practice, this presents a conundrum because, ideally, all available observations would be used

Fig. 4 | **Application of a stochastic ensemble Kalman filter for estimating three parameters of a zero-dimensional (single box) NPZD model in a twin experiment using example code[57].** **a** | Ten phytoplankton observations (green dots) are generated by sampling a model simulation with known parameters (+ signs in panels **c** and **d**). This simulation represents the 'true' solution. **b** | The spread in model solutions prior to assimilation or analysis. The forecast ensemble includes the model simulations (transparent thin lines) for all parameter values (small blue dots in panels **c** and **d**) of the prior estimate. Dashed lines show the ensemble mean. **c** | Cost function values (in greyscale) over the phytoplankton mortality and maximum growth rates, shown with the forecast ensemble of parameters (small blue dots), the parameter ensemble after analysis (small green diamonds), the mean values of the parameter ensemble with standard deviation before and after analysis (large blue dot and green diamond, respectively, with transparent ovals) and the true parameter values (+ sign). Red arrow indicates the change in parameter estimate during the analysis step due to assimilation. For these two parameters, the analysis moves the mean parameter estimates close to the 'truth'. The contraction in the size of the transparent ovals before (blue) and after (green) analysis indicates improved confidence in the accuracy of the parameter estimates. **d** | Same as panel **c** but for phytoplankton mortality and nutrient remineralization rates. The estimate for the nutrient remineralization rate is not improved by the analysis, indicating that the observations contain little information on this parameter. **e** | Model ensemble after assimilation, represented in the same way as panel **b**. The ensemble mean for phytoplankton matches the observations very well, especially compared with panel **b**. However, not all parameter updates resulted in the 'true' values (see nutrient remineralization rate in panel **d**), which illustrates the underdetermination problem often encountered in parameter optimization. NPZD, nutrient–phytoplankton–zooplankton–detritus.

in assimilation to obtain the best possible estimates, and observations withheld from assimilation may be correlated to those used in assimilation, meaning they are not truly independent. Decorrelation scales must be considered when deciding what represents truly independent observations. Perhaps the most convincing assessment of an assimilative model is an ongoing test of its predictions against observations as they become available. A closely related approach can be used to assess sequential state estimates in hindcast mode. In this approach, misfits between observations and model forecasts at the current time are compared with misfits between current observations and analysis from the previous update step. The underlying idea is to test whether the forecast outperforms the persistence model, which assumes that the previous analysis is also the best forecast.

Challenges to model evaluation

The rigour of any model evaluation depends on the observational data set available. Despite major efforts in ocean observation in the twentieth and early twenty-first century, the ocean's biogeochemical state remains under-observed in many critical aspects. This has hampered the evaluation and systematic improvement of OBMs[89]. Major biogeochemical observation efforts include ocean colour satellites, coordinated ship-based initiatives to obtain global three-dimensional distributions, time series sites for a broad suite of observations, networks of ships of opportunity and numerous individual, investigator-driven cruises. Although each approach is valuable, biogeochemical under-sampling remains a problem due to the cost and effort involved in ship-based measurements. This limits sampling to a few instances in space and time. Additionally, satellite observations of ocean colour only provide information about plankton-related properties at the very surface of the ocean. The maturation of autonomous platforms — profiling floats and gliders — and miniaturized biogeochemical sensors has paved the way for cost-effective and routine observation of a broad suite of biogeochemical properties[90,91]. Autonomous observation technology is quickly becoming an additional, complementary data source, making it feasible to observe the global ocean in near-real time at an unprecedented spatial and temporal resolution, with an accuracy sufficient to detect climate-induced changes[92].

A technical challenge to rigorous model evaluation is that many model variables are not in the same currency or unit as observations. For example, chlorophyll is non-linearly related to phytoplankton biomass. Plankton observations can be in the unit of cells per volume or biovolume per volume, whereas the model uses moles of carbon or nitrogen per volume. Large zooplankton is often measured in towed nets, resulting in a spatially integrated measure. Definitions of plankton size classes are typically not well defined for models. The necessary conversions introduce uncertainties when comparing observations and models. Another difficulty is accessing existing observations. Whereas many of the major coordinated observational initiatives have provided sustained access to observations, often through individual repositories, small, investigator-driven efforts are typically not well positioned to guarantee long-term access to specialized observations. Instead, they depend on national or international repositories, such as NOAA MEDS. A unified data management approach, with standardized meta-data requirements and data formats that ensure discoverability and accessibility of existing data sets, has been identified by the oceanographic community as a common goal[93,94] and will greatly benefit OBM evaluation and improvement.

Model evaluation relies on climatologies, satellite-based estimates of chlorophyll and primary production, comprehensive time series for a small number of sites, focused process studies and, increasingly, global autonomous data sets. Typically, an available data set includes fewer properties than the OBM's state and only a small number of the biogeochemical transformations are observed. As a result, an excellent agreement between model and observation does not guarantee that the model's representation of unobserved properties and fluxes is correct, or that the model is a skilful predictive tool. Internal model errors may compensate to reach a seemingly correct result when judged by a limited data set. However, the result may have been obtained for the wrong reasons. It is therefore critical to continually evaluate models and their predictions against sustained observational data streams with an increasing breadth of observables.

Applications

OBM applications range from scientific, for example building fundamental understanding or for hypothesis testing, to practical, such as producing forecasts and model-derived products. This section gives key examples to illustrate the breadth and importance of different applications, but should not be considered a comprehensive description of OBM applications.

Decorrelation scales
The *e*-folding scales of the autocorrelation function of the property under consideration; the distance or period over which the autocorrelation decreases by a factor of 1/*e*.

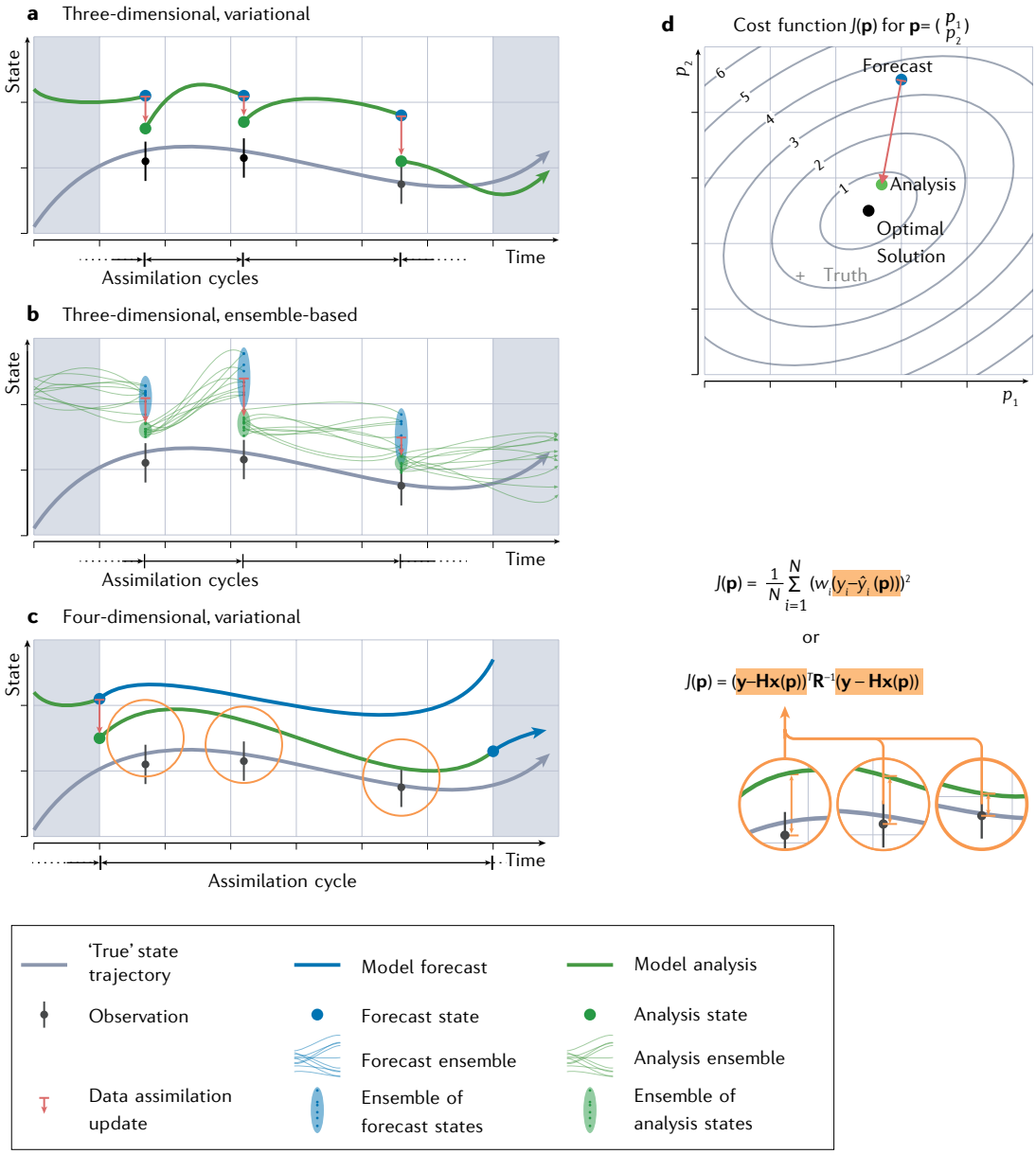


Fig. 5 | **Illustration of state estimation and parameter optimization. a–d** | In both state estimation (panels **a–c**) and parameter optimization (panel **d**), the goal is to use observations of the ocean's true state to obtain more accurate model estimates, by improving either model parameters or the model state itself. In parameter optimization, the misfit is measured by a cost function *J*(**p**). The initial guess (analysis) is systematically improved by minimizing the cost function. The parameters at the minimum are referred to as the optimal parameters. In state estimation, the model state variables are updated sequentially, as observations become available, in a series of forecast and update steps. **H**, linear operator; **p**, control vector; *T*, matrix transpose; w_i and **R**, weights that each observation contributes to the cost function; **x**(**p**), model state; **y**, vector of the available observations.

Ocean carbon accounting

The global ocean absorbs about a quarter of contemporary human emissions of CO_2 to the atmosphere[95]. OBMs have been central in quantifying the patterns and rates of ocean anthropogenic CO_2 uptake. This uptake occurs via natural physical–chemical gas exchange at the air–sea interface, followed by ocean circulation that transports surface water with excess CO_2 into the ocean interior[96]. OBMs are pivotal for characterizing future ocean CO_2 uptake and its sensitivity to ocean climate change under different policy scenarios[97]. Synthesis of ocean carbon observational and model information is invaluable for efforts to quantify the contemporary global carbon budget[98,99]. It is also needed to assess the predictability of global-scale atmosphere–ocean CO_2 flux relevant for carbon policy and management[100]. Partly because of the large CO_2 uptake capacity of the ocean, several approaches have been proposed to enhance ocean uptake through deliberate CO_2 removal or negative emissions technologies[101]. Rapid decarbonization of the global economy is needed to meet the international Paris Climate Agreement to keep global surface

warming well below +2.0 °C, relative to pre-industrial conditions. Coupled carbon–climate models indicate that society must meet roughly net-zero human CO_2 emissions by the middle of the twenty-first century. Given the challenges of abating all human CO_2 emissions from energy and transportation systems, a substantial amount of deliberate CO_2 removal may be required[97].

Significant knowledge gaps exist for all ocean-based CO_2 removal approaches. Unknowns range from the efficacy of net CO_2 uptake to permanence of carbon storage, method verification or carbon accounting, scalability and environmental impacts[102]. In conjunction with laboratory and field experiments, OBMs are central to resolving these questions across a range of scales: local, regional and global. Deliberate CO_2 removal will be challenging to verify because it represents a relatively small perturbation of the large natural uptake of anthropogenic CO_2 and natural background variations. Ocean circulation would transport added CO_2 away from the site of deliberate manipulation and dilute signals. Alterations of the ocean's carbonate system and nutrient inventory would have downstream effects ranging from desirable — countering ocean acidification for alkalinity enhancements — to counterproductive, for example by diminishing the additionality of CO_2 removal, or enhancing acidification. OBMs in combination with well-resolved, comprehensive observation will be central to verification of CO_2 removal and carbon accounting.

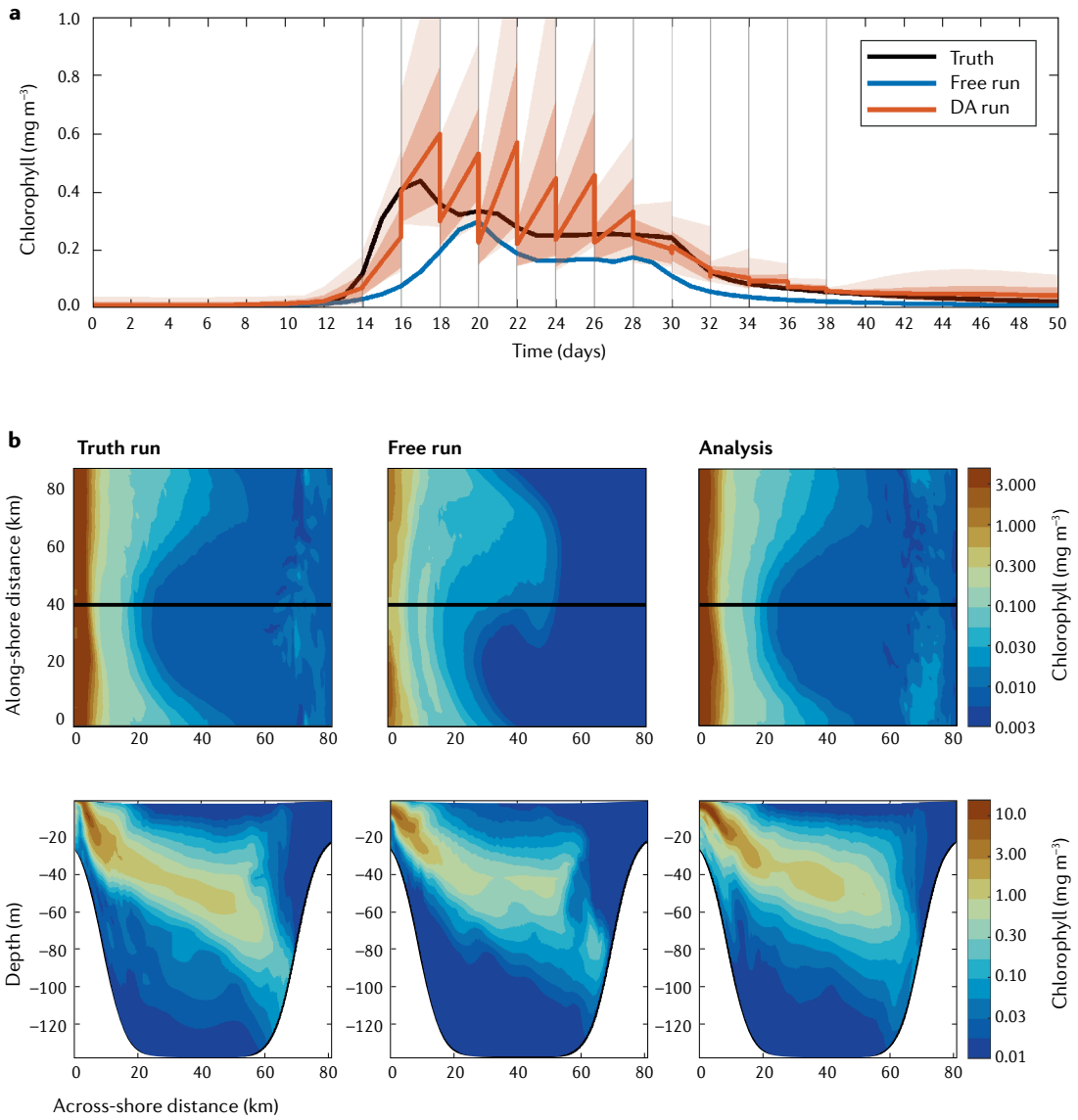


Fig. 6 | **Application of ensemble-based state estimation for a three-dimensional model, using example code[81].** **a** | Mean surface chlorophyll concentration in the 'truth' run (black line), the perturbed free run (blue line), the model ensemble (red line) of the data assimilation (DA) run, and the ensemble spread (full range shown in light red, ±1 standard deviation of the ensemble in darker red). Grey vertical lines indicate the timing of assimilation steps. **b** | Surface and vertical distributions of chlorophyll concentrations on day 16 in the 'truth' run, the free run and the analysis (mean of the ensemble after the assimilation step) from left to right. The location of the vertical transect is indicated by the black line in the middle panels.

Box 3 | **Common statistical metrics for model evaluation**

The root mean-square error (RMSE), defined as $\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(m_i - o_i)^2}$, is a measure of the overall distance between observational data points o_i ($i = 1, \ldots, n$) and their model equivalents m_i.

The bias, b, defined as $b = \frac{1}{n}\sum_{i=1}^{n}(m_i - o_i)$, measures to what degree the model overestimates ($b > 0$) or underestimates ($b < 0$) the observational data.

The correlation coefficient, r, defined as $r = (\sum_{i=1}^{n}(m_i - \overline{m})(o_i - \overline{o}))/\sqrt{\sum_{i=1}^{n}(m_i - \overline{m})^2 \sum_{i=1}^{n}(o_i - \overline{o})^2}$, measures to what degree the observations and their model equivalents are linearly related. For $r = 1$ they would perfectly relate to each other or be perfectly correlated. For $r = 0$ there would be no relation, and for $r = -1$ they would be perfectly anti-correlated, meaning whenever observations increase the model equivalents decrease by a proportional amount.

A climatology, $\overline{o}$, is the long-term average of an observational data set $o_i (i = 1, \ldots, n)$. A climatology can be spatially resolved, for example a gridded field, or calculated for a single location. It can be a temporal average over all data (annual climatology), temporally resolved by month (monthly climatology), day (daily climatology) or another averaging interval.

The model skill or model efficiency, m_e, defined as $m_e = 1 - (\sum_{i=1}^{n}(m_i - o_i)^2)/(\sum_{i=1}^{n}(\overline{o} - o_i)^2)$, measures whether a model results in better ($0 < m_e < 1$) or worse ($m_e < 0$) predictions than an observation-based climatology $\overline{o}$.

The Z-score, Z_i, defined as $Z_i = (m_i - \mu)/\sigma$, relates model output m_i to corresponding observational data of the same property. Assuming the observational data are normally distributed, with mean μ and standard deviation σ, the Z-score indicates the probability of encountering the value m_i in the data set given natural variability. A Z-score of 0 occurs when m_i is equal to the mean of the observations. Z-scores above 1 or below −1 indicate that m_i is outside one standard deviation of the observations.

Previous OBM studies have explored some of these questions for different CO_2 removal techniques, including ocean iron fertilization of high-nitrate, low-chlorophyll regions[103–105], artificial upwelling of nutrients into the surface ocean[106], macroalgae farming[107], seawater alkalinity enhancement[108,109] and, more generally, the permanence of CO_2 removal[110].

Ocean ecosystem health

Deoxygenation. Dissolved oxygen is an important measure of ocean ecosystem health because oxygen is essential for supporting aerobic aquatic life. Long-term observations indicate that the oxygen content of the global ocean has declined by more than 2% over the past five decades[111]. This has raised concerns about profound effects on ocean biogeochemical cycles and marine ecosystems[112]. The observed oxygen loss of the global ocean is projected to continue[113,114], primarily because of a decrease in oxygen solubility, increases in biological oxygen demand and reduced ventilation of the deep ocean from global warming[115,116]. In addition to the large-scale climate effect, coastal waters are affected by growing anthropogenic nutrient inputs that lead to a worldwide expansion of coastal hypoxia[117,118].

Global OBMs and ESMs have been used to understand why the global ocean oxygen content changes[119] and to make future projections under different emission scenarios[113,114,120]. These models consistently project continued and accelerating deoxygenation, but underestimate observed deoxygenation rates and fail to accurately reproduce observed patterns and temporal variability of oxygen changes[116]. Notable differences in the simulated intensity and spatial patterns of oxygen projections among ESMs[120,121] point to deficiencies in mechanistic understanding and modelling capabilities. Likely factors limiting current models have been identified in multi-model comparisons and sensitivity experiments, including insufficient model resolution[122], inaccuracies in ocean mixing parameterizations[123,124] and incorrect model representation of biological processes[119].

Regional OBMs are widely used to improve understanding of coastal oxygen dynamics and to guide management in coastal regions[125,126]. They allow examination and quantification of the factors governing oxygen variability and hypoxia formation[127–130]. Alongside this, they can be used to project changes in oxygen supply under climate change[131,132] and to understand the consequences of hypoxia on the marine food web[133,134]. Regional OBMs are used for a range of applied purposes. For example, they can evaluate how hypoxia would be affected by different nutrient reduction scenarios[135,136], investigate the compounding effects of anthropogenic nutrient inputs and climate change on hypoxia[136–138], provide for seasonal forecasts[139,140] and allow exploration of eco-engineering strategies for hypoxia mitigation[141]. Data-assimilative OBMs provide short-term ecological forecasts, including for oxygen, in various coastal systems by optimally combining models and observations[73].

Acidification. Ocean uptake of anthropogenic CO_2 slows its atmospheric accumulation and, in turn, slows climate change. However, it also changes seawater chemistry by elevating dissolved inorganic carbon and aqueous CO_2, while reducing pH and carbonate mineral saturation states[142]. Because it shifts seawater pH towards acidic conditions, this process is referred to as ocean acidification. Acidification likely has deleterious impacts on ocean ecosystems and coastal human communities that depend on marine resources[143,144]. OBMs are used extensively to quantify past and future rates and patterns of ocean acidification.

Global OBMs simulate future reductions in surface planktonic $CaCO_3$ production and elevated, shallower $CaCO_3$ remineralization[145–147]. These simulations suggest that open-ocean surface acidification is controlled largely by the choice of atmosphere CO_2 scenario. This is due to relatively rapid air–sea CO_2 gas exchange on annual and longer timescales. By contrast, subsurface acidification is more strongly dependent on simulated ocean ventilation rates, which differ across models[148]. Decadal prediction systems using ensemble forecasts from ESMs have demonstrable skill for surface pH variations for up to 5 years[149].

In coastal ecosystems, acidification is compounded by eutrophication, acidic freshwater discharge, coastal upwelling and terrestrial organic carbon inputs. Regional OBMs are used to analyse the synergy between acidification and eutrophication[150] and for characterizing the highly variable physical and biogeochemical conditions[151,152]. These models have also been used to quantify the time of emergence when anthropogenic changes exceed natural variability[153]. Further applications include investigating how anthropogenic CO_2 trends amplify the frequency of extreme acidification events[154] and compound events with overlapping extremes of acidification, marine heatwaves and deoxygenation[155].

Eutrophication
An excessive supply of plant nutrients to a body of water, often due to input from land.

The biological impacts of ocean acidification likely vary across different marine environments — coral reefs, wetlands, shallow coastal systems and pelagic planktonic systems — with effects extending across scales from organisms to the community and whole ecosystem. This is accompanied by positive and negative effects for different taxonomic groups[144]. For planktonic systems, elevated aqueous CO_2 is projected to increase primary and secondary productivity, alter community structure and, perhaps, increase the frequency of harmful algal blooms. At the same time, reduced carbonate mineral saturation states are projected to lower the competitiveness of calcifying plankton species[156]. This poses considerable challenges for OBMs because existing model structures and parameterizations are tailored to present conditions. They are not necessarily able to account for functional changes at the organism and community levels due to acidification.

Fisheries. Primary production by phytoplankton sustains the marine food web. However, OBMs typically include only species on the lowest trophic levels, such as phytoplankton and zooplankton, with the implicit assumption that predation on plankton by higher trophic levels, for example fish, can be represented by an additional mortality term in the zooplankton equation. The clear dependence of fishery yields on ecosystem primary production[157] led modellers of higher trophic level processes to use simulated primary production from OBMs to force models in the early 2000s. Examples include a study of the impact of climate change on tuna populations in the tropical oceans[158], and a study using IPCC-type models to assess the impact of climate on living marine resources[159]. This type of work continues, as exemplified by the recent FishMIP model intercomparison project[160] where a set of global, upper trophic level models were forced with outputs from ESM projections. These projections reveal that average animal biomass in the global ocean could decline by 17–19% in high emission scenarios and 5–7% in low emission scenarios by 2100 (REFS.[161,162]). This is primarily due to increasing temperature, decreasing primary production and changes in species ranges. These OBM projections have been used to project changes in fish catch potential[163] and global fishery revenues[164]. Although these results have been widely used in international assessments[165,166], the projections are subject to large uncertainties with a substantial uncertainty contribution from OBM-projected lower trophic level biomasses and production[161].

More recently, direct coupling of OBMs to higher trophic level models has enabled examination of top-down control from higher trophic levels on planktonic ecosystems and marine biogeochemical cycles in general. For example, a study that coupled an OBM to an upper trophic level model with explicit representation of vertically migrating organisms, which feed at the surface but excrete and respire at greater depth, estimated that diurnal vertical migration contributes significantly to the biological carbon pump. It could amount to 20% of the vertical carbon flux due to settling particles[167]. Diagnostic analyses of the contribution of zooplankton's diel vertical migration to biological carbon export have yielded similar results[168].

Observing system design

Observing system simulation experiments (OSSEs) are a type of data assimilation experiment in which synthetic, or simulated, observations are used to design new, or modify existing, observing systems. OSSEs have their origin in numerical weather prediction[169] but are increasingly used for ocean models[170], including OBMs[65,76]. Typically, OSSEs are performed when a new observing system is designed, or in anticipation of a proposed change to an existing observing system, for example when a new instrument type or sensor becomes available. An OSSE provides information about the instrument combination and configuration that leads to the greatest improvement in forecast skill of a data assimilation system. This is particularly valuable information given the expense of new satellite-based sensors or in situ ocean observing arrays. Beyond guiding observing system design, OSSEs can also be used to prepare for new data, to develop and improve data processing, and to streamline the data assimilation system, before new data sets become available.

OSSEs produce a simulated set of observations and use a data assimilation system to examine their impact on the system's predictive skill. The simulated observations, which include representations of the existing operational observing system and the proposed additions, should be obtained from an independent model simulation that satisfies a few criteria[170]. This independent simulation is referred to as the nature run. The simulated observations sampled from the nature run are then assimilated into a data-assimilative model that closely resembles the operational system. This is referred to as the perturbation run. By comparing the perturbation run with a control run, where only simulated observations from the existing system were assimilated, the potential benefits of the new observational assets or configuration can be quantified. As the nature run can be sampled at any desired location and frequency, many different observing system configurations can be assessed. However, for an OSSE to yield reliable conclusions, discrepancies between the simulated observations and the data-assimilative model must have the same error statistics as the differences between real observations and the data-assimilative model.

Creating realistic simulated observations with representative error statistics is non-trivial and should follow the principles laid out in REF.[170]. Identical twin OSSEs use the same model to perform the nature and perturbation runs, where discrepancies between the two are introduced by modifying one or multiple model inputs, for example parameters, initial conditions or physical forcing. Although identical twin OSSEs are relatively easy to set up, they are likely to produce simulated observations with non-representative errors, leading to unrealistically optimistic forecast skill estimates[65]. The more desirable, fraternal twin OSSEs use a different model for generating observations from the assimilative system. OSSEs require careful calibration, comparison of the nature run with reality, and an examination of observation

innovation statistics and forecast compared with the assimilation of real data.

Reproducibility and data deposition

In alignment with efforts to make ocean observations freely and easily accessible, alongside a general movement to greater transparency in research dissemination, for example through open-access publishing, it is becoming standard for OBM codes and output to be made freely available. All of the widely used OBM codes are community efforts, with shared code repositories and active user groups (TABLE 1). Code archiving services such as github.com enable scientists to maintain and share their individual code repositories and collaborate with others. Archiving of model output presents a larger challenge because the size of raw model output easily exceeds available storage capacities. This makes some subsampling and curation necessary. Community-driven archiving services such as zenodo.org enable permanent and open archiving of data sets, including OBM output, with assignment of persistent digital object identifiers. Some physical and biogeochemical simulation outputs from more formal model intercomparison activities are archived and available through centralized data repositories[20,171].

Table 1 | **Important repositories examples for OBM codes, outputs and observations**

Repository type	Abbreviation	Name
Codes	ROMS	Regional Ocean Modeling System
	MOM	Modular Ocean Model
	HYCOM	Hybrid Coordinate Ocean Model
	NEMO	Nucleus for European Modelling of the Ocean
	MITgcm	Massachusetts Institute of Technology General Circulation Model
Outputs	OMIP	Ocean Model Intercomparison Project
	CMIP6	Coupled Model Intercomparison Project Phase 6, output data repository
	C4MIP	Coupled Climate Carbon Cycle Model Intercomparison Project
	RECCAP-2	Regional Carbon Cycle Assessment and Processes 2
	MAREMIP	MARine Ecosystem Model Intercomparison Project
	IOOS COMT	Integrated Ocean Observing System's Coastal and Ocean Modeling Testbed
Observations	GEOTRACES	An International Study of the Marine Biogeochemical Cycles of Trace Elements and Isotopes
	GO-SHIP	Global Ocean Ship-based Hydrographic Investigations Program
	SOCAT	Surface Ocean CO_2 Atlas
	GLODAP	Global Ocean Data Analysis Project
	BGC-Argov	Biogeochemical Argo
	BCO-DMO	Biological and Chemical Oceanography Data Management Office
	World Ocean Atlas 2018	Ocean hydrography, oxygen and nutrients
	SeaBASS	NASA SeaWiFS Bio-optical Archive and Storage System
	MAREDAT	MARine Ecosystem Biomass DATa

OBM, ocean biogeochemical model.

The ocean modelling community has a history of embracing model intercomparison. Different modelling groups collaborate to define a target simulation with a standard set of implementation and evaluation criteria[172,173]. By comparing different model architectures against common criteria, it is possible to identify shortcomings in individual models that can be remedied and assess model uncertainties. If these uncertainties cannot be eliminated, they must be considered when interpreting results. The IPCC has a well-established process for intercomparison of ESMs, referred to as the Coupled Model Intercomparison Project (CMIP), that includes intercomparison projects for ocean physical and biogeochemical models[20,171]. The latest round, CMIP6, corresponds to the 6th Assessment Report of the IPCC and includes an intercomparison of the ocean biogeochemical components of ESMs[148,174,175]. Coordinated efforts to compare multiple OBMs in a systematic fashion date back to the mid 1990s and early 2000s, with the first two phases of the Ocean Carbon Model Intercomparison Project (OCMIP-1 and OCMIP2)[173]. The OCMIP team created common frameworks for ocean tracer and biogeochemical experiments. These include natural and anthropogenic radiocarbon; chlorofluorocarbons; and abiotic and biotic carbon and nutrient cycling. Experiment packages included a specified atmospheric boundary condition for trace gases, such as chlorofluorocarbons and CO_2, ocean biogeochemical parameterizations and standardized model diagnostics and output. Individual groups implemented the OCMIP experimental protocols in different ocean physical general circulation models with different physical forcing and ocean circulation[112]. The resulting range of simulated ocean tracer and biogeochemical fields could then be related back to differences in the underlying physics, rather than differences in the biogeochemical components[176], and model skill could be evaluated against a common set of observed metrics[177].

Intercomparison is also well established for regional models, for example within the Integrated Ocean Observing System's Coastal Ocean Modeling Testbed (IOOS COMT)[178]. As part of the COMT, different physical models for simulation and prediction of coastal hypoxia in the Mississippi River outflow region were compared using the same simple oxygen model[179]. The aim was to distinguish model to model differences due to model physics from those resulting from different biological model formulations[180]. This approach is similar to the OCMIP protocols. An alternate, but equally useful, approach is to compare different biogeochemical parameterizations within the same ocean physical model. This is particularly effective when including parameter optimization to minimize uncertainties in parameter choices, focusing the intercomparison on structural model differences[45,53,54].

OBM intercomparison activities rely on available, open-access, high-quality ocean data sets for model development and evaluation[20,89]. In many cases, gridded data products and climatologies are more desirable than

compilations of raw ocean data profiles. The construction of data products that span many individual data collectors and time frames requires substantial effort by domain experts from oceanography, data analysis and data management. The visibility of ocean data management as a distinct and critical element of ocean research has risen substantially over the past couple of decades. Multiple aspects, from informatics and cyber infrastructure to research culture, create incentives for scientists to share data through recognized data repositories and to properly cite data products[181,182]. Current community ocean data efforts revolve around the emergence of new standards for data to be findable, accessible, interoperable and reusable (FAIR)[93]. Data system elements can include an agreement on common sampling and analytical measurement methods, data quality control protocols and assurance checks, laboratory and field intercomparison activities, analytical standards and reference materials, data ontologies and vocabularies, data reporting requirements and formatting standards, and support for data accessibility and archiving through data repositories.

The physical hydrography community has a long history of building unified data products for temperature, salinity, nutrients and oxygen[183]. Standard products, such as the World Ocean Atlas 2018, are created by the National Oceanic and Atmospheric Administration (NOAA) and other ocean data centres[184]. The ocean chemical tracer and biogeochemical community has pursued similar efforts to compile and, where feasible, grid water column data relevant to OBMs. The Global Ocean Data Analysis Project (GLODAP), for example, compiled data on ocean circulation tracers — radiocarbon and chlorofluorocarbon — and dissolved inorganic carbon system variables, including alkalinity, based on historical global ocean surveys dating back to the 1970s (REF.[185]). The GLODAP product includes objectively analysed, gridded spatial maps for both observed properties and derived products, such as anthropogenic CO_2. The GLODAP data product is a living data set, most recently updated as GLODAPv2.2020 (REF.[186]). Many of the large-scale, international ocean observing efforts are built on collaborations to generate similar publicly accessible, standardized data products that are routinely updated with new field observations. Relevant examples for OBM efforts include the GO-SHIP hydrographic programme[187], surface ocean CO_2 observations[188], time series[189], plankton products[190,191] and quality-controlled BGC-Argo products[192,193].

Limitations and optimizations

Application of OBMs is subject to computational limitations, for example limits in CPU time and disk space. This requires compromises in spatial and process resolution achievable in the required domain size, integration time and model ensemble size. For example, the resolution of most ESMs is too coarse to capture ocean mesoscale features, although some modelling groups are exploring global ocean mesoscale plankton simulations[194]. Given current computing resources, increasing the spatial resolution of a global model with century integration times is typically only possible by drastically reducing the number of biogeochemical state variables[26]. Regional models are affordable at much higher spatial resolution, but require boundary conditions from larger-scale models and much shorter integration times. Practical workarounds include approaches for upscaling and downscaling[195]; nested domains, including two-way coupling, where information flows from the large to small scale and vice versa[196]; and adaptive grids or model structures, although these model structures are difficult to implement and rarely used.

Uncertainties in OBM parameters and structures are a major issue limiting the models' predictive skill. Many model parameters cannot be determined experimentally and have relatively large plausible ranges. Parameter optimization approaches should, theoretically, be able to address this as an inverse problem, where information contained in biogeochemical observations is used to infer the underlying parameters. However, in practice, observations are often limited by resolution and breadth of variable types, meaning many model parameters are unconstrained by optimization[44,50]. The issue of undetermined parameters worsens with increasing model complexity, for example when increasing the number of state variables. This is because the number of poorly known parameters multiplies and the degrees of freedom increase[27]. The strong non-linearities characterizing OBM equations also contribute to the challenges. One manifestation of underdetermination is a cancellation of errors, where the model seemingly agrees with available observations, but does so because underlying errors compensate for each other. This includes the possibility of the biogeochemical model compensating for shortcomings in the physical model component. This is problematic for future climate projections, where compensating errors give a plausible simulated present, but limited confidence in future behaviour[197]. Whenever different errors compensate for each other, the OBM might perform acceptably for the tuned observational period, but not outside it. This limits both the mechanistic insights that can be gained from the model and its use as a predictive tool. Slow climate-biogeochemical feedback, for example, is difficult to probe with current observations. A prudent approach is to apply Occam's razor, for example by limiting the number of biogeochemical state variables to those necessary and focusing on processes which have a conceptual or theoretical understanding of climate sensitivity. In practice, this depends on the scientific or practical challenges being addressed and can be subjective.

Related to the underdetermination problem is the issue of structural uncertainties in OBMs. Equations governing biogeochemical state variables are empirical and, except for mass conservation, are not derived from fundamental laws and first principles such as the Navier–Stokes equations. As a result, there are no universally agreed upon parameterizations or optimal model structures. Although most models have converged on a similar, intermediate-complexity structure, the model structures and predictive capabilities are not rigorously tested. Until recently, observational data sets have been too limited in space, time or breadth, for a thorough OBM validation. The rise of global autonomous biogeochemical observation networks[91,191] is beginning to

alleviate this problem and will likely prove transformative for further development of OBMs.

A limitation to the use of ESMs for projecting future conditions is that projections cannot be validated. The use of past climatic changes is one possibility, but has the drawback of observational constraints being limited to palaeo proxies. Additionally, there is no close analogue to the current global warming in the recent geological past, with the current continental plate configuration. Assessment of ESMs for present-day conditions[38,174] provides information about model biases, but does not necessarily describe their behaviour outside the observed conditions. Again, the issues of underdetermination and compensating errors lead to large and poorly quantified uncertainties. In numerical weather prediction it is generally true that an ensemble of different models represents a better forecast than individual models because model uncertainties partly cancel. This notion has been adopted in climate projections where ensembles of ESMs are commonly used. However, evaluation of CMIP5 and CMIP6 ensembles for present-day conditions shows that the assumption of cancellation of errors does not generally hold[38]. This is unsurprising when considering that many models share similar shortcomings and biases. A recent development is the use of explainable and empirical inter-model relationships between characteristics of the present-day conditions and long-term climate projections, the emergent constraint approach[198]. It has enabled a reduction in projection uncertainties in ESMs and has been applied successfully to marine biogeochemistry[199,200].

Another major challenge when using OBMs for future projections is that a fixed biogeochemical structure may not account for functional changes in biological communities due to acclimation and adaptation to new environmental conditions. One potential workaround is to enable emergent communities[22]. However, this is computationally costly and not feasible in the integration times required for future projections with current computational capabilities. Another approach, which requires ongoing input of well-resolved biogeochemical observations, is to allow for adaptive model parameters[55,56]. This is potentially feasible for nowcasts but not projections. Whether the current status quo prevails or new approaches are adopted, a continuous evaluation of OBM results against a well-resolved, broad suite of observations from a sustained global ocean observing system will be paramount as climate-related changes manifest.

Outlook

Anthropogenic perturbations of the global carbon and nitrogen cycles are leading to ocean warming, acidification, deoxygenation and coastal eutrophication. These place stresses on ocean ecosystems and compound direct human impacts, such as overfishing or trawling. Prevention, mitigation and adaptation to the negative effects of these stressors present a formidable challenge. Skilful OBMs that provide robust, accurate and actionable information are key to developing an appropriate response, to enable optimal outcomes. This will require a rigorous quantitative assessment and validation of OBMs, including their structural uncertainties, model refinement and method development. Ocean modelling already benefits from computational advances in traditional ocean and climate models, alongside new approaches using machine learning and other advanced techniques. The OBM simulation tasks further depend on continued expansion of the global ocean observing system into more biogeochemical and ecosystem variables, taking advantage of cost-effective, autonomous platforms and sensors, plus remote sensing.

Given an expanded and sustained biogeochemical observing system, the development of well-constrained, operational OBMs will be feasible in the near term. They will benefit a broad range of scientific purposes and practical applications. The need for accurate information in some specific applications will require development of tailored OBM applications. This includes ocean CO_2 removal and carbon accounting; coastal eutrophication; fisheries; marine diseases; and harmful algal blooms. New applications focused on vulnerability and impacts will drive integration of the current OBM generation with specialized application models and more direct engagement with stakeholders. Demand may grow for near-term to seasonal, to multi-year forecasting products. A continued commitment to open-source code and open-access principles in dissemination of OBM results and derived products should be prioritized by the research community and funders.

Code availability

Example code for a nutrient–phytoplankton–zooplankton–detritus (NPZD) model, parameter optimization and state estimation can be found in REFS.[57,81,201], respectively.

Published online: 22 September 2022

1. Riley, G. A. Factors controlling phytoplankton population on George's Bank. *J. Mar. Res.* **6**, 54–73 (1946).
2. Evans, G. T. & Parslow, J. S. A model of annual plankton cycles. *Biol. Oceanogr.* **3**, 327–347 (1985).
3. Fasham, M. J. R., Ducklow, H. W. & McKelvie, S. M. A nitrogen-based model of plankton dynamics in the oceanic mixed layer. *J. Mar. Res.* **48**, 591–639 (1990).
This work is a seminal early example of an OBM applied to time-series data.
4. Franks, P. J. S., Wroblewski, J. S. & Flierl, G. R. Behavior of a simple plankton model with food-level acclimation by herbivores. *Mar. Biol.* **91**, 121–129 (1986).
5. Sarmiento, J. L. et al. A seasonal three-dimensional ecosystem model of nitrogen cycling in the North Atlantic Euphotic Zone. *Glob. Biogeochem. Cycles* **7**, 417–450 (1993).
This regional model of the North Atlantic is probably the first true OBM, that is, an ocean circulation model with explicit representation of plankton dynamics.
6. Revelle, R. & Suess, H. E. Carbon dioxide exchange between atmosphere and ocean and the question of an increase of atmospheric CO_2 during the past decades. *Tellus* **9**, 18–27 (1957).
7. Sarmiento, J. L. & Toggweiler, J. R. A new model for the role of the oceans in determining atmospheric pCO_2. *Nature* **308**, 621–624 (1984).
8. Siegenthaler, U. & Wenk, T. Rapid atmospheric CO_2 variations and ocean circulation. *Nature* **308**, 624–626 (1984).
9. Maier-Reimer, E. & Hasselmann, K. Transport and storage of CO_2 in the ocean — an inorganic ocean-circulation carbon cycle model. *Clim. Dyn.* **2**, 63–90 (1987).
10. Maier-Reimer, E. Geochemical cycles in an ocean general circulation model. Preindustrial tracer distributions. *Glob. Biogeochem. Cycles* **7**, 645–677 (1993).
This seminal paper describes one of the first marine biogeochemical models of the global ocean.
11. Six, K. D. & Maier-Reimer, E. Effects of plankton dynamics on seasonal carbon fluxes in an ocean general circulation model. *Glob. Biogeochem. Cycles* **10**, 559–583 (1996).
12. Sarmiento, J. L. & Gruber, N. *Ocean Biogeochemical Dynamics* (Princeton Univ. Press, 2006).
13. Glover, D. M., Jenkins, W. J. & Doney, S. C. *Modeling Methods for Marine Science* (Cambridge Univ. Press, 2011).

14. Franks, P. J. S. NPZ models of plankton dynamics: their construction, coupling to physics, and application. *J. Oceanogr.* **58**, 379–387 (2002).
15. Gentleman, W., Leising, A., Frost, B., Strom, S. & Murray, J. Functional responses for zooplankton feeding on multiple resources: a review of assumptions and biological dynamics. *Deep. Sea Res. Part II Top. Stud. Oceanogr.* **50**, 2847–2875 (2003).
16. Le Quéré, C. et al. Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Glob. Chang. Biol.* **11**, 2016–2040 (2005).
17. Cullen, J. J. Subsurface chlorophyll maximum layers: enduring enigma or mystery solved? *Ann. Rev. Mar. Sci.* **7**, 207–239 (2015).
18. Fennel, K. & Boss, E. Subsurface maxima of phytoplankton and chlorophyll: steady-state solutions from a simple model. *Limnol. Oceanogr.* **48**, 1521–1534 (2003).
19. Geider, R. J., MacIntyre, H. L. & Kana, T. M. Dynamic model of phytoplankton growth and acclimation: responses of the balanced growth rate and the chlorophyll a: carbon ratio to light, nutrient-limitation and temperature. *Mar. Ecol. Prog. Ser.* **148**, 187–200 (1997).
20. Orr, J. C. et al. Biogeochemical protocols and diagnostics for the CMIP6 Ocean Model Intercomparison Project (OMIP). *Geosci. Model. Dev.* **10**, 2169–2199 (2017).
This work presents a framework detailing common protocols for including ocean biogeochemistry and chemical tracers in ESMs.
21. Lam, P. & Kuypers, M. M. M. Microbial nitrogen cycling processes in oxygen minimum zones. *Ann. Rev. Mar. Sci.* **3**, 317–345 (2011).
22. Follows, M. J., Dutkiewicz, S., Grant, S. & Chisholm, S. W. Emergent biogeography of microbial communities in a model ocean. *Science* **315**, 1843–1846 (2007).
This paper is the first to explore competition among a large number of phytoplankton groups within a global ocean model.
23. Dutkiewicz, S. et al. Dimensions of marine phytoplankton diversity. *Biogeosciences* **17**, 609–634 (2020).
24. Armstrong, R. A. Grazing limitation and nutrient limitation in marine ecosystems: steady state solutions of an ecosystem model with multiple food chains. *Limnol. Oceanogr.* **39**, 597–608 (1994).
25. Banas, N. S. Adding complex trophic interactions to a size-spectral plankton model: emergent diversity patterns and limits on predictability. *Ecol. Modell.* **222**, 2663–2675 (2011).
26. Galbraith, E. D., Gnanadesikan, A., Dunne, J. P. & Hiscock, M. R. Regional impacts of iron–light colimitation in a global biogeochemical model. *Biogeosciences* **7**, 1043–1064 (2010).
27. Denman, K. L. Modelling planktonic ecosystems: parameterizing complexity. *Prog. Oceanogr.* **57**, 429–452 (2003).
28. Haidvogel, D. B. & Beckmann, A. *Numerical Ocean Circulation Modeling* (Imperial College Press, 1999).
29. Haltiner, G. J. & Williams, R. T. *Numerical Prediction and Dynamic Meteorology* (Wiley, 1980).
30. Roache, P. J. *Fundamentals of Computational Fluid Dynamics* (Hermosa, 1998).
31. Foucart, C., Mirabito, C., Haley, P. J. & Lermusiaux, P. F. J. High-order discontinuous Galerkin methods for nonhydrostatic ocean processes with a free surface. *OCEANS 2021: San Diego–Porto* https://doi.org/10.23919/OCEANS44145.2021.9705767 (2021).
32. Schourup-Kristensen, V., Wekerle, C., Wolf-Gladrow, D. A. & Völker, C. Arctic Ocean biogeochemistry in the high resolution FESOM 1.4-REcoM2 model. *Prog. Oceanogr.* **168**, 65–81 (2018).
33. Zang, Z. et al. Spatially varying phytoplankton seasonality on the northwest Atlantic Shelf: a model-based assessment of patterns, drivers, and implications. *ICES J. Mar. Sci.* **78**, 1920–1934 (2021).
34. Brennan, C. E., Blanchard, H. & Fennel, K. Putting temperature and oxygen thresholds of marine animals in context of environmental change: a regional perspective for the Scotian Shelf and Gulf of St. Lawrence. *PLoS ONE* **11**, e0167411 (2016).
35. Claret, M. et al. Rapid coastal deoxygenation due to ocean circulation shift in the northwest Atlantic. *Nat. Clim. Chang.* **8**, 868–872 (2018).
36. Rutherford, K. & Fennel, K. Diagnosing transit times on the northwestern North Atlantic continental shelf. *Ocean. Sci.* **14**, 1207–1221 (2018).
37. Bourgeois, T. et al. Coastal-ocean uptake of anthropogenic carbon. *Biogeosciences* **13**, 4167–4185 (2016).
38. Laurent, A., Fennel, K. & Kuhn, A. An observation-based evaluation and ranking of historical Earth system model simulations in the northwest North Atlantic Ocean. *Biogeosciences* **18**, 1803–1822 (2021).
39. Rutherford, K. & Fennel, K. Elucidating coastal ocean carbon transport processes: a novel approach applied to the northwest North Atlantic Shelf. *Geophys. Res. Lett.* **49**, e2021GL097614 (2022).
40. Saba, V. S. et al. Enhanced warming of the northwest Atlantic Ocean under climate change. *J. Geophys. Res. Ocean.* **121**, 118–132 (2016).
41. Sweeney, C. et al. Impacts of shortwave penetration depth on large-scale ocean circulation and heat transport. *J. Phys. Oceanogr.* **35**, 1103–1119 (2005).
42. Bonan, G. B. & Doney, S. C. Climate, ecosystems, and planetary futures: the challenge to predict life in Earth system models. *Science* **359**, eaam8328 (2018).
43. Matear, R. J. Parameter optimization and analysis of ecosystem models using simulated annealing: a case study at Station P. *J. Mar. Res.* **53**, 571–607 (1995).
44. Fennel, K., Losch, M., Schroter, J. & Wenzel, M. Testing a marine ecosystem model: sensitivity analysis and parameter optimization. *J. Mar. Syst.* **28**, 45–63 (2001).
45. Friedrichs, M. A. M. et al. Assessment of skill and portability in regional marine biogeochemical models: role of multiple planktonic groups. *J. Geophys. Res.* **112**, 1–22 (2007).
46. Mattern, J. P. & Edwards, C. A. Simple parameter estimation for complex models — testing evolutionary techniques on 3-dimensional biogeochemical ocean models. *J. Mar. Syst.* **165**, 139–152 (2017).
47. Laurent, A., Fennel, K., Wilson, R., Lehrter, J. & Devereux, R. Parameterization of biogeochemical sediment–water fluxes using in situ measurements and a diagenetic model. *Biogeosciences* **13**, 77–94 (2016).
48. Wilson, R. F., Fennel, K. & Paul Mattern, J. Simulating sediment–water exchange of nutrients and oxygen: a comparative assessment of models against mesocosm observations. *Cont. Shelf Res.* **63**, 69–84 (2013).
49. Thacker, W. C. The role of the Hessian matrix in fitting models to measurements. *J. Geophys. Res. Ocean.* **94**, 6177–6196 (1989).
50. Ward, B. A., Friedrichs, M. A. M., Anderson, T. R. & Oschlies, A. Parameter optimisation techniques and the problem of underdetermination in marine biogeochemical models. *J. Mar. Syst.* **81**, 34–43 (2010).
51. Schartau, M. et al. Reviews and syntheses: parameter identification in marine planktonic ecosystem modelling. *Biogeosciences* **14**, 1647–1701 (2017).
52. Gregg, W. W. et al. Skill assessment in ocean biological data assimilation. *J. Mar. Syst.* **76**, 16–33 (2009).
53. Bagniewski, W., Fennel, K., Perry, M. J. & D'Asaro, E. A. Optimizing models of the North Atlantic spring bloom using physical, chemical and bio-optical observations from a Lagrangian float. *Biogeosciences* **8**, 1291–1307 (2011).
54. Kuhn, A. M., Fennel, K. & Berman-frank, I. Modelling the biogeochemical effects of heterotrophic and autotrophic N_2 fixation in the Gulf of Aqaba (Israel), Red Sea. *Biogeosciences* **15**, 7379–7401 (2018).
55. Mattern, J. P., Fennel, K. & Dowd, M. Periodic time-dependent parameters improving forecasting abilities of biological ocean models. *Geophys. Res. Lett.* **41**, 6848–6854 (2014).
56. Kitagawa, G. A self-organizing state-space model. *J. Am. Stat. Assoc.* **93**, 1203–1215 (1998).
57. Mattern, J. P. Visualizing parameter and state estimation for a zero-dimensional ocean biological model. *GitHub* https://doi.org/10.5281/zenodo.6994739 (2022).
58. Evensen, G. The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean. Dyn.* **53**, 343–367 (2003).
This influential paper proposes the now widely used EnKF.
59. Kalman, R. E. A new approach to linear filtering and prediction problems. *J. Basic Eng.* **82**, 35–45 (1960).
60. Humpherys, J., Redd, P. & West, J. A fresh look at the Kalman filter. *SIAM Rev.* **54**, 801–823 (2012).
61. Jazwinski, A. R. *Stochastic Processes and Filtering Theory* (Academic, 1970).
62. Pham, D. T., Verron, J. & Roubaud, M. C. A singular evolutive extended Kalman filter for data assimilation in oceanography. *J. Mar. Syst.* **16**, 323–340 (1998).
63. van Leeuwen, P. J. A consistent interpretation of the stochastic version of the ensemble Kalman filter. *Q. J. R. Meteorol. Soc.* **146**, 2815–2825 (2020).
64. Yu, L. et al. Insights on multivariate updates of physical and biogeochemical ocean variables using an ensemble Kalman filter and an idealized model of upwelling. *Ocean. Model.* **126**, 13–28 (2018).
65. Yu, L. et al. Evaluation of nonidentical versus identical twin approaches for observation impact assessments: an ensemble-Kalman-filter-based ocean assimilation application for the Gulf of Mexico. *Ocean. Sci.* **15**, 1801–1814 (2019).
66. Wang, B., Fennel, K. & Yu, L. Can assimilation of satellite observations improve subsurface biological properties in a numerical model? A case study for the Gulf of Mexico. *Ocean. Sci.* **17**, 1141–1156 (2021).
67. Sakov, P. & Oke, P. R. A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters. *Tellus A Dyn. Meteorol. Oceanogr.* **60**, 361–371 (2008).
68. Houtekamer, P. L. & Zhang, F. Review of the ensemble Kalman filter for atmospheric data assimilation. *Mon. Weather. Rev.* **144**, 4489–4532 (2016).
69. Mattern, J. P., Song, H., Edwards, C. A., Moore, A. M. & Fiechter, J. Data assimilation of physical and chlorophyll a observations in the California current system using two biogeochemical models. *Ocean. Model.* **109**, 55–71 (2017).
70. Wang, B., Fennel, K., Yu, L. & Gordon, C. Assessing the value of biogeochemical Argo profiles versus ocean color observations for biogeochemical model optimization in the Gulf of Mexico. *Biogeosciences* **17**, 4059–4074 (2020).
71. Fiechter, J., Broquet, G., Moore, A. M. & Arango, H. G. A data assimilative, coupled physical–biological model for the Coastal Gulf of Alaska. *Dyn. Atmos. Ocean.* **52**, 95–118 (2011).
72. Moore, A. M. et al. The regional ocean modeling system (ROMS) 4-dimensional variational data assimilation systems: part III — observation impact and observation sensitivity in the California Current System. *Prog. Oceanogr.* **91**, 74–94 (2011).
73. Fennel, K. et al. Advancing marine biogeochemical and ecosystem reanalyses and forecasts as tools for monitoring and managing ecosystem health. *Front. Mar. Sci.* **6**, 89 (2019).
74. Teruzzi, A., Bolzon, G., Salon, S., Lazzari, P. & Solidoro, C. Assimilation of coastal and open sea biogeochemical data to improve phytoplankton simulation in the Mediterranean Sea. *Ocean. Model.* **132**, 46–60 (2018).
75. Cossarini, G. et al. Towards operational 3D-Var assimilation of chlorophyll biogeochemical-Argo float data into a biogeochemical model of the Mediterranean Sea. *Ocean. Model.* **133**, 112–128 (2019).
76. Ford, D. Assimilating synthetic biogeochemical-Argo and ocean colour observations into a global ocean model to inform observing system design. *Biogeosciences* **18**, 509–534 (2021).
77. Song, H., Edwards, C. A., Moore, A. M. & Fiechter, J. Data assimilation in a coupled physical–biogeochemical model of the California current system using an incremental lognormal 4-dimensional variational approach: part 3 — assimilation in a realistic context using satellite and in situ observations. *Ocean. Model.* **106**, 159–172 (2016).
78. Courtier, P., Thépaut, J.-N. & Hollingsworth, A. A strategy for operational implementation of 4D-Var, using an incremental approach. *Q. J. R. Meteorol. Soc.* **120**, 1367–1387 (1994).
79. Gordon, N. J., Salmond, D. J. & Smith, A. F. M. in *IEE Proc. F-radar and Signal Processing* Vol. 140 107–113 (IET Digital Library, 1993).
80. Mattern, J. P., Dowd, M. & Fennel, K. Particle filter-based data assimilation for a three-dimensional biological ocean model and satellite observations. *J. Geophys. Res. Ocean.* **118**, 2746–2760 (2013).
81. Mattern, J. P., Yu, L., Wang, B. & Fennel, K. Ensemble Kalman filter application for an ocean biogeochemical model in an idealized 3-dimensional channel. *GitHub* https://doi.org/10.5281/zenodo.6974184 (2022).
82. Rothstein, L. M. et al. Modeling ocean ecosystems: the PARADIGM program. *Oceanography* **19**, 22–51 (2006).
83. Lehmann, M. K., Fennel, K. & He, R. Statistical validation of a 3-D bio-physical model of the western North Atlantic. *Biogeosciences* **6**, 1961–1974 (2009).
84. Taylor, K. E. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res. Atmos.* **106**, 7183–7192 (2001).

85. Jolliff, J. K. et al. Summary diagrams for coupled hydrodynamic–ecosystem model skill assessment. *J. Mar. Syst.* **76**, 64–82 (2009).
86. Stow, C. A. et al. Skill assessment for coupled biological/physical models of marine systems. *J. Mar. Syst.* **76**, 4–15 (2009).
This paper presents a tutorial on common statistical approaches to model-data skill assessment for OBMs.
87. Doney, S. C. et al. Skill metrics for confronting global upper ocean ecosystem–biogeochemistry models against field and remote sensing data. *J. Mar. Syst.* **76**, 95–112 (2009).
88. Mattern, J. P., Fennel, K. & Dowd, M. Introduction and assessment of measures for quantitative model-data comparison using satellite images. *Remote. Sens.* **2**, 794–818 (2010).
89. Capotondi, A. et al. Observational needs supporting marine ecosystems modeling and forecasting: from the global ocean to regional and coastal systems. *Front. Mar. Sci.* https://doi.org/10.3389/fmars.2019.00623 (2019).
90. Roemmich, D. et al. On the future of Argo: a global, full-depth, multi-disciplinary array. *Front. Mar. Sci.* **6**, 439 (2019).
91. Chai, F. et al. Monitoring ocean biogeochemistry with autonomous platforms. *Nat. Rev. Earth Environ.* **1**, 315–326 (2020).
This work reviews autonomous approaches to measuring ocean biogeochemical properties, which will likely prove transformative for OBM validation and assimilation.
92. Johnson, K. S. et al. Biogeochemical sensor performance in the SOCCOM profiling float array. *J. Geophys. Res. Ocean.* **122**, 6416–6436 (2017).
93. Tanhua, T. et al. Ocean FAIR data services. *Front. Mar. Sci.* **6**, 440 (2019).
94. Révelard, A. et al. Ocean integration: the needs and challenges of effective coordination within the ocean observing system. *Front. Mar. Sci.* https://doi.org/10.3389/fmars.2021.737671 (2022).
95. Friedlingstein, P. et al. Global Carbon Budget 2021. *Earth Syst. Sci. Data* **14**, 1917–2005 (2022).
96. Khatiwala, S. et al. Global ocean storage of anthropogenic carbon. *Biogeosciences* **10**, 2169–2191 (2013).
97. IPCC. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge Univ. Press, 2021).
98. Hauck, J. et al. Consistency and challenges in the ocean carbon sink estimate for the global carbon budget. *Front. Mar. Sci.* **7**, 571720 (2020).
99. Crisp, D. et al. How well do we understand the land–ocean–atmosphere carbon cycle? *Rev. Geophys.* **60**, e2021RG000736 (2022).
100. Ilyina, T. et al. Predictable variations of the carbon sinks and atmospheric CO_2 growth in a multi-model framework. *Geophys. Res. Lett.* **48**, e2020GL090695 (2021).
101. Gattuso, J.-P. et al. Ocean solutions to address climate change and its effects on marine ecosystems. *Front. Mar. Sci.* https://doi.org/10.3389/fmars.2018.00337 (2018).
102. National Academies of Sciences, Engineering, and Medicine. *A Research Strategy for Ocean-based Carbon Dioxide Removal and Sequestration* (National Academies, 2022).
103. Aumont, O. & Bopp, L. Globalizing results from ocean in situ iron fertilization studies. *Glob. Biogeochem. Cycles* https://doi.org/10.1029/2005GB002591 (2006).
104. Jin, X., Gruber, N., Frenzel, H., Doney, S. C. & McWilliams, J. C. The impact on atmospheric CO_2 of iron fertilization induced changes in the ocean's biological pump. *Biogeosciences* **5**, 385–406 (2008).
105. Oschlies, A., Koeve, W., Rickels, W. & Rehdanz, K. Side effects and accounting aspects of hypothetical large-scale Southern Ocean iron fertilization. *Biogeosciences* **7**, 4017–4035 (2010).
106. Dutreuil, S., Bopp, L. & Tagliabue, A. Impact of enhanced vertical mixing on marine biogeochemistry: lessons for geo-engineering and natural variability. *Biogeosciences* **6**, 901–912 (2009).
107. Bach, L. T. et al. Testing the climate intervention potential of ocean afforestation using the Great Atlantic Sargassum Belt. *Nat. Commun.* **12**, 2556 (2021).
108. Ilyina, T., Wolf-Gladrow, D., Munhoven, G. & Heinze, C. Assessing the potential of calcium-based artificial ocean alkalinization to mitigate rising atmospheric CO_2 and ocean acidification. *Geophys. Res. Lett.* **40**, 5909–5914 (2013).
109. Feng, E. Y., Koeve, W., Keller, D. P. & Oschlies, A. Model-based assessment of the CO_2 sequestration potential of coastal ocean alkalinization. *Earth's Futur.* **5**, 1252–1266 (2017).
110. Siegel, D. A., DeVries, T., Doney, S. C. & Bell, T. Assessing the sequestration time scales of some ocean-based carbon dioxide reduction strategies. *Environ. Res. Lett.* **16**, 104003 (2021).
111. Schmidtko, S., Stramma, L. & Visbeck, M. Decline in global oceanic oxygen content during the past five decades. *Nature* **542**, 335–339 (2017).
112. Doney, S. C., Bopp, L. & Long, M. C. Historical and future trends in ocean climate and biogeochemistry. *Oceanography* **27**, 108–119 (2014).
113. Bopp, L., Resplandy, L., Untersee, A., Le Mezo, P. & Kageyama, M. Ocean (de)oxygenation from the Last Glacial Maximum to the twenty-first century: insights from Earth system models. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **375**, 20160323 (2017).
114. Takano, Y., Ito, T. & Deutsch, C. Projected centennial oxygen trends and their attribution to distinct ocean climate forcings. *Glob. Biogeochem. Cycles* **32**, 1329–1349 (2018).
115. Levin, L. A. Manifestation, drivers, and emergence of open ocean deoxygenation. *Ann. Rev. Mar. Sci.* **10**, 229–260 (2018).
116. Oschlies, A., Brandt, P., Stramma, L. & Schmidtko, S. Drivers and mechanisms of ocean deoxygenation. *Nat. Geosci.* **11**, 467–473 (2018).
117. Breitburg, D. et al. Declining oxygen in the global ocean and coastal waters. *Science* **359**, eaam7240 (2018).
118. Rabalais, N. N. et al. Eutrophication-driven deoxygenation in the coastal ocean. *Oceanography* **27**, 172–183 (2014).
119. Andrews, O., Buitenhuis, E., Le Quéré, C. & Suntharalingam, P. Biogeochemical modelling of dissolved oxygen in a changing ocean. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **375**, 20160328 (2017).
120. Cocco, V. et al. Oxygen and indicators of stress for marine life in multi-model global warming projections. *Biogeosciences* **10**, 1849–1868 (2013).
121. Bopp, L. et al. Multiple stressors of ocean ecosystems in the 21st century: projections with CMIP5 models. *Biogeosciences* **10**, 6225–6245 (2013).
122. Couespel, D., Lévy, M. & Bopp, L. Oceanic primary production decline halved in eddy-resolving simulations of global warming. *Biogeosciences* **18**, 4321–4349 (2021).
123. Bahl, A., Gnanadesikan, A. & Pradal, M.-A. Variations in ocean deoxygenation across earth system models: isolating the role of parameterized lateral mixing. *Glob. Biogeochem. Cycles* **33**, 703–724 (2019).
124. Lévy, M., Resplandy, L., Palter, J. B., Couespel, D. & Lachkar, Z. in *Ocean Mixing* Ch. 13 (eds Meredith, M. & Naveira Garabato, A. B. T.-O. M.) 329–344 (Elsevier, 2022).
125. Fennel, K. & Testa, J. M. Biogeochemical controls on coastal hypoxia. *Ann. Rev. Mar. Sci.* **11**, 105–130 (2019).
This review of coastal hypoxia puts forward a simple non-dimensional number to elucidate key factors controlling hypoxia in diverse coastal systems.
126. Peña, M. A., Katsev, S., Oguz, T. & Gilbert, D. Modeling dissolved oxygen dynamics and hypoxia. *Biogeosciences* **7**, 933–957 (2010).
127. Irby, I. D. et al. Challenges associated with modeling low-oxygen waters in Chesapeake Bay: a multiple model comparison. *Biogeosciences* **13**, 2011–2028 (2016).
128. Zhang, H., Fennel, K., Laurent, A. & Bian, C. A numerical model study of the main factors contributing to hypoxia and its interannual and short-term variability in the East China Sea. *Biogeosciences* **17**, 5745–5761 (2020).
129. Li, Y., Li, M. & Kemp, W. M. A budget analysis of bottom-water dissolved oxygen in Chesapeake Bay. *Estuaries Coasts* **38**, 2132–2148 (2015).
130. Yu, L., Fennel, K., Laurent, A., Murrell, M. C. & Lehrter, J. C. Numerical analysis of the primary processes controlling oxygen dynamics on the Louisiana shelf. *Biogeosciences* **12**, 2063–2076 (2015).
131. Laurent, A., Fennel, K., Ko, D. & Lehrter, J. Climate change projected to exacerbate impacts of coastal eutrophication in the northern Gulf of Mexico. *J. Geophys. Res. Ocean.* **123**, (2018).
132. Ni, W., Li, M., Ross, A. C. & Najjar, R. G. Large projected decline in dissolved oxygen in a eutrophic estuary due to climate change. *J. Geophys. Res. Ocean.* **124**, 8271–8289 (2019).
133. LaBone, E. D., Rose, K. A., Justic, D., Huang, H. & Wang, L. Effects of spatial variability on the exposure of fish to hypoxia: a modeling analysis for the Gulf of Mexico. *Biogeosciences* **18**, 487–507 (2021).
134. de Mutsert, K., Steenbeek, J., Cowan, J. H. & Christensen, V. in *Modeling Coastal Hypoxia* (eds. Justic, D. et al.) 377–400 (Springer International, 2017).
135. Fennel, K. & Laurent, A. N and P as ultimate and proximate limiting nutrients in the northern Gulf of Mexico: implications for hypoxia reduction strategies. *Biogeosciences* **15**, 3121–3131 (2018).
136. Saraiva, S. et al. Baltic Sea ecosystem response to various nutrient load scenarios in present and future climates. *Clim. Dyn.* **52**, 3369–3387 (2019).
137. Irby, I. D., Friedrichs, M. A. M., Da, F. & Hinson, K. E. The competing impacts of climate change and nutrient reductions on dissolved oxygen in Chesapeake Bay. *Biogeosciences* **15**, 2649–2668 (2018).
138. Kessouri, F. et al. Coastal eutrophication drives acidification, oxygen loss, and ecosystem change in a major oceanic upwelling system. *Proc. Natl Acad. Sci. USA* **118**, e2018856118 (2021).
139. Laurent, A. & Fennel, K. Time-evolving, spatially explicit forecasts of the northern Gulf of Mexico Hypoxic Zone. *Environ. Sci. Technol.* **53**, 14449–14458 (2019).
140. Matli, V. R. R. et al. Fusion-based hypoxia estimates: combining geostatistical and mechanistic models of dissolved oxygen variability. *Environ. Sci. Technol.* **54**, 13016–13025 (2020).
141. Yu, L. & Gan, J. Mitigation of eutrophication and hypoxia through oyster aquaculture: an ecosystem model evaluation off the Pearl River Estuary. *Environ. Sci. Technol.* **55**, 5506–5514 (2021).
142. Feely, R. A., Doney, S. C. & Cooley, S. R. Ocean acidification: present conditions and future changes in a high-CO_2 world. *Oceanography* **22**, 36–47 (2009).
143. Licker, R. et al. Attributing ocean acidification to major carbon producers. *Environ. Res. Lett.* **14**, 124060 (2019).
144. Doney, S. C., Busch, D. S., Cooley, S. R. & Kroeker, K. J. The impacts of ocean acidification on marine ecosystems and reliant human communities. *Annu. Rev. Environ. Resour.* **45**, 83–112 (2020).
145. Gehlen, M. et al. The fate of pelagic $CaCO_3$ production in a high CO_2 ocean: a model study. *Biogeosciences* **4**, 505–519 (2007).
146. Ilyina, T., Zeebe, R. E., Maier-Reimer, E. & Heinze, C. Early detection of ocean acidification effects on marine calcification. *Glob. Biogeochem. Cycles* https://doi.org/10.1029/2008GB003278 (2009).
147. Krumhardt, K. M. et al. Coccolithophore growth and calcification in an acidified ocean: insights from community earth system model simulations. *J. Adv. Model. Earth Syst.* **11**, 1418–1437 (2019).
148. Kwiatkowski, L. et al. Twenty-first century ocean warming, acidification, deoxygenation, and upper-ocean nutrient and primary production decline from CMIP6 model projections. *Biogeosciences* **17**, 3439–3470 (2020).
This work assesses the projected evolution of ocean biogeochemistry under twenty-first-century climate change across a suite of ESMs.
149. Brady, R. X., Lovenduski, N. S., Yeager, S. G., Long, M. C. & Lindsay, K. Skillful multiyear predictions of ocean acidification in the California Current System. *Nat. Commun.* **11**, 2166 (2020).
150. Laurent, A. et al. Eutrophication-induced acidification of coastal waters in the northern Gulf of Mexico: insights into origin and processes from a coupled physical–biogeochemical model. *Geophys. Res. Lett.* **44**, 946–956 (2017).
151. Hauri, C. et al. A regional hindcast model simulating ecosystem dynamics, inorganic carbon chemistry, and ocean acidification in the Gulf of Alaska. *Biogeosciences* **17**, 3837–3857 (2020).
152. Rutherford, K., Fennel, K., Atamanchuk, D., Wallace, D. & Thomas, H. A modelling study of temporal and spatial pCO_2 variability on the biologically active and temperature-dominated Scotian Shelf. *Biogeosciences* **18**, 6271–6286 (2021).
153. Hauri, C. et al. Spatiotemporal variability and long-term trends of ocean acidification in the California Current System. *Biogeosciences* **10**, 193–216 (2013).
154. Hauri, C. et al. Modulation of ocean acidification by decadal climate variability in the Gulf of Alaska. *Commun. Earth Environ.* **2**, 191 (2021).

155. Gruber, N., Boyd, P. W., Frölicher, T. L. & Vogt, M. Biogeochemical extremes and compound events in the ocean. *Nature* **600**, 395–407 (2021).
156. Dutkiewicz, S. et al. Impact of ocean acidification on the structure of future phytoplankton communities. *Nat. Clim. Chang.* **5**, 1002–1006 (2015).
157. Pauly, D. & Christensen, V. Primary production required to sustain global fisheries. *Nature* **374**, 255–257 (1995).
158. Loukos, H., Monfray, P., Bopp, L. & Lehodey, P. Potential changes in skipjack tuna (*Katsuwonus pelamis*) habitat from a global warming scenario: modelling approach and preliminary results. *Fish. Oceanogr.* **12**, 474–482 (2003).
159. Stock, C. A. et al. On the use of IPCC-class models to assess the impact of climate on living marine resources. *Prog. Oceanogr.* **88**, 1–27 (2011).
160. Tittensor, D. P. et al. A protocol for the intercomparison of marine fishery and ecosystem models: Fish-MIP v1.0. *Geosci. Model. Dev.* **11**, 1421–1442 (2018).
161. Lotze, H. K. et al. Global ensemble projections reveal trophic amplification of ocean biomass declines with climate change. *Proc. Natl Acad. Sci. USA* **116**, 12907–12912 (2019).
162. Tittensor, D. P. et al. Next-generation ensemble projections reveal higher climate risks for marine ecosystems. *Nat. Clim. Chang.* **11**, 973–981 (2021).
163. Cheung, W. W. L. et al. Large-scale redistribution of maximum fisheries catch potential in the global ocean under climate change. *Glob. Chang. Biol.* **16**, 24–35 (2010).
164. Lam, V. W. Y., Cheung, W. W. L., Reygondeau, G. & Sumaila, U. R. Projected change in global fisheries revenues under climate change. *Sci. Rep.* **6**, 32607 (2016).
165. IPCC. *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate* (Cambridge Univ. Press, 2019).
166. IPBES. *Global Assessment Report of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services* (IPBES Secretariat, 2019).
167. Aumont, O., Maury, O., Lefort, S. & Bopp, L. Evaluating the potential impacts of the diurnal vertical migration by marine organisms on marine biogeochemistry. *Glob. Biogeochem. Cycles* **32**, 1622–1643 (2018).
168. Archibald, K. M., Siegel, D. A. & Doney, S. C. Modeling the impact of zooplankton diel vertical migration on the carbon export flux of the biological pump. *Glob. Biogeochem. Cycles* **33**, 181–199 (2019).
169. Arnold, C. P. & Dey, C. H. Observing-systems simulation experiments: past, present, and future. *Bull. Am. Meteorol. Soc.* **67**, 687–695 (1986).
170. Halliwell, G. R. et al. Rigorous evaluation of a fraternal twin ocean OSSE system for the open Gulf of Mexico. *J. Atmos. Ocean. Technol.* **31**, 105–130 (2014).
171. Griffies, S. M. et al. OMIP contribution to CMIP6: experimental and diagnostic protocol for the physical component of the ocean model intercomparison project. *Geosci. Model. Dev.* **9**, 3231–3296 (2016).
172. Chassignet, E. P. et al. DAMÉE-NAB: the base experiments. *Dyn. Atmos. Ocean.* **32**, 155–183 (2000).
173. Orr, J. C. On ocean carbon-cycle model comparison. *Tellus B Chem. Phys. Meteorol.* **51**, 509–510 (1999).
174. Séférian, R. et al. Tracking improvement in simulated marine biogeochemistry between CMIP5 and CMIP6. *Curr. Clim. Chang. Rep.* **6**, 95–119 (2020).
175. Canadell, J. G. et al. in *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge Univ. Press, 2021).
176. Najjar, R. G. et al. Impact of circulation on export production, dissolved organic matter, and dissolved oxygen in the ocean: results from phase II of the Ocean Carbon-cycle Model Intercomparison Project (OCMIP-2). *Global Biogeochem. Cycles* https://doi.org/10.1029/2006GB002857 (2007).
177. Matsumoto, K. et al. Evaluation of ocean carbon cycle models with data-based metrics. *Geophys. Res. Lett.* https://doi.org/10.1029/2003GL018970 (2004).
178. Luettich, R. A. Jr et al. A test bed for coastal and ocean modeling. *Eos* https://doi.org/10.1029/2017EO078243 (2017).
179. Yu, L., Fennel, K. & Laurent, A. A modeling study of physical controls on hypoxia generation in the northern Gulf of Mexico. *J. Geophys. Res. Ocean.* **120**, 5019–5039 (2015).
180. Fennel, K. et al. Effects of model physics on hypoxia simulations for the northern Gulf of Mexico: a model intercomparison. *J. Geophys. Res. Ocean.* **121**, 5731–5750 (2016).
181. Glover, D. M. et al. The US JGOFS data management experience. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **53**, 793–802 (2006).
182. Baker, K. S. & Chandler, C. L. Enabling long-term oceanographic research: changing data practices, information management strategies and informatics. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **55**, 2132–2142 (2008).
183. Boyer, T. et al. Objective analyses of annual, seasonal, and monthly temperature and salinity for the World Ocean on a 0.25° grid. *Int. J. Climatol.* **25**, 931–945 (2005).
184. Garcia, H. E., Boyer, T. P., Baranova, O. K. & Locarnini, R. A. *World Ocean Atlas 2018: Product Documentation* (ed. Mishonov, A.) (NOAA, 2019).
185. Key, R. M. et al. A global ocean carbon climatology: results from Global Data Analysis Project (GLODAP). *Glob. Biogeochem. Cycles* https://doi.org/10.1029/2004GB002247 (2004).
186. Olsen, A. et al. An updated version of the global interior ocean biogeochemical data product, GLODAPv2.2020. *Earth Syst. Sci. Data* **12**, 3653–3678 (2020).
187. Sloyan, B. M. et al. The Global Ocean Ship-based Hydrographic Investigations Program (GO-SHIP): a platform for integrated multidisciplinary ocean science. *Front. Mar. Sci.* **6**, 445 (2019).
188. Wanninkhof, R. et al. A surface ocean CO_2 reference network, SOCONET and associated marine boundary layer CO_2 measurements. *Front. Mar. Sci.* **6**, 400 (2019).
189. Benway, H. M. et al. Ocean time series observations of changing marine ecosystems: an era of integration, synthesis, and societal applications. *Front. Mar. Sci.* **6**, 393 (2019).
190. Buitenhuis, E. T. et al. MAREDAT: towards a world atlas of MARine Ecosystem DATa. *Earth Syst. Sci. Data* **5**, 227–239 (2013).
191. Lombard, F. et al. Globally consistent quantitative observations of planktonic ecosystems. *Front. Mar. Sci.* **6**, 196 (2019).
192. Bittig, H. C. et al. A BGC-Argo guide: planning, deployment, data handling and usage. *Front. Mar. Sci.* https://doi.org/10.3389/fmars.2019.00502 (2019).
193. Maurer, T. L., Plant, J. N. & Johnson, K. S. Delayed-mode quality control of oxygen, nitrate, and pH data on SOCCOM biogeochemical profiling floats. *Front. Mar. Sci.* **8**, 683207 (2021).
194. Harrison, C. S., Long, M. C., Lovenduski, N. S. & Moore, J. K. Mesoscale effects on carbon export: a global perspective. *Glob. Biogeochem. Cycles* **32**, 680–703 (2018).
195. Katavouta, A. & Thompson, K. R. Downscaling ocean conditions with application to the Gulf of Maine, Scotian Shelf and adjacent deep ocean. *Ocean. Model.* **104**, 54–72 (2016).
196. Debreu, L., Marchesiello, P., Penven, P. & Cambon, G. Two-way nesting in split-explicit ocean models: algorithms, implementation and validation. *Ocean. Model.* **49–50**, 1–21 (2012).
197. Löptien, U. & Dietze, H. Reciprocal bias compensation and ensuing uncertainties in model-based climate projections: pelagic biogeochemistry versus ocean mixing. *Biogeosciences* **16**, 1865–1881 (2019).
198. Eyring, V. et al. Taking climate model evaluation to the next level. *Nat. Clim. Chang.* **9**, 102–110 (2019).
199. Kwiatkowski, L. et al. Emergent constraints on projections of declining primary production in the tropical oceans. *Nat. Clim. Chang.* **7**, 355–358 (2017).
200. Terhaar, J., Kwiatkowski, L. & Bopp, L. Emergent constraint on Arctic Ocean acidification in the twenty-first century. *Nature* **582**, 379–383 (2020).
201. Fennel, K. A simple one-dimensional NPZD model with graphical user interface. *GitHub* https://doi.org/10.5281/zenodo.6993508 (2022).
202. Kuhn, A. M., Fennel, K. & Mattern, J. P. Progress in oceanography model investigations of the North Atlantic spring bloom initiation. *Prog. Oceanogr.* **138**, 176–193 (2015).

Acknowledgements

K.F. and B.W. acknowledge support from the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Program (RGPIN-2014-03938), the Canada Foundation for Innovation (Innovation Fund 39902) and the Ocean Frontier Institute. J.P.M. was supported by the Simons Foundation (CBIOMES award ID: 549949). S.C.D. acknowledges support from the US National Science Foundation via the Center for Chemical Currencies of a Microbial Planet (National Science Foundation (NSF) 2019589). L.B. acknowledges support from the European Union's Horizon 2020 research and innovation programmes COMFORT (grant agreement no. 820989) and ESM2025 (grant agreement no. 101003536). L.Y. acknowledges support from the Center for Ocean Research in Hong Kong and Macau.

Author contributions

Introduction (K.F. and L.B.); Experimentation (S.C.D., K.F., J.P.M., A.M.M. and B.W.); Results (K.F. and J.P.M.); Applications (S.C.D., L.Y., L.B., J.P.M. and K.F.); Reproducibility and data deposition (S.C.D. and K.F.); Limitations and optimizations (K.F. and S.C.D.); Outlook (K.F.); Overview of the Primer (K.F.).

Competing interests

The authors declare no competing interests.

Peer review information

Nature Reviews Methods Primers thanks Yvette Spitz, Zhengui Wang, Peng Xiu and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.