# SeapoPym v0.1: Implementation of the SEAPODYM low and mid trophic levels in Python with a flexible optimization framework

Jules Victor Lehodey[1,2,3], Alexandre Mignot[4], Alexandre Ganachaud[1,2], Sarah Albernhe[4], and Simon Nicol[5]

[1]LEGOS, University of Toulouse, CNRS, IRD, CNES, UT, Toulouse, France
[2]MARBEC, University of Montpellier, CNRS, IRD, IFREMER, Sète, France
[3]Université de la Nouvelle-Calédonie, ISEA - Institut des Sciences Exactes et Appliquées (EA 7484), Noumea, New Caledonia
[4]Mercator Ocean International, Toulouse, France
[5]Pacific Community (SPC), Noumea, New Caledonia

**Correspondence:** Jules Victor Lehodey (lehodey.jules@gmail.com)

**Abstract.** SEAPODYM-LMTL, the low and mid trophic level component of SEAPODYM, simulates mesozooplankton and micronekton biomass globally as an advection-diffusion-reaction system driven by physical and biogeochemical forcing. Its mesozooplankton parameterization remains incompletely calibrated, and its operational implementation couples the biological equations to spatial transport, so evaluating a parameter set requires running the full spatial model, which makes automated calibration costly. We present SeapoPym v0.1, an open-source Python re-implementation of the SEAPODYM-LMTL biological model that solves the dynamics locally, without transport. We apply it to the single epipelagic mesozooplankton group and estimate its five biological parameters with a covariance matrix adaptation evolution strategy (CMA-ES). Comparison with the operational product shows that omitting transport matters most in strongly advective regions and in cold high-latitude waters, where the long zooplankton life cycle keeps the biomass exposed to advection. At six contrasting stations the difference between the two models stays between 6 and 12% of the simulated biomass, five to seven times smaller than the model-observation gap where in-situ records allow that comparison. A Sobol analysis attributes the magnitude of the biomass to the energy-transfer and mortality parameters, and the timing of the seasonal peak to the recruitment parameters. Twin experiments then show that parameter identifiability depends on the environmental regime sampled. Wherever the search converged, energy transfer and mortality were recovered, whereas recruitment was recovered only in cold water. A single cold station constrained all five parameters as well as the six stations combined, so recovery follows the information content of the sampled regime rather than the number of stations. These results hold for noise-free synthetic observations generated by the transport-free model itself and driven by the exact forcing. The next step is to repeat them under realistic sampling and forcing error, then calibrate the model against in-situ records.

## 1 Introduction

Mesozooplankton connect primary production to higher trophic levels and contribute to ocean carbon export. They are conventionally defined as planktonic organisms ranging from $0.2$ to $20$ $\mathrm{mm}$ (Sieburth et al., 1978), dominated by copepods, although

the size class is taxonomically diverse (Moriarty and O'Brien, 2013; Drago et al., 2022). They provide a major food source for micronekton (2–20 $\mathrm{cm}$), including mesopelagic fish, euphausiids and gelatinous organisms (Brodeur et al., 2005; St. John et al., 2016). Mesozooplankton also contribute to the transformation and export of organic matter (Steinberg et al., 2000, 2002; Stukel et al., 2022; Siegel et al., 2023). Changes in their biomass or seasonal timing can therefore affect both prey availability and carbon export, making their representation a leading source of uncertainty in global biogeochemical models (Rohr et al., 2023).

The SEAPODYM-LMTL model provides operational global simulations of mesozooplankton and micronekton biomass. It is the Low and Mid-Trophic Levels (LMTL) component of SEAPODYM, a model developed to simulate tuna and tuna-like populations (Lehodey et al., 2008). SEAPODYM-LMTL represents their prey environment through one mesozooplankton and six micronekton functional groups (Lehodey et al., 2010; Conchon, 2016). It couples biological processes with advection and diffusion and is driven by currents, temperature, euphotic depth and net primary production (NPP). The resulting biomass fields are distributed through the Copernicus Marine Service for ecosystem and fisheries applications (Titaud et al., 2024).

The mesozooplankton parameterization of the SEAPODYM-LMTL model remains incompletely calibrated, and its operational product shows substantial differences from observations. Of its five biological parameters, only the energy-transfer coefficient has been formally calibrated against mesozooplankton observations. The four parameters controlling the temperature dependence of recruitment and mortality are derived from the literature or manual tuning (Huntley and Lopez, 1992; Gillooly, 2000; Gillooly et al., 2002; Conchon, 2016; Titaud et al., 2024). When compared to in-situ observations (direct mesozooplankton biomass measurements) from the COPEPOD database (O'Brien, 2014), the operational SEAPODYM-LMTL product yields a coefficient of determination of $R^2 = 0.11$, while its global mean biomass is $1.19\,\mathrm{g\,C\,m^{-2}}$, compared with $0.76\,\mathrm{g\,C\,m^{-2}}$ in the observations (Titaud et al., 2024). These differences motivate a systematic recalibration of the biological component.

The operational SEAPODYM-LMTL implementation does not provide a practical route for automated parameter calibration. Its biological equations are embedded in a non-public C++ code and tightly coupled to the spatial transport component through an implicit solver. Evaluating each parameter set would therefore require running the full spatial model, whereas an optimization typically requires thousands of such evaluations. This makes systematic calibration computationally impractical at global scale or at fine resolution, and prevents the biological component from being tested independently.

An open and lightweight implementation that separates the biological equations from spatial transport and can be coupled directly to optimization algorithms is therefore needed. For this, we developed SeapoPym, an open-source Python reimplementation of the SEAPODYM-LMTL biological model, together with a flexible framework for automated parameter optimization. SeapoPym isolates the biological equations from spatial transport and solves the temperature- and NPP-driven dynamics locally in a zero-dimensional configuration, allowing many parameter sets to be evaluated efficiently. Two experiments establish whether SeapoPym is a valid basis for parameter optimization. We verify that it converges to the analytical steady state of the reference equations, and we quantify how far it departs from the operational SEAPODYM-LMTL product.

Two further experiments, under controlled conditions, establish whether the five biological parameters can be recovered, and from which observations. The first, a Sobol sensitivity analysis, determines what each parameter controls, the magnitude of the biomass or the timing of its seasonal peak. The second, a set of twin experiments, measures identifiability, the capacity

to recover the true parameter values from the data. The optimization framework must recover known parameter values from synthetic biomass time series the model generated itself, across contrasting environmental regimes from cold and productive to warm and oligotrophic. These experiments are a prerequisite to any real-data calibration.

All abbreviations used in this paper are listed in Appendix C.

## 2 Material and methods

### 2.1 SEAPODYM-LMTL (the reference model)

SEAPODYM (Spatial Ecosystem and Populations Dynamics Model) was originally developed to simulate the dynamics of tuna and tuna-like populations under the combined influence of fishing and the environment (Lehodey et al., 2008). Its lower trophic levels are represented by a dedicated component, SEAPODYM-LMTL (Low and Mid-Trophic Levels), which the SeapoPym model re-implements and which we describe here. In its first versions, SEAPODYM-LMTL represented the lower trophic levels as a single biomass pool, the prey ("forage") available to tuna (Lehodey et al., 1998; Lehodey, 2001). That pool was later split into one mesozooplankton group and six micronekton groups (Lehodey et al., 2010; Conchon, 2016). It simulates their biomass as a system of advection-diffusion-reaction equations driven by physical and biogeochemical forcing (Lehodey et al., 2010). Its operational hindcast is distributed through the Copernicus Marine Service (product GLOBAL_MULTIYEAR_BGC_001_033) as a set of Essential Ocean Variables for downstream ecosystem and fisheries applications (E.U. Copernicus Marine Service Information (CMEMS), 2024; Titaud et al., 2024). The operational implementation is written in C++, with the biological reaction terms tightly coupled to both advection and diffusion terms within the numerical solver.

SEAPODYM-LMTL represents the pelagic habitat with three vertical layers whose boundaries follow the euphotic depth $Z_{\mathrm{eu}}$: an epipelagic layer (down to $1.5\,Z_{\mathrm{eu}}$), an upper mesopelagic layer (to $4.5\,Z_{\mathrm{eu}}$), and a lower mesopelagic layer (to $10.5\,Z_{\mathrm{eu}}$, capped at 1000 m) (Lehodey et al., 2010, 2015). The six micronekton groups are defined by their diel vertical migration between these layers: three stay in a single layer over the day-night cycle, and three migrate, tracking the light field (Lehodey et al., 2010). For the migrating groups, the temperature and currents they experience are averaged over the layers they occupy by day and by night, weighted by day length (Lehodey et al., 2010). Mesozooplankton, which are the focus of this study, are represented by a single group which inhabits the epipelagic layer, does not migrate (Conchon, 2016), and is independent of other groups (see also Appendix A). Diel vertical migration therefore does not apply to it, and the vertical structure reduces to the epipelagic layer alone.

The biology of the mesozooplankton group rests on three coupled processes: energy intake from primary production, aging through an age-structured zooplankton production ($p$), and its recruitment ($R$) into a zooplankton biomass ($B$) pool that is subject to mortality. A fixed fraction $E$ of NPP enters the youngest production class, and this production ages as a daily cohort until it is recruited into the biomass pool at a temperature-dependent recruitment age $\tau_r$ (Lehodey et al., 2010; Conchon, 2016). The biomass pool then declines through a temperature-dependent mortality rate $\lambda$, so that at steady state the biomass equals the recruitment flux divided by the mortality rate, $B_{\mathrm{eq}} = R/\lambda$ (Lehodey, 2001). The recruitment age and the mortality timescale $1/\lambda$ both shorten with temperature, following the exponential form expressed in the normalized temperature $\bar{T} =$

$T/(1+T/273)$ (Lehodey et al., 2010), after the metabolic theory of Gillooly (2000); Gillooly et al. (2002). Production ages by one class per day and the model resolves at most eleven daily cohorts, set by the maximum recruitment age ($\tau_{r_0} = 10.38\,\mathrm{d}$, reached at the reference temperature $T_{\mathrm{ref}} = 0^{\circ}\mathrm{C}$). The full equations and their numerical solution are given in Appendix A. The spatial transport that SeapoPym omits, an advection-diffusion scheme shared with the upper-trophic-level model, is described in Sibert et al. (1999) and Lehodey et al. (2008).

The reference values of the five biological parameters are listed in Table 1. The recruitment and mortality temperature relationships are taken from the literature on copepod development and metabolism (Huntley and Lopez, 1992; Gillooly, 2000; Gillooly et al., 2002), whereas the energy-transfer coefficient was estimated by maximum likelihood against the COPEPOD mesozooplankton database, where it reaches $E \approx 0.167$, above the canonical 10% trophic-transfer efficiency (Titaud et al., 2024). These values have been refined across successive versions of the operational SEAPODYM-LMTL product, but the five parameters have never been estimated jointly: the micronekton energy-transfer coefficients were optimized separately against acoustic data (Lehodey et al., 2015) and in a synthetic observing-system experiment (Delpech et al., 2020), while the zooplankton recruitment and mortality parameters were manually tuned (Conchon, 2016).

### 2.2 SeapoPym (the model and its optimization framework)

The SeapoPym model is an open-source Python re-implementation of the SEAPODYM-LMTL biological reaction terms, that was developed for this study and released under the GPLv3 license (Lehodey, 2026). It reproduces the recruitment, aging and mortality described above (see also Appendix A), driven by the same temperature and NPP forcing, but omits spatial transport: the dynamics are solved locally, in a zero-dimensional (0D) configuration. The quantity SeapoPym computes is the time evolution of zooplankton carbon biomass, in $\mathrm{g\,C\,m^{-2}}$. The model is run independently at each grid point. Because the dynamics are time-dependent and age-structured, two cells with the same mean temperature and mean NPP produce different biomass, so the local forcing cannot be reduced to its mean. The formulation is identical across the functional groups of the SEAPODYM-LMTL model, which differ in their vertical-layer occupancy. We restrict this study to the single epipelagic mesozooplankton group.

In the SeapoPym code, each biological process is implemented as an independent function (a kernel), and a simulation is assembled by chaining these kernels into a pipeline, while a configuration layer enforces data integrity and dimensional consistency through the Pint library (https://pint.readthedocs.io/). Processes that do not depend on the previous time step (temperature normalization and layer averaging, day length, energy intake defined as a fraction $E$ of primary production, temperature-dependent mortality rates, and recruitment window) are calculated directly and vectorized over the entire simulation period using NumPy (Harris et al., 2020). The two recursive processes, the aging and recruitment of production cohorts and the biomass update from one step to the next, are evaluated as time loops compiled with Numba (Lam et al., 2015), given that each step depends on the preceding one. Thanks to this near-universal vectorization and compilation limited to the recursive parts, a single run is fast, while multiple simulations (covering grid cells, sensitivity analysis or optimization parameters sets) can be launched in parallel via Xarray (Hoyer and Hamman, 2017) and Dask (Dask Development Team, 2016), whether on a laptop or a computing cluster.

The SeapoPym framework includes a parameter-estimation component (optimizer) built on the same principle of modularity. The model definition, the observations to be fitted, the cost function and the optimization algorithm constitute independent components, each defined by a minimal interface and thus interchangeable. The component handling observations accepts various data types (time series or spatial fields), and the cost function itself is a replaceable module, allowing the minimized metric to be changed without altering the rest of the framework. The optimization algorithm is interchangeable based on the same principle. The framework currently offers a genetic algorithm as well as CMA-ES. The present study employs CMA-ES (Sect. 2.6). It is this separation of model, data, cost function and optimizer that makes SeapoPym a general-purpose experimental framework rather than a rigid processing pipeline.

An overview of the full modeling and experimental workflow, including the experiments described below, is given in Fig. D1 (Appendix D).

### 2.3 Forcing, CMEMS product, stations and observations

SeapoPym is forced with the same temperature and NPP fields as the operational SEAPODYM-LMTL product, the Copernicus Marine Service "Global Ocean Low and Mid Trophic Levels Biomass Content Hindcast" (GLOBAL_MULTIYEAR_BGC_001_033, https://doi.org/10.48670/moi-00020) (Titaud et al., 2024). The temperature is the layer average over the epipelagic layer (named $T_{\mathrm{env}}$ in Appendix A) from the GLORYS12 reanalysis (Lellouche et al., 2021) whose depth is set by the euphotic depth (Sect. 2.1). The NPP and euphotic depth come from the same satellite ocean-color product, in which the NPP is computed with the VGPM algorithm (Behrenfeld and Falkowski, 1997) and is supplied as a vertically integrated value. This preparation is done within the product, not by SeapoPym, and the fields cover 1998 to 2019. These fields are provided at daily resolution, matching the model's one-day time step. The first two years are discarded as spin-up, justified by the model's equilibration time (Sect. 3.1). The global mean distributions of the temperature and primary-production forcing, together with the reference SEAPODYM-LMTL biomass and the current field, are shown in Fig. 1.

For the sensitivity and twin experiments, the model is run at the location of six existing observation stations chosen to span a large temperature and primary-production range (Fig. 2). In order of increasing temperature, these are the Barents Sea (BARENTS, 75°N, 40°E), a subarctic northeast Pacific location in the Line P region (PAPA, 50°N, 132°W), the Bay of Biscay (BISCAY, 45.5°N, 4°W), the Canary region (CANARY, 30°N, 13°W), the Bermuda Atlantic Time-series Study site (BATS, 32°N, 64°W) and the Hawaii Ocean Time-series station ALOHA (HOT, 23°N, 158°W). Seasonality does not follow mean temperature alone across these stations. HOT and BATS share a similar mean temperature and productivity, yet the seasonal range of their forcing differs, with a primary-production inter-quartile range about twice as wide at BATS as at HOT (Fig. 2).

SeapoPym does not simulate transport. It is therefore necessary to verify whether this simplification is acceptable. To determine whether the resulting error is significant or negligible, we compare it to the existing discrepancy between the operational SEAPODYM-LMTL product and in-situ observations. For this purpose, we use long-term in-situ zooplankton observational data from the two stations for which these time series are freely available: HOT (Station ALOHA) and BATS. These data originate from the Hawaii Ocean Time-series (HOT) and the Bermuda Atlantic Time-series Study (BATS) programs, in the form of total mesozooplankton net samples (200 $\mu$m). We retain the epipelagic zooplankton samples and exclude the larger

size fraction (>5 mm), consistent with the single epipelagic mesozooplankton group modeled here. The samples are reported as dry-weight biomass per unit volume ($\mathrm{mg\,m^{-3}}$). Dry-weight biomass is converted to carbon using a uniform factor of 0.4, since carbon represents approximately 40% of zooplankton dry weight (Omori, 1969). This conversion factor is used for both stations. This value is then multiplied by the epipelagic-layer thickness ($1.5\,Z_{\mathrm{eu}}$, Sect. 2.1), taken from the product's time-varying euphotic depth, to obtain carbon biomass per unit area (in $\mathrm{g\,C\,m^{-2}}$), which is directly comparable to the model results. These observations serve only to compare the models' order of magnitude to the in-situ zooplankton observations (Sect. 2.4) rather than for calibration.

### 2.4 Implementation validation and inter-model differences

The first validation experiment for the SeapoPym model compares the model's convergence toward the analytical solution $B_{\mathrm{eq}} = R/\lambda$ (Sect. 2.1, Appendix A) under constant forcing. This experiment is conducted at four different temperatures (0, 10, 20 and 30 °C) and a fixed NPP of $300\ \mathrm{mg\,C\,m^{-2}\,d^{-1}}$, and we analyze both the absolute difference at equilibrium and the time required to reach that equilibrium. The biomass starts from zero, and we take the time to equilibrium as the time to come within $1\%$ of the analytical steady state.

The second experiment compares SeapoPym with the operational SEAPODYM-LMTL product, which includes transport, over the global ocean and the 2000 to 2019 period (Sect. 2.3). This comparison does not separate the effect of transport from that of the re-implementation. We quantify this difference at each grid cell with three complementary metrics (Eqs. (1)–(3)), mapped over the global ocean in Fig. 4. Each metric compares the SeapoPym biomass $\hat{y}_i$ with the reference biomass $y_i$ from SEAPODYM-LMTL, at time step $i$ among the $n$ steps of the series. The bias is the mean signed error, so opposite differences cancel and it can stay near zero even where the two models differ most. The root mean square error (RMSE) does not cancel in this way and measures the magnitude of the difference, giving more weight to large errors. The mean absolute percentage error (MAPE) expresses the difference relative to the local biomass, which matters where biomass is low.

$$\mathrm{bias} = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i) \tag{1}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{2}$$

$$\mathrm{MAPE} = \frac{100}{n}\sum_{i=1}^{n}\frac{|\hat{y}_i - y_i|}{|y_i|} \tag{3}$$

To assess whether this difference is significant, we compare it to the discrepancy already existing between the operational SEAPODYM-LMTL product and in-situ observations. At the HOT and BATS stations, where in-situ time series are available, we use the root mean square error to quantify the two discrepancies of the operational SEAPODYM-LMTL product, one with

the SeapoPym model and one with the in-situ observations (Fig. 5). We then calculate their ratio, allowing us to place the inter-model difference in the context of the gap between the operational SEAPODYM-LMTL product and the in-situ observations. This comparison focuses solely on the order of magnitude.

### 2.5 Sobol sensitivity analysis

A Sobol sensitivity analysis makes it possible to determine what each parameter controls prior to calibration. It is a global, variance-based method that apportions the variance of a model output among the parameters. It is widely used to identify and rank the driving factors of environmental models (Pianosi et al., 2016; Sobol', 2001). For each parameter, it provides a first-order index $S_1$, representing the fraction of output variance attributable to that parameter alone, as well as a total index $S_T$, that represents the fraction attributable to that parameter and all its interactions with others. The difference $S_T - S_1$ corresponds to the portion of a parameter's influence exerted only in combination with other parameters, isolated through variance decomposition (Saltelli, 2002). The analysis relies on output variance, a quantity independent of the cost function minimized during the optimization (Sect. 2.6).

We run SeapoPym at the six station locations (Sect. 2.3), with the two-year spin-up used throughout, and we vary the five biological parameters over the same ranges as the twin experiments (Sect. 2.6). The result of a sensitivity analysis depends on the chosen descriptors (Pianosi et al., 2016), so we reduce each simulation to two complementary scalars. The first descriptor is the base-10 logarithm of the mean biomass, which measures the overall level, taken in logarithm because the mean spans several orders of magnitude across parameters sets. The second descriptor is the day of year of the biomass maximum, which measures the timing of the seasonal peak. We compute the indices separately at each station location, so that any dependence of the sensitivity on the environment is visible.

We compute the Sobol indices with the SALib library (Herman and Usher, 2017; Iwanaga et al., 2022). Its Saltelli sampler `sample_sobol` generates the sets of parameter values at which the model is run, and its estimator `sobol.analyze` returns $S_1$ and $S_T$ with 95% bootstrap confidence intervals. For the $D = 5$ parameters, a base sample size $N$ yields $N(D+2)$ parameter sets, one per model run. Only first- and total-order indices are computed, which is what fixes the count to $N(D+2)$. The Saltelli sample and the bootstrap confidence intervals both use a fixed random seed, so the design and the reported intervals are reproducible. Rather than fixing $N$ in advance, we choose it by a convergence test on the indices. We raise $N$ through a doubling sequence, re-estimate the indices at each value, and stop at the smallest $N$ that meets two conditions. The first follows Sarrazin et al. (2016) and requires precision: the half-width of the 95% bootstrap confidence interval of each index must stay below 0.05. The second requires stability: each index must change by less than 0.05 from the previous $N$. Both conditions are first met at $N = 8192$, that is $N(D+2) = 57\,344$ model runs.

### 2.6 CMA-ES: parameter optimization

We assess identifiability with twin experiments. From the reference parameters (Table 1) we simulate the synthetic observations at each station over the 2000 to 2019 analysis period, then we use the optimizer to recover those parameters by fitting these synthetic observations either separately or together. Each candidate runs its own two-year spin-up over 1998 and 1999 within

a full 1998 to 2019 integration, and the cost is scored on the 2000 to 2019 analysis window (Sect. 2.3). Because the synthetic observations are produced by the transport-free model itself and used without added noise, the target contains no transport signal, and an exact recovery is possible in principle. Twin experiments of this kind were used on the same model family by Delpech et al. (2020) for the micronekton energy-transfer coefficients.

We recover the parameters with CMA-ES, that is an algorithm of adaptation evolution strategy using the covariance matrix (Hansen and Ostermeier, 2001; Hansen, 2016). CMA-ES is a derivative-free method that adapts a multivariate normal search distribution to the local shape of the cost, which suits a smooth but possibly ill-conditioned cost without requiring its gradient. We use pycma, its reference implementation (Hansen et al., 2019), with the five parameters mapped to the unit box (linearly normalized between 0 and 1) from the search ranges of Table 1 and the standard CMA-ES settings. The default population size, i.e. the number of parameter sets to test in one iteration, is defined by $\lfloor 4 + 3\ln D \rfloor = 8$ for $D = 5$ parameters (Hansen, 2016). The initial step size is $\sigma_0 = 0.30$, and candidate solutions that leave the unit box are mapped back into it by a smooth transformation, pycma's default boundary handling. The search stops on the standard CMA-ES convergence criteria, when the step size and the spread of the cost fall below pycma's default tolerances (step-size tolerance $10^{-4}$). A generation cap is used solely as a safety measure, as it is never reached during the experiments.

A single optimization can settle in a local optimum that depends on its starting point, so we repeat the search from twenty random starts (seeds 0 to 19), hereafter called restarts, and keep the best. We assess identifiability by comparing the parameters recovered by the run achieving the best score with the reference. Parameters matching the reference are well constrained by the synthetic observations, whereas a recovered parameter set that departs from the reference while it gets the best score suggests a problem of equifinality (Beven, 2006).

The cost function is the root mean square error normalized by the mean of the synthetic observations (NRMSE). Following Eq. (2), $\hat{y}$ refers to the simulated biomass and $y$ refers to the synthetic observations, so that stations with different biomasses enter on a comparable scale. For the joint experiment, the cost function is the mean of the per-station NRMSE. We conduct seven experiments, each of the six stations individually and one joint experiment (MERGED) that combines all six. The recovered parameter values are reported with the results.

## 3 Results

### 3.1 Implementation validation and the impact of transport

Under constant forcing, the simulated biomass converges to the analytical steady state $B_{\mathrm{eq}} = R/\lambda$ at every temperature tested, the relative departure falling below $0.01\%$ given a long enough integration (Fig. 3). It reaches within $1\%$ of this equilibrium in about fifteen days at $30^{\circ}\mathrm{C}$ and about two years at $0^{\circ}\mathrm{C}$, the adjustment slowing as temperature falls. This slow adjustment in cold water sets the two-year spin-up used throughout the study, which brings even the coldest water to within about $1\%$ of equilibrium, so the analysis at every station begins from a near-equilibrium state.

The difference between SeapoPym and the operational SEAPODYM-LMTL product measures the combined effect of omitting transport and the re-implementation. As Fig. 4 shows, it is non-uniform across the ocean. In absolute terms, the RMSE is

largest where currents and gradients are strong, along the Gulf Stream, the Kuroshio and the Antarctic Circumpolar Current, and in cold high-latitude waters where the long life cycle of zooplankton leaves the biomass exposed to advection for a longer period. The MAPE, calculated relative to local biomass, reduces the weights of these high-biomass currents and instead highlights regions of lower biomass, particularly around the equator in the Pacific, Indian and Atlantic Oceans. The bias introduces a directional dimension. Its opposite-signed fronts mark neighboring areas where the 0D model shows higher and lower biomass than the two-dimensional (2D) reference, a dipole consistent with biomass displaced by currents, such as at the confluence in the South Atlantic between the warm Brazil Current and the cold Malvinas Current. This signed pattern is consistent with transport redistributing biomass between adjacent regions, although a static difference map does not resolve the displacement itself.

To assess the significance of this difference, we compare it to the discrepancy already existing between the operational SEAPODYM-LMTL product and the in-situ observations. At HOT and BATS stations the difference between the 0D and 2D models is five to seven times smaller than the gap between the operational SEAPODYM-LMTL product and the observations (Fig. 5, see Table B1 for station-specific values). The two models yield similar results, yet both deviate from the observations in terms of amplitude: they underestimate values at HOT and overestimate them at BATS. This justifies the use of local 0D optimization in these regimes. This finding applies to both HOT and BATS, which are warm-water oligotrophic stations with in-situ measurement series. The comparison focuses solely on magnitude, as the observations are not used for calibration.

### 3.2 Sobol sensitivity analysis

Sensitivity analysis of biomass magnitude reveals a distribution of roles (Fig. 6). This magnitude is determined by energy transfer $E$ and the two mortality parameters, with the balance between these factors shifting along the temperature gradient. First-order indices for $E$ and $\lambda_0$ decrease from the coldest station to the warmest station, whereas the index for $\gamma_\lambda$ rises from a value near zero at BARENTS to approximately 0.6 at HOT. The thermal term $\gamma_\lambda$ is inactive near the reference temperature but dominant in warm waters. Recruitment parameters have no impact on this magnitude, and for this descriptor, the indices are additive, summing to one at each station.

The peak timing is governed by the recruitment parameters, while energy transfer has no effect. At BARENTS, BISCAY, CANARY and BATS the first-order indices sum to 0.6 to 0.8, so recruitment acts mostly directly, whereas that sum falls to 0.33 at PAPA and 0.17 at HOT, so interactions carry most of the timing variance. The two recruitment parameters respond to temperature differently. At BARENTS the first-order index of $\tau_{r_0}$ reaches 0.72, and its total-order index falls with warming, from 0.96 at BARENTS to 0.40 at BATS. The first-order index of $\gamma_{\tau_r}$ rises over the same gradient, from near zero at BARENTS to 0.60 at BATS. HOT departs from both trends: the first-order indices of $\tau_{r_0}$ and $\gamma_{\tau_r}$ fall below 0.1 while their total-order indices reach 0.68 and 0.82, so recruitment acts on the timing almost entirely through interactions.

Taken together, these two descriptors distinguish the respective roles of the parameters: magnitude is determined by energy transfer and mortality, whereas timing is determined by recruitment. At steady state, biomass results from a multiplicative combination of the parameters (Appendix A). Therefore, its logarithm is expressed as a sum of distinct terms: $\log B = \log E - \log \lambda_0 - \gamma_\lambda \bar{T} + \log \mathrm{NPP}$, where $\bar{T}$ and NPP are fixed by the station and where recruitment does not appear.

The sensitivity analysis confirms this simplified scheme, showing additive indices for magnitude and a negligible effect of recruitment. However, sensitivity does not imply identifiability: knowing what each parameter controls does not determine whether it can be estimated from the data. This is a point that is demonstrated by the twin experiments presented in Sect. 3.3.

### 3.3 Twin experiments: convergence and parameter recovery

We read parameter recovery from the best restart of each experiment (Table 2). At the cold stations BARENTS, PAPA and BISCAY, and in the joint MERGED experiment, all five parameters are recovered to better than 0.1% of the reference, recruitment included. At the warm stations CANARY and BATS the energy transfer and the two mortality parameters are recovered to the same precision, but the two recruitment parameters $\tau_{r_0}$ and $\gamma_{\tau_r}$ miss the reference by 40 to 160%.

To validate convergence we use the cost reached by the experiments that recover the parameters as a benchmark. The cold stations and MERGED reach an NRMSE of order $10^{-4}$ (Table 2), and CANARY and BATS reach it as well, so both converged (Fig. 7). HOT's best restart, by contrast, stays near $10^{-2}$, two orders of magnitude above, so none of its twenty restarts converged and we leave HOT out of the recovery analysis.

Three conclusions follow. At CANARY and BATS the search reaches the cost of a full recovery yet leaves the recruitment parameters far from the reference, so a good fit does not identify the parameters, the equifinality of Beven (2006). At the cold stations, by contrast, the search both converges and recovers all five parameters. In MERGED, the joint search over the six stations converges and recovers every parameter to better than 0.1% of the reference, so the non-identifiable warm stations do not degrade the joint estimate.

## 4 Discussion

The first question this study addressed is whether a transport-free re-implementation is a valid basis for parameter optimization. SeapoPym converges to the analytical steady state of the reference equations, and its departure from the operational SEAPODYM-LMTL product is small enough for local 0D optimization to be valid, at least at the two warm stations where in-situ observations allow that comparison. The second question is whether zooplankton observations carry enough information to recover the five biological parameters of the SEAPODYM-LMTL model. Using the SeapoPym model coupled to a CMA-ES optimizer, we found that the answer depends on the environmental regime sampled. The energy-transfer and mortality parameters were recovered at every station where the search converged, whereas the recruitment parameters were recovered only in cold water.

The influence of advection on zooplankton distribution has long been recognized and is often pronounced, particularly in regions characterized by intense physical dynamics, such as frontal zones, upwelling systems, and areas with strong currents (e.g., Banse, 1964; Edvardsen et al., 2006; Mackas and Coyle, 2005). Our comparison with the operational SEAPODYM-LMTL product confirms that transport cannot be ignored in such areas, especially when these conditions prevail in cold-water environments. This result aligns with theoretical predictions stating that physical transport becomes significant when biological timescales exceed physical retention timescales. This scenario is more likely in cold waters, where zooplankton generation

times can exceed one year (Conover, 1988; Kosobokova, 1999; Daase et al., 2013). Nevertheless, the version of our model that omits transport remains suitable for simulating and estimating mesozooplankton parameters in most warm- and temperate-water regions, outside of localized dynamic zones. This is the case at our six stations, where the 0D-to-2D difference stays between 6 and 12% (Table B1). This difference is smaller than the model-observation gap only where we could measure that gap, at the two warm stations HOT and BATS. Confirming it in cold water would require in-situ series there.

The Sobol sensitivity analysis assigns a distinct role to each parameter: energy transfer and the two mortality parameters govern the magnitude of the biomass, while the two recruitment parameters govern the timing of the seasonal peak, with their relative predominance depending on the temperature regime. This division shapes what a calibration against real data can and cannot separate. At steady state the mean biomass depends on the energy transfer $E$, the mortality $\lambda$, and the primary production through the ratio $\frac{E \times \mathrm{NPP}}{\lambda}$ (Appendix A), so these three are confounded in the magnitude. A bias in the NPP forcing magnitude could therefore be absorbed into biased estimates of $E$ and $\lambda$, indistinguishable in the mean biomass from a genuine parameter error. The recruitment parameters play an analogous role for timing, absorbing errors in the seasonal phase of the forcing. Our twin experiments recover $E$ and the mortality parameters because the forcing is identical for the target and the candidates. Against uncertain forcing these same parameters could instead take values that compensate for it, so their calibrated estimates should be interpreted as effective quantities rather than independently identified ones.

The same reasoning applies to transport. Since the synthetic observations are generated by SeapoPym, they contain no transport signal to be absorbed. These twin experiments therefore do not allow us to assess whether the optimized parameters would take biased values to compensate for the omission of transport. Assessing this would require the same experiment with a target that carries a transport signal, such as the operational product. The recovery presented here should not be interpreted as proof that the estimates would remain unbiased in that case.

At the cold stations the mean recruitment age exceeds two daily cohorts (Table B1), and the two recruitment parameters are recovered. At CANARY and BATS, where the mean recruitment age is close to one daily cohort, the recruitment parameters are not recovered, even though the fit reaches the same cost as at the cold stations. Consequently, the quality of the fit alone does not guarantee the validity of the calibration. Sensitivity does not imply identifiability either, since at BATS the first-order index of $\gamma_{\tau_r}$ on the peak timing reaches 0.60. A likely reason lies in how recruitment is discretized. Age classes are one time step wide, and a class is completely absorbed into the biomass (Appendix A), so two recruitment ages within the same age class produce the same recruitment and the cost changes in steps rather than continuously. The recruitment age also decreases exponentially with temperature, so only one or two age classes are resolved in warm water. Crossing a class boundary there demands either a wider temperature range or a larger parameter change than at the cold stations. This warm-water equifinality would therefore stem from both the exponential behavior of recruitment and the numerical solution the model uses. A smoother recruitment, or a finer time and age step, would smooth the cost function, but we have not tested whether either restores identifiability. Recruitment parameters must therefore be determined from independent data or estimated in conjunction with an informative cold station, as demonstrated by the MERGED experiment.

These results have direct implications for observing-system design. Recovering all five parameters requires sampling at least one cold-water regime, because recruitment is identifiable there. A single cold station, sampled across its full seasonal range,

constrained the five parameters as well as the six stations combined. What matters is the information content of the regime, not the number of stations. The cold station must also have weak enough currents for the 0D approximation to hold, as at BARENTS, PAPA and BISCAY, where the 0D-to-2D difference is of the same order, in relative terms (MAPE), as at the warm stations. Cold but strongly advective regions carry the same seasonal signal, but exploiting it would require a version of the model that resolves transport.

These findings are a best case. The synthetic observations are noise-free, complete at every time step, and driven by the exact forcing, whereas in-situ series are sparse and irregular and the forcing carries its own error. All three will lower identifiability and widen parameter uncertainty in real applications. Working with such observations will also require a cost function that represents their error, such as a Gaussian likelihood that weights the misfit by the observation-error variance, or a log-normal likelihood suited to the positive, skewed distribution of plankton biomass (Schartau et al., 2017). A second limitation is our diagnosis of equifinality. We identified it from the best restart alone, which showed that parameter sets far from the reference reached the same cost, but this single optimum does not reveal the whole family of parameter sets that fit the observations equally well. Mapping that family would require sampling the parameter space rather than a single optimum, either as a behavioral ensemble drawn by random sampling (Beven and Binley, 1992) or as a Bayesian estimate of the parameter posterior (Tarantola, 2005).

The first step beyond this study is to repeat these experiments under realistic conditions, as an observing system simulation experiment in which measurement error, temporal gaps, and forcing uncertainty are added to the synthetic observations (Hoffman and Atlas, 2016). Such an experiment would show how much of the identifiability reported here still holds under those conditions. It would also let us vary the sampling effort, which we held fixed, since every station used the same complete twenty-year record. We do not know whether a long continuous record outperforms shorter records spread over several stations, nor where along the temperature gradient recruitment identifiability breaks down. The second step is to calibrate the model against real data. The long-term series at HOT and BATS are complete, but this study shows that they cannot constrain recruitment. A cold and weakly advective station is needed as well. The COPEPOD database is a logical candidate, although a long-term series in cold water would constrain recruitment more directly. The SeapoPym framework is built for both steps.

## 5 Conclusions

We presented here SeapoPym, an open-source Python re-implementation of the SEAPODYM-LMTL model for mesozooplankton biology. It is coupled with a modular framework in which the model, observations, cost function, and optimizer are interchangeable components. We validated this re-implementation against the operational SEAPODYM-LMTL product and its analytical steady state. We determined the roles of the five parameters using a Sobol analysis, then tested their identifiability using twin experiments at six station locations with different environmental conditions. The difference between the 0D and 2D models, which combines the omission of transport and the re-implementation, ranges from 6 to 12% of the simulated biomass at all six stations (Table B1). At HOT and BATS, the only stations with in-situ records, it is five to seven times smaller than the

discrepancy between the model and the observations, so it is second-order there. The 0D model therefore serves as a valid test bed for assessing parameter identifiability where transport is a second-order effect.

The identifiability of the parameters depends on the environmental conditions observed. In the twin experiments, the energy-transfer and mortality parameters were correctly estimated at every station where the search converged, while the recruitment parameters were estimated only under cold-water conditions. At warm stations, the optimization reaches parameter sets far from the reference that minimize the cost as well as at the cold stations. Cost convergence, therefore, does not imply that the parameters are correctly estimated. A single cold station constrained all five parameters as well as the six stations combined. Parameter recovery is therefore determined by the information content of the sampled regime, not by the number of stations. In locations where the difference between the 0D and 2D models is second-order, the 0D model can be calibrated using observations made under cold, low-current conditions, where recruitment is informative. In the absence of such cold observations, recruitment parameters must be set or constrained based on independent, process-based knowledge.

The twin experiments use noise-free observations, complete at every time step, generated by the transport-free model itself and driven by the exact forcing. The identifiability they show is therefore a best case, and it leaves open whether the parameters would absorb the missing transport once the target contains it. In-situ series are sparse and irregular, and the forcing carries its own error, so real conditions will reduce the identifiability. One warm station, HOT, did not converge to the low cost reached at the other stations and departed from them for reasons that remain open. Calibrating the model against real data will require a cold, weakly advective station with a long time series, where recruitment is informative and the 0D approximation holds, and a cost function that represents observational error.

## Appendix A: Underlying model equations

The full SEAPODYM-LMTL model resolves several functional groups across three vertical layers, with diel vertical migration linking the layers (Sect. 2.1, Lehodey et al. 2010). The equations below describe the single epipelagic mesozooplankton group studied here. The functional groups are not coupled to one another, each driven independently by primary production, and the epipelagic group does not migrate, so neither inter-group terms nor diel vertical migration appear here. The group is represented by two coupled state variables, an age-structured production field $p(t,\tau)$ and an age-aggregated biomass pool $B(t)$. All equations are written in continuous form, the discretization of time and age being introduced only in the numerical-solution part.

### A1 Biomass

The biomass $B$ $[\mathrm{g\,C\,m^{-2}}]$ follows the balance between the recruitment flux $R$ and mortality:

$$\frac{dB}{dt} = R(t) - \lambda(t)\,B(t) \tag{A1}$$

where $R$ [$\mathrm{g\,C\,m^{-2}\,d^{-1}}$] is the recruitment flux from the aging production field (defined below) and $\lambda$ [$\mathrm{d^{-1}}$] is the mortality rate. At steady state this gives the equilibrium biomass $B_{\mathrm{eq}} = R/\lambda$. Mortality rises exponentially with temperature:

$$\lambda(t) = \lambda_0 \exp(\gamma_\lambda \bar{T}(t)) \tag{A2}$$

where $\lambda_0$ is the mortality rate at the reference temperature and $\gamma_\lambda$ [$^\circ\mathrm{C}^{-1}$] its thermal sensitivity. The temperature entering the metabolic terms is the normalized temperature of the SEAPODYM-LMTL model (Lehodey et al., 2010), after the metabolic theory of Gillooly (2000); Gillooly et al. (2002),

$$\bar{T}(t) = \frac{T(t)}{1 + T(t)/273}, \qquad T(t) = \max(T_{\mathrm{env}}(t), T_{\mathrm{ref}}), \quad T_{\mathrm{ref}} = 0^\circ\mathrm{C}, \tag{A3}$$

where $T_{\mathrm{env}}$ is the temperature forcing (Sect. 2.3) and $T_{\mathrm{ref}}$ the reference temperature, with $T$ floored at $T_{\mathrm{ref}}$ so that the recruitment age cannot exceed the maximum age the model resolves (below).

### A2 Production

The production field $p(t,\tau)$ [$\mathrm{g\,C\,m^{-2}\,d^{-1}}$] describes how production is distributed across age $\tau$. It ages at unit velocity and is conserved while it ages, following the loss-free McKendrick-Von Foerster transport equation (McKendrick, 1926; Von Foerster, 1959):

$$\frac{\partial p}{\partial t} + \frac{\partial p}{\partial \tau} = 0 \tag{A4}$$

Its dynamics are set by two boundaries, a source at age zero and a sink at the recruitment age. At age zero, a fraction $E$ of NPP is injected (the source):

$$p(t, \tau = 0) = E\,\mathrm{NPP}(t), \tag{A5}$$

with NPP the net primary production [$\mathrm{g\,C\,m^{-2}\,d^{-1}}$] and $E$ the dimensionless transfer efficiency between primary production and zooplankton production. The sink at the recruitment age is described next.

### A3 Recruitment

Production is recruited into the biomass pool when it reaches the temperature-dependent recruitment age $\tau_r(t)$:

$$\tau_r(t) = \tau_{r_0} \exp(\gamma_{\tau_r} \bar{T}(t)), \tag{A6}$$

where $\tau_{r_0}$ is the recruitment age at the reference temperature (hence the maximum age represented in the model) and $\gamma_{\tau_r}$ [$^\circ\mathrm{C}^{-1}$] its thermal sensitivity. This sink is an absorbing boundary, $p(t,\tau) = 0$ for $\tau > \tau_r(t)$, and the recruitment flux feeding the biomass is the production reaching that age:

$$R(t) = p(t, \tau_r(t)). \tag{A7}$$

Recruitment moves production from the field $p$ into the pool $B$, a transfer and not a loss. Mortality acts only on $B$, through $\lambda$, and any loss during development is absorbed into the efficiency $E$.

### A4 Numerical solution

Age and time are discretized into steps $\Delta\tau$ and $\Delta t$, and production is carried in discrete cohorts (age classes) indexed by $a$, advanced over time steps indexed by $n$. Because production ages at unit velocity, we set $\Delta t = \Delta\tau \equiv \Delta = 1\,\mathrm{day}$, a Courant number of one, so one time step moves a cohort up by exactly one age class. The number of resolved age classes is set by the maximum recruitment age, about eleven daily classes at the reference temperature ($\tau_{r_0} = 10.38\,\mathrm{d}$). Each step then reduces to three operations on the production field, by region of age:

$$p_a^{n+1} = \begin{cases} E\,\mathrm{NPP}^{\,n+1} & a = 0 \qquad \text{(source at age zero)} \\ p_{a-1}^{n} & 0 < a\Delta < \tau_r^{n+1} \qquad \text{(loss-free aging)} \\ 0 & a\Delta \geq \tau_r^{n+1} \qquad \text{(recruitment)} \end{cases} \tag{A8}$$

The production that crosses the recruitment age (the classes set to zero) feeds the biomass, giving the recruitment flux $R^{n+1} = \sum_{a\geq 1}(p_{a-1}^n - p_a^{n+1})$. At a Courant number of one these operations are exact, so the production needs no integration scheme. Only the biomass does, because of its mortality term.

The biomass is integrated with an implicit (backward Euler) scheme, evaluating the right-hand side at the new time step:

$$B^{n+1} = \frac{B^n + \Delta\, R^{n+1}}{1 + \Delta\, \lambda^{n+1}}. \tag{A9}$$

This scheme is unconditionally stable. A fully explicit update is not: in warm water and over the high mortality rates explored during the experiments (Sect. 2.5 and 2.6), it produces spurious oscillations and negative biomass. The implicit form lets the whole parameter range be integrated with a single time step.

## Appendix B: Per-station summary

Table B1 gathers, for each of the six stations, the mean temperature, the recruitment age at the mean and warmest points of its forcing, the difference between the 0D and 2D models, and, where in-situ records exist, the structural gap between the operational SEAPODYM-LMTL product and the observations.

## Appendix C: Abbreviations

The abbreviations and acronyms used in this paper are listed in Table C1.

## Appendix D: Workflow overview

Figure D1 gives an overview of the modeling and experimental workflow. Temperature and NPP drive both models, whereas the currents drive only the SEAPODYM-LMTL model (with transport), SeapoPym being transport-free. The figure also shows the in-situ observations, the CMA-ES optimizer, and the four experiments together with the components each one uses.

*Code availability.* The SeapoPym model code is open-source, distributed under the GPLv3 license, and available at https://github.com/SeapoPym/seapopym. The exact version described in this paper is SeapoPym v0.1.1, archived on Zenodo (https://doi.org/10.5281/zenodo.21838667, Lehodey 2026). The repository includes the source code, installation instructions, and a user manual. A dedicated reproducibility deposit accompanying this paper, providing the scripts, data and configuration necessary to reproduce the experiments and figures presented here, is openly available at https://github.com/SeapoPym/SeapoPym-v0.1-Reproducibility (release v2.0.0) and archived on Zenodo (https://doi.org/10.5281/zenodo.21849337, Lehodey et al. 2026). Its frozen experiment outputs are included, so every figure and table can be redrawn without repeating a calibration or a sensitivity analysis. The deposit also pins the complete software environment used for the results in this paper, which were produced with Python 3.12, pycma 4.4.4 for the optimization, and SALib 1.5.2 for the sensitivity analysis.

*Data availability.* The temperature and NPP fields forcing the model, together with the mesozooplankton biomass from the operational SEAPODYM-LMTL product and the epipelagic current field, are distributed by the Copernicus Marine Service (CMEMS) in the product *Global Ocean Low and Mid Trophic Levels Biomass Content Hindcast* (GLOBAL_MULTIYEAR_BGC_001_033, https://doi.org/10.48670/moi-00020, Titaud et al. 2024). Within this product, the epipelagic-layer temperature derives from the GLORYS12 reanalysis (Lellouche et al., 2021) and the vertically integrated NPP from satellite ocean color through the VGPM algorithm (Behrenfeld and Falkowski, 1997). These fields cover the period 1998 to 2019. The in-situ mesozooplankton observations, used only for the magnitude comparison of Sect. 2.4 and Fig. 5 and not for calibration, are openly available. The Bermuda Atlantic Time-series Study zooplankton biomass (Steinberg et al., 2001) is available at https://bats.bios.asu.edu/bats-data/ and archived at https://doi.org/10.5281/zenodo.10182499, and the Hawaii Ocean Time-series macrozooplankton record (Karl and Lukas, 1996) through the HOT-DOGS system (https://hahana.soest.hawaii.edu/hot/hot-dogs/) and archived at https://doi.org/10.5281/zenodo.10850539. Both in-situ datasets were obtained through the Simons Collaborative Marine Atlas Project (Simons CMAP, https://simonscmap.com/) as the datasets `BATS_Zooplankton_Biomass` and `HOT_Macrozooplankton_v2022`.

*Author contributions.* JVL led the development of the SeapoPym software, designed the methodology, performed the simulations and sensitivity analyses, and wrote and revised the manuscript. AM and AG supervised the PhD work, provided guidance on the methodology, and

contributed to the review and revision of the manuscript. SA tested new functionalities, provided feedback on the optimization framework, and contributed to the software validation. SN conceptualized the project, secured funding, supervised the overall research direction, and reviewed the manuscript. All authors approved the final version.

*Competing interests.* The authors declare that they have no conflict of interest.

*Acknowledgements.* We gratefully acknowledge the support and assistance of the Pacific Community (SPC) and the Institut de Recherche pour le Développement (IRD). We also thank the Mercator Ocean International for hosting JVL and providing computational resources, and the Copernicus Marine Service for providing the physical and biogeochemical data products. Finally, the authors are grateful to Patrick Lehodey for his careful proofreading of the manuscript and for his insightful discussions. AI-assisted tools were used for code development, code documentation, and language editing of the manuscript. All code was reviewed and validated by the authors. All scientific content, analyses, and interpretations were produced by the authors.

*Financial support.* Financial support for this work was provided by the Government of New Zealand via a grant from the Ministry of Foreign Affairs and Trade to the Pacific Community (WPG-0103601, DOC-4119683, ACT-0103048) to implement the Climate Science for Ensuring Pacific Tuna Access project. SA's contribution was funded by the European Union under grant agreement no. 101136748 (BioEcoOcean).

## References


Banse, K.: On the vertical distribution of Zooplankton in the sea, Prog. Oceanogr., 2, 53–125, https://doi.org/10.1016/0079-6611(64)90003-5, 1964.

Behrenfeld, M. J. and Falkowski, P. G.: Photosynthetic rates derived from satellite-based chlorophyll concentration, Limnol. Oceanogr., 42, 1–20, https://doi.org/10.4319/lo.1997.42.1.0001, 1997.

Beven, K.: A manifesto for the equifinality thesis, J. Hydrol., 320, 18–36, https://doi.org/10.1016/j.jhydrol.2005.07.007, 2006.

Beven, K. and Binley, A.: The future of distributed models: Model calibration and uncertainty prediction, Hydrol. Process., 6, 279–298, https://doi.org/10.1002/hyp.3360060305, 1992.

Brodeur, R. D., Seki, M. P., Pakhomov, E. A., and Suntsov, A. V.: Micronekton - What are they and why are they important?, PICES Press, 13, 7–11, 2005.

Conchon, A.: Modélisation du zooplancton et du micronecton marins, Ph.D. thesis, Université de La Rochelle, https://theses.hal.science/tel-01661554, nNT: 2016LAROS011, 2016.

Conover, R. J.: Comparative life histories in the genera Calanus and Neocalanus in high latitudes of the northern hemisphere, Hydrobiologia, 167–168, 127–142, https://doi.org/10.1007/bf00026299, 1988.

Daase, M., Falk-Petersen, S., Varpe, Ø., Darnis, G., Søreide, J. E., Wold, A., Leu, E., Berge, J., Philippe, B., and Fortier, L.: Timing of reproductive events in the marine copepod Calanus glacialis: a pan-Arctic perspective, Can. J. Fish. Aquat. Sci., 70, 871–884, https://doi.org/10.1139/cjfas-2012-0401, 2013.

Dask Development Team: Dask: Library for dynamic task scheduling, https://dask.org, software, 2016.

Delpech, A., Conchon, A., Titaud, O., and Lehodey, P.: Influence of oceanic conditions in the energy transfer efficiency estimation of a micronekton model, Biogeosciences, 17, 833–850, https://doi.org/10.5194/bg-17-833-2020, 2020.

Drago, L., Panaïotis, T., Irisson, J.-O., Babin, M., Biard, T., Carlotti, F., Coppola, L., Guidi, L., Hauss, H., Karp-Boss, L., Lombard, F., McDonnell, A. M. P., Picheral, M., Rogge, A., Waite, A. M., Stemmann, L., and Kiko, R.: Global Distribution of Zooplankton Biomass Estimated by In Situ Imaging and Machine Learning, Front. Mar. Sci., 9, https://doi.org/10.3389/fmars.2022.894372, 2022.

Edvardsen, A., Pedersen, J. M., Slagstad, D., Semenova, T., and Timonin, A.: Distribution of overwintering Calanus in the North Norwegian Sea, Ocean Sci., 2, 87–96, https://doi.org/10.5194/os-2-87-2006, 2006.

E.U. Copernicus Marine Service Information (CMEMS): Global Ocean Low and Mid Trophic Levels Biomass Content Hindcast, https://doi.org/10.48670/moi-00020, product GLOBAL_MULTIYEAR_BGC_001_033, 2024.

Gillooly, J. F.: Effect of body size and temperature on generation time in zooplankton, J. Plankton Res., 22, 241–251, https://doi.org/10.1093/plankt/22.2.241, 2000.

Gillooly, J. F., Charnov, E. L., West, G. B., Savage, V. M., and Brown, J. H.: Effects of size and temperature on developmental time, Nature, 417, 70–73, https://doi.org/10.1038/417070a, 2002.

Hansen, N.: The CMA Evolution Strategy: A Tutorial, arXiv preprint arXiv:1604.00772, https://doi.org/10.48550/arXiv.1604.00772, 2016.

Hansen, N. and Ostermeier, A.: Completely Derandomized Self-Adaptation in Evolution Strategies, Evol. Comput., 9, 159–195, https://doi.org/10.1162/106365601750190398, 2001.

Hansen, N., Akimoto, Y., and Baudis, P.: CMA-ES/pycma on Github, Zenodo, DOI:10.5281/zenodo.2559634, https://doi.org/10.5281/zenodo.2559634, 2019.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E.: Array programming with NumPy, Nature, 585, 357–362, https://doi.org/10.1038/s41586-020-2649-2, 2020.

Herman, J. and Usher, W.: SALib: An open-source Python library for Sensitivity Analysis, J. Open Source Softw., 2, 97, https://doi.org/10.21105/joss.00097, 2017.

Hoffman, R. N. and Atlas, R.: Future Observing System Simulation Experiments, Bull. Amer. Meteor. Soc., 97, 1601–1616, https://doi.org/10.1175/BAMS-D-15-00200.1, 2016.

Hoyer, S. and Hamman, J.: xarray: N-D labeled Arrays and Datasets in Python, J. Open Res. Softw., 5, 10, https://doi.org/10.5334/jors.148, 2017.

Huntley, M. E. and Lopez, M. D. G.: Temperature-Dependent Production of Marine Copepods: A Global Synthesis, Am. Nat., 140, 201–242, https://doi.org/10.1086/285410, 1992.

Iwanaga, T., Usher, W., and Herman, J.: Toward SALib 2.0: Advancing the accessibility and interpretability of global sensitivity analyses, Socio-Environ. Syst. Model., 4, 18 155, https://doi.org/10.18174/sesmo.18155, 2022.

Karl, D. M. and Lukas, R.: The Hawaii Ocean Time-series (HOT) program: Background, rationale and field implementation, Deep-Sea Res. Pt. II, 43, 129–156, https://doi.org/10.1016/0967-0645(96)00005-7, 1996.

Kosobokova, K. N.: The reproductive cycle and life history of the Arctic copepod Calanus glacialis in the White Sea, Polar Biol., 22, 254–263, https://doi.org/10.1007/s003000050418, 1999.

Lam, S. K., Pitrou, A., and Seibert, S.: Numba: a LLVM-based Python JIT compiler, in: Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC, SC15, p. 1–6, ACM, https://doi.org/10.1145/2833157.2833162, 2015.

Lehodey, J. V.: SeapoPym/seapopym: SeapoPym v0.1.1, Zenodo, https://doi.org/10.5281/zenodo.21838667, software, 2026.

Lehodey, J. V., Mignot, A., Ganachaud, A., Albernhe, S., and Nicol, S.: Reproducibility deposit for SeapoPym v0.1: Implementation of the SEAPODYM low and mid trophic levels in Python with a flexible optimization framework, Zenodo, https://doi.org/10.5281/zenodo.21849337, software, 2026.

Lehodey, P.: The pelagic ecosystem of the tropical Pacific Ocean: dynamic spatial modelling and biological consequences of ENSO, Prog. Oceanogr., 49, 439–468, https://doi.org/10.1016/s0079-6611(01)00035-0, 2001.

Lehodey, P., André, J.-M., Bertignac, M., Hampton, J., Stoens, A., Menkes, C., Memery, L., and Grima, N.: Predicting skipjack tuna forage distributions in the equatorial Pacific using a coupled dynamical bio-geochemical model, Fish. Oceanogr., 7, 317–325, https://doi.org/10.1046/j.1365-2419.1998.00063.x, 1998.

Lehodey, P., Senina, I., and Murtugudde, R.: A spatial ecosystem and populations dynamics model (SEAPODYM) – Modeling of tuna and tuna-like populations, Prog. Oceanogr., 78, 304–318, https://doi.org/10.1016/j.pocean.2008.06.004, 2008.

Lehodey, P., Murtugudde, R., and Senina, I.: Bridging the gap from ocean models to population dynamics of large marine predators: A model of mid-trophic functional groups, Prog. Oceanogr., 84, 69–84, https://doi.org/10.1016/j.pocean.2009.09.008, 2010.

Lehodey, P., Conchon, A., Senina, I., Domokos, R., Calmettes, B., Jouanno, J., Hernandez, O., and Kloser, R.: Optimization of a micronekton model with acoustic data, ICES J. Mar. Sci., 72, 1399–1412, https://doi.org/10.1093/icesjms/fsu233, 2015.

Lellouche, J.-M., Greiner, E., Bourdallé-Badie, R., et al.: The Copernicus Global 1/12° Oceanic and Sea Ice GLORYS12 Reanalysis, Frontiers in Earth Science, 9, 698 876, https://doi.org/10.3389/feart.2021.698876, 2021.

Mackas, D. and Coyle, K.: Shelf–offshore exchange processes, and their effects on mesozooplankton biomass and community composition patterns in the northeast Pacific, Deep-Sea Res. Pt. II, 52, 707–725, https://doi.org/10.1016/j.dsr2.2004.12.020, 2005.

McKendrick, A. G.: Applications of Mathematics to Medical Problems, Proceedings of the Edinburgh Mathematical Society, 44, 98–130, https://doi.org/10.1017/s0013091500034428, 1926.

Moriarty, R. and O'Brien, T. D.: Distribution of mesozooplankton biomass in the global ocean, Earth Syst. Sci. Data, 5, 45–55, https://doi.org/10.5194/essd-5-45-2013, 2013.

O'Brien, T. D.: COPEPOD: The Global Plankton Database. An overview of the 2014 database contents, processing methods, and access interface, Tech. Rep. NOAA Tech. Memo. NMFS-F/ST-37, U.S. Department of Commerce, NOAA, https://www.st.nmfs.noaa.gov/copepod/, 2014.

Omori, M.: Weight and chemical composition of some important oceanic zooplankton in the North Pacific Ocean, Mar. Biol., 3, 4–10, https://doi.org/10.1007/BF00355587, 1969.

Pianosi, F., Beven, K., Freer, J., Hall, J. W., Rougier, J., Stephenson, D. B., and Wagener, T.: Sensitivity analysis of environmental models: A systematic review with practical workflow, Environ. Model. Softw., 79, 214–232, https://doi.org/10.1016/j.envsoft.2016.02.008, 2016.

Rohr, T., Richardson, A. J., Lenton, A., Chamberlain, M. A., and Shadwick, E. H.: Zooplankton grazing is the largest source of uncertainty for marine carbon cycling in CMIP6 models, Communications Earth & Environment, 4, 212, https://doi.org/10.1038/s43247-023-00871-w, 2023.

Saltelli, A.: Making best use of model evaluations to compute sensitivity indices, Comput. Phys. Commun., 145, 280–297, https://doi.org/10.1016/s0010-4655(02)00280-1, 2002.

Sarrazin, F., Pianosi, F., and Wagener, T.: Global Sensitivity Analysis of environmental models: Convergence and validation, Environ. Model. Softw., 79, 135–152, https://doi.org/10.1016/j.envsoft.2016.02.005, 2016.

Schartau, M., Wallhead, P., Hemmings, J., Löptien, U., Kriest, I., Krishna, S., Ward, B. A., Slawig, T., and Oschlies, A.: Reviews and syntheses: parameter identification in marine planktonic ecosystem modelling, Biogeosciences, 14, 1647–1701, https://doi.org/10.5194/bg-14-1647-2017, 2017.

Sibert, J. R., Hampton, J., Fournier, D. A., and Bills, P. J.: An advection-diffusion-reaction model for the estimation of fish movement parameters from tagging data, with application to skipjack tuna (Katsuwonus pelamis), Can. J. Fish. Aquat. Sci., 56, 925–938, https://doi.org/10.1139/f99-017, 1999.

Sieburth, J. M., Smetacek, V., and Lenz, J.: Pelagic ecosystem structure: Heterotrophic compartments of the plankton and their relationship to plankton size fractions 1, Limnol. Oceanogr., 23, 1256–1263, https://doi.org/10.4319/lo.1978.23.6.1256, 1978.

Siegel, D. A., DeVries, T., Cetinić, I., and Bisson, K. M.: Quantifying the Ocean's Biological Pump and Its Carbon Cycle Impacts on Global Scales, Annu. Rev. Mar. Sci., 15, 329–356, https://doi.org/10.1146/annurev-marine-040722-115226, 2023.

Sobol', I.: Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, Math. Comput. Simul., 55, 271–280, https://doi.org/10.1016/s0378-4754(00)00270-6, 2001.

St. John, M. A., Borja, A., Chust, G., Heath, M., Grigorov, I., Mariani, P., Martin, A. P., and Santos, R. S.: A Dark Hole in Our Understanding of Marine Ecosystems and Their Services: Perspectives from the Mesopelagic Community, Front. Mar. Sci., 3, https://doi.org/10.3389/fmars.2016.00031, 2016.

Steinberg, D. K., Carlson, C. A., Bates, N. R., Goldthwait, S. A., Madin, L. P., and Michaels, A. F.: Zooplankton vertical migration and the active transport of dissolved organic and inorganic carbon in the Sargasso Sea, Deep-Sea Res. Pt. I, 47, 137–158, https://doi.org/10.1016/s0967-0637(99)00052-7, 2000.

Steinberg, D. K., Carlson, C. A., Bates, N. R., Johnson, R. J., Michaels, A. F., and Knap, A. H.: Overview of the US JGOFS Bermuda Atlantic Time-series Study (BATS): a decade-scale look at ocean biology and biogeochemistry, Deep-Sea Res. Pt. II, 48, 1405–1447, https://doi.org/10.1016/S0967-0645(00)00148-X, 2001.

Steinberg, D. K., Goldthwait, S. A., and Hansell, D. A.: Zooplankton vertical migration and the active transport of dissolved organic and inorganic nitrogen in the Sargasso Sea, Deep-Sea Res. Pt. I, 49, 1445–1461, https://doi.org/10.1016/s0967-0637(02)00037-7, 2002.

Stukel, M. R., Décima, M., and Landry, M. R.: Quantifying biological carbon pump pathways with a data-constrained mechanistic model ensemble approach, Biogeosciences, 19, 3595–3624, https://doi.org/10.5194/bg-19-3595-2022, 2022.

Tarantola, A.: Inverse Problem Theory and Methods for Model Parameter Estimation, Society for Industrial and Applied Mathematics, Philadelphia, PA, ISBN 978-0-89871-572-9, https://doi.org/10.1137/1.9780898717921, 2005.

Titaud, O., Conchon, A., Mérillet, L., Goeman, N., and Magliano, H.: Quality Information Document, Global Ocean Low and Mid Trophic Levels Biomass Content Hindcast (GLOBAL_MULTIYEAR_BGC_001_033), Tech. Rep. CMEMS-GLO-QUID-001-033, Issue 4.2, Copernicus Marine Service, https://documentation.marine.copernicus.eu/QUID/CMEMS-GLO-QUID-001-033.pdf, 2024.

Von Foerster, H.: Some Remarks on Changing Populations, in: The Kinetics of Cellular Proliferation, edited by Stohlman, J. F., pp. 382–407, Grune and Stratton, New York, 1959.

**Table 1.** Biological parameters of the mesozooplankton group: symbol, meaning, operational reference value with its unit, the range explored in the sensitivity and twin experiments (Sects. 2.5 and 2.6), the origin of the reference value, and the Appendix A equation in which each parameter appears. The reference temperature is $T_{\mathrm{ref}} = 0\,^{\circ}\mathrm{C}$. The recruitment age $\tau_r$ and the mortality rate $\lambda$ both vary with temperature as $\exp(\gamma \bar{T})$ (Appendix A), so $\lambda$ rises and $\tau_r$ falls as temperature increases.

| Symbol | Meaning | Reference | Range | Origin | Eq. |
|---|---|---|---|---|---|
| $E$ | Energy-transfer coefficient, the fraction of NPP entering the youngest production class | 0.1668 | 0 to 0.5 | Maximum-likelihood estimate against the COPEPOD database (Titaud et al., 2024) | A5 |
| $\tau_{r_0}$ | Recruitment age at the reference temperature | 10.38 d | 0.001 to 300 d | Copepod development and temperature relationship (Huntley and Lopez, 1992; Gillooly et al., 2002; Lehodey et al., 2010) | A6 |
| $\gamma_{\tau_r}$ | Temperature coefficient of the recruitment age | $-0.11\,^{\circ}\mathrm{C}^{-1}$ | $-0.25$ to $-0.001\,^{\circ}\mathrm{C}^{-1}$ | As above | A6 |
| $\lambda_0$ | Mortality rate at the reference temperature | $0.0067\,\mathrm{d}^{-1}$ | 0.001 to $0.1\,\mathrm{d}^{-1}$ | Operational value manually tuned (Conchon, 2016), temperature form from metabolic theory (Gillooly et al., 2002) | A2 |
| $\gamma_\lambda$ | Temperature coefficient of the mortality rate | $0.15\,^{\circ}\mathrm{C}^{-1}$ | 0.001 to $0.25\,^{\circ}\mathrm{C}^{-1}$ | As above | A2 |

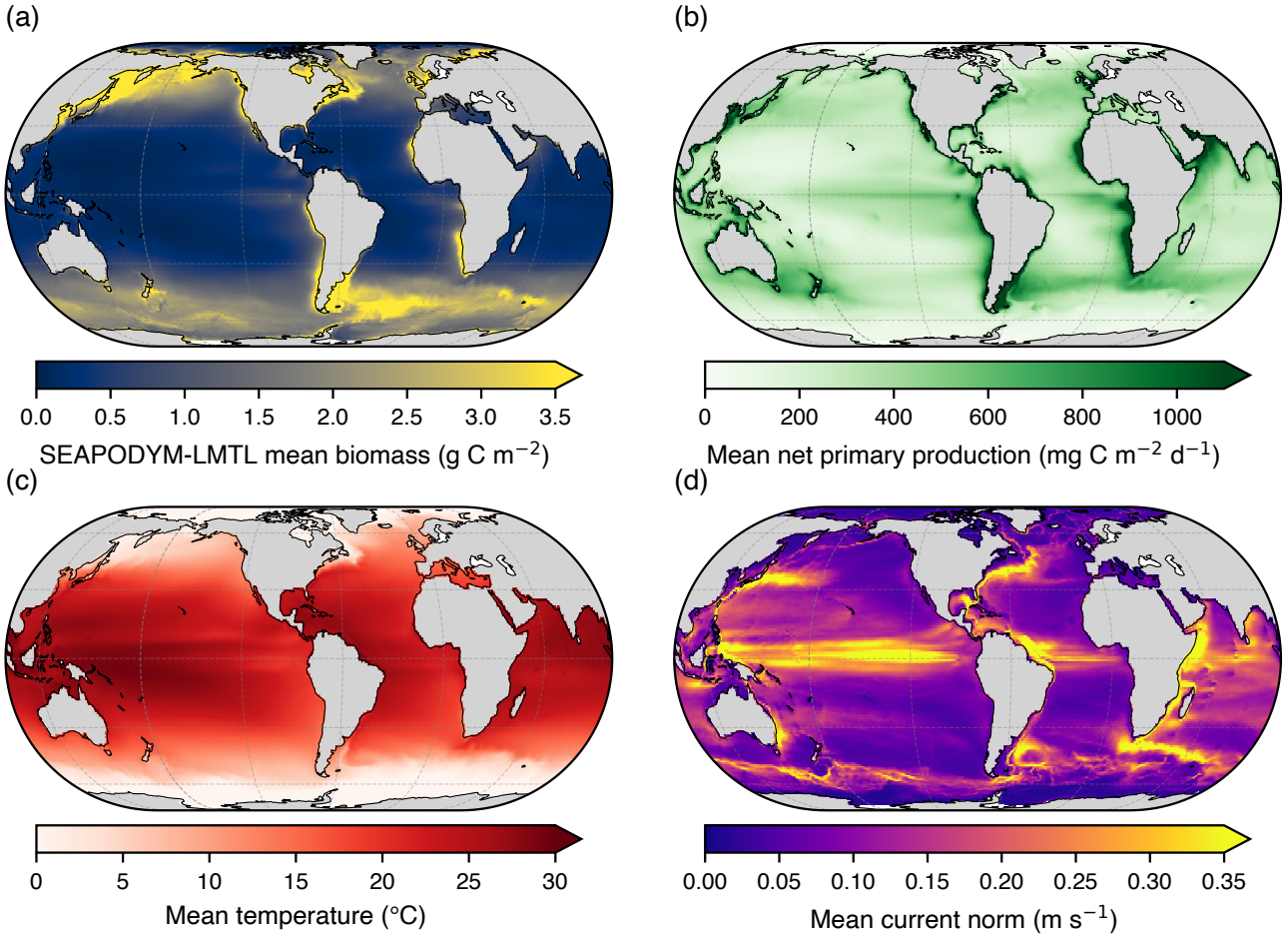


**Figure 1.** Global ocean maps of the time-mean forcing and the reference biomass, averaged over 2000 to 2019. (a) Mean mesozooplankton biomass of the operational SEAPODYM-LMTL product, the two-dimensional reference that SeapoPym is compared against in Fig. 4. (b) Mean vertically integrated NPP. (c) Mean temperature of the epipelagic layer. (d) Mean current norm of the epipelagic layer.

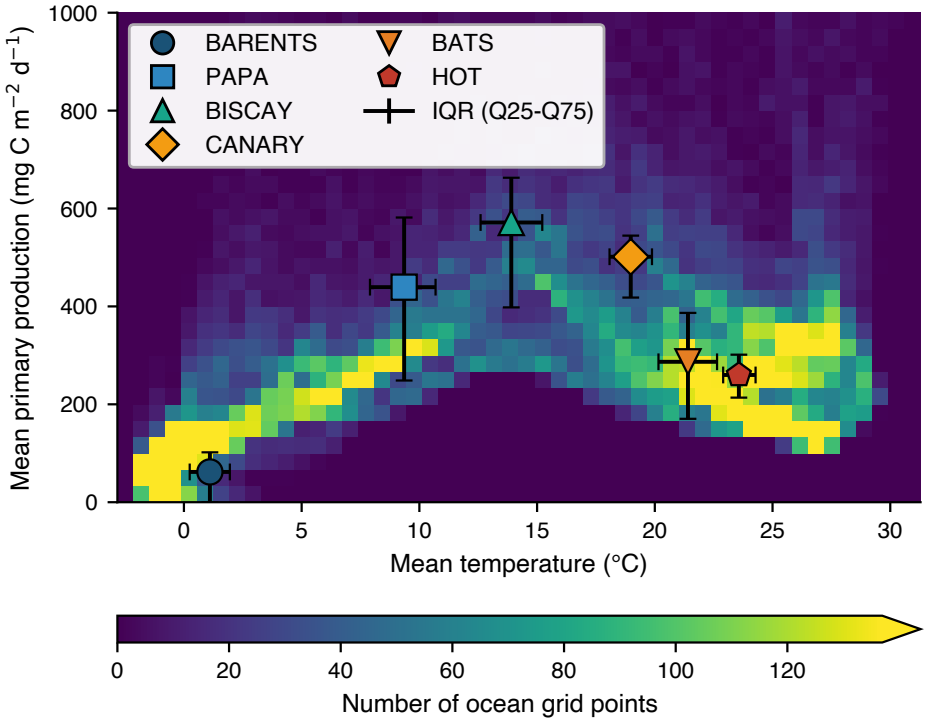


**Figure 2.** The six stations placed in the global temperature and primary-production space. The color field is a two-dimensional histogram of the ocean grid cells, binned by their 2000 to 2019 mean temperature and mean NPP, the color giving the number of cells in each bin. Each marker is the mean position of one station, and the bars span its interquartile range (25th to 75th percentile) in temperature (horizontal) and in production (vertical) over the same period. Station color encodes mean temperature from cold blue to warm red, a convention kept across every figure.

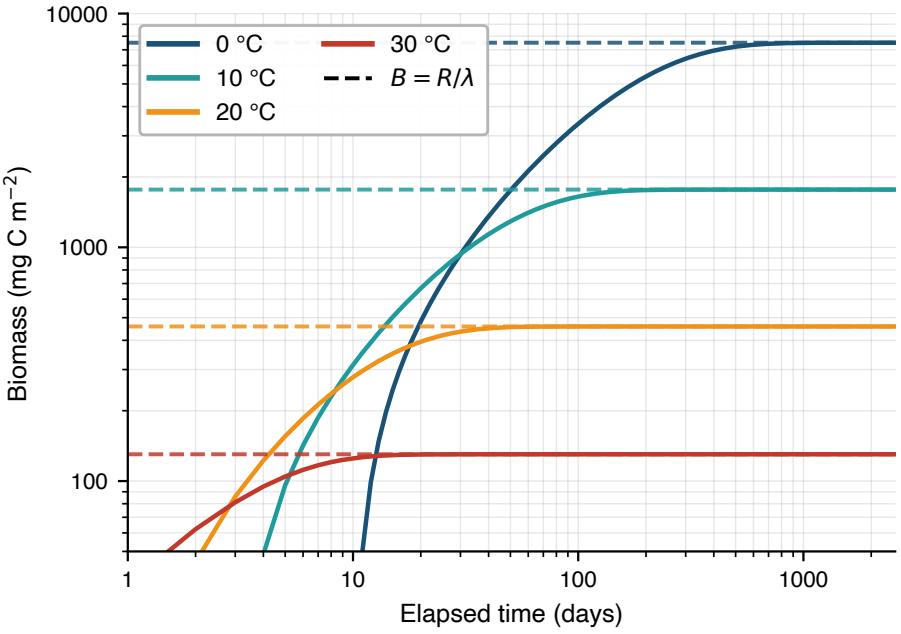


**Figure 3.** Convergence of the SeapoPym biomass to its analytical steady state under constant forcing. Each solid line is the simulated biomass at one of four constant temperatures (0, 10, 20 and 30 °C), with NPP held at 300 mg C m$^{-2}$ d$^{-1}$, and the matching dashed line is the analytical equilibrium $B_{\mathrm{eq}} = R/\lambda$ (Appendix A). Line color encodes temperature. Both axes are logarithmic. The biomass reaches the analytical value at every temperature, within 0.01 percent. The time to approach it lengthens as temperature falls, reaching within 1 percent in about fifteen days at 30 °C and about two years at 0 °C.

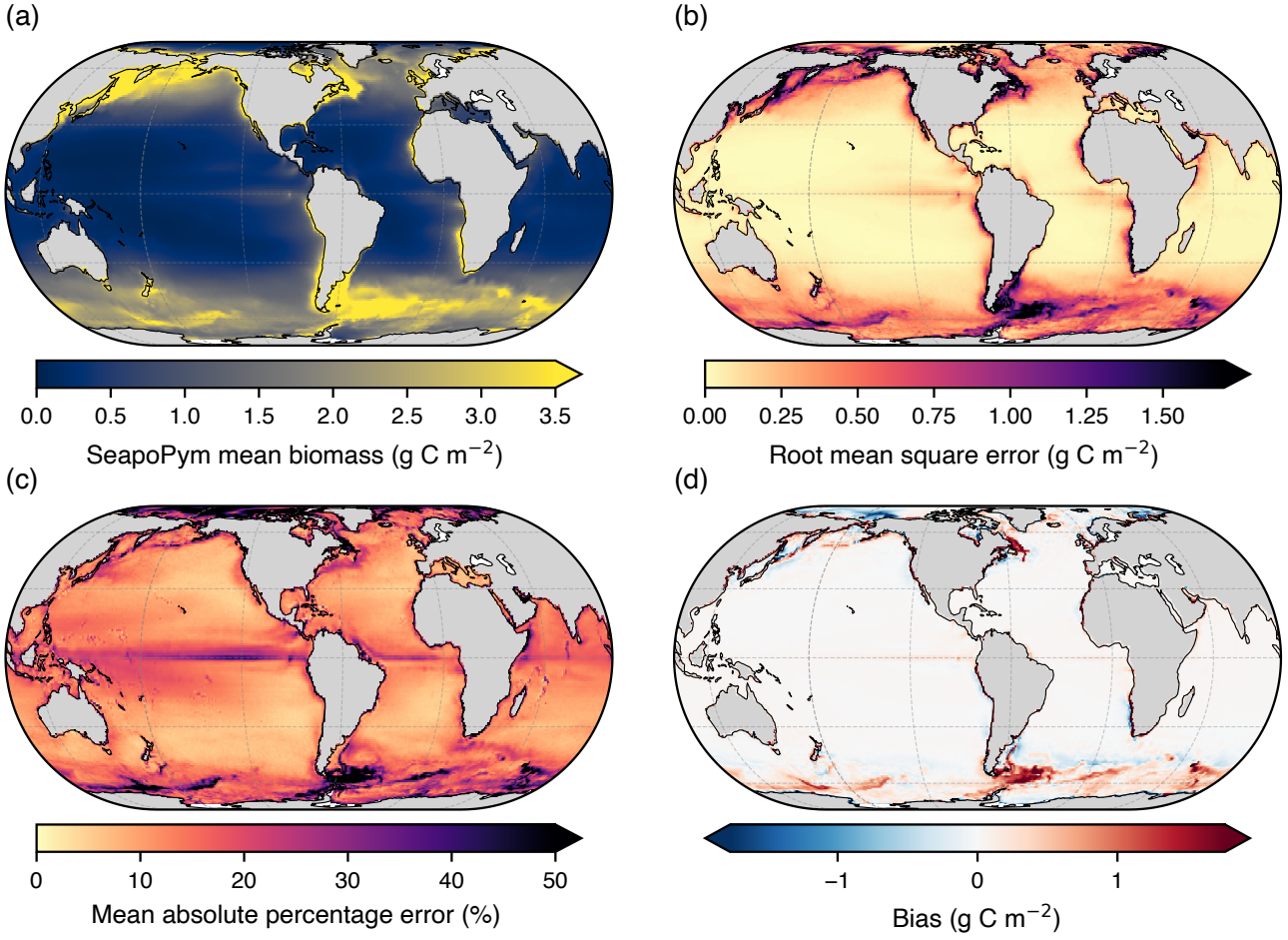


**Figure 4.** Effect of neglecting transport, shown as SeapoPym (0D) against the operational SEAPODYM-LMTL product (2D), as time means over 2000 to 2019. (a) SeapoPym mean biomass, on the same scale as the reference biomass in Fig. 1a. (b) Root mean square error between the two models. (c) Mean absolute percentage error, the difference taken relative to the local biomass. (d) Bias, the mean signed difference of SeapoPym minus the 2D reference. The difference is largest along the western-boundary currents and the Antarctic Circumpolar Current and in cold high-latitude waters, and smallest in the warm subtropical gyres.

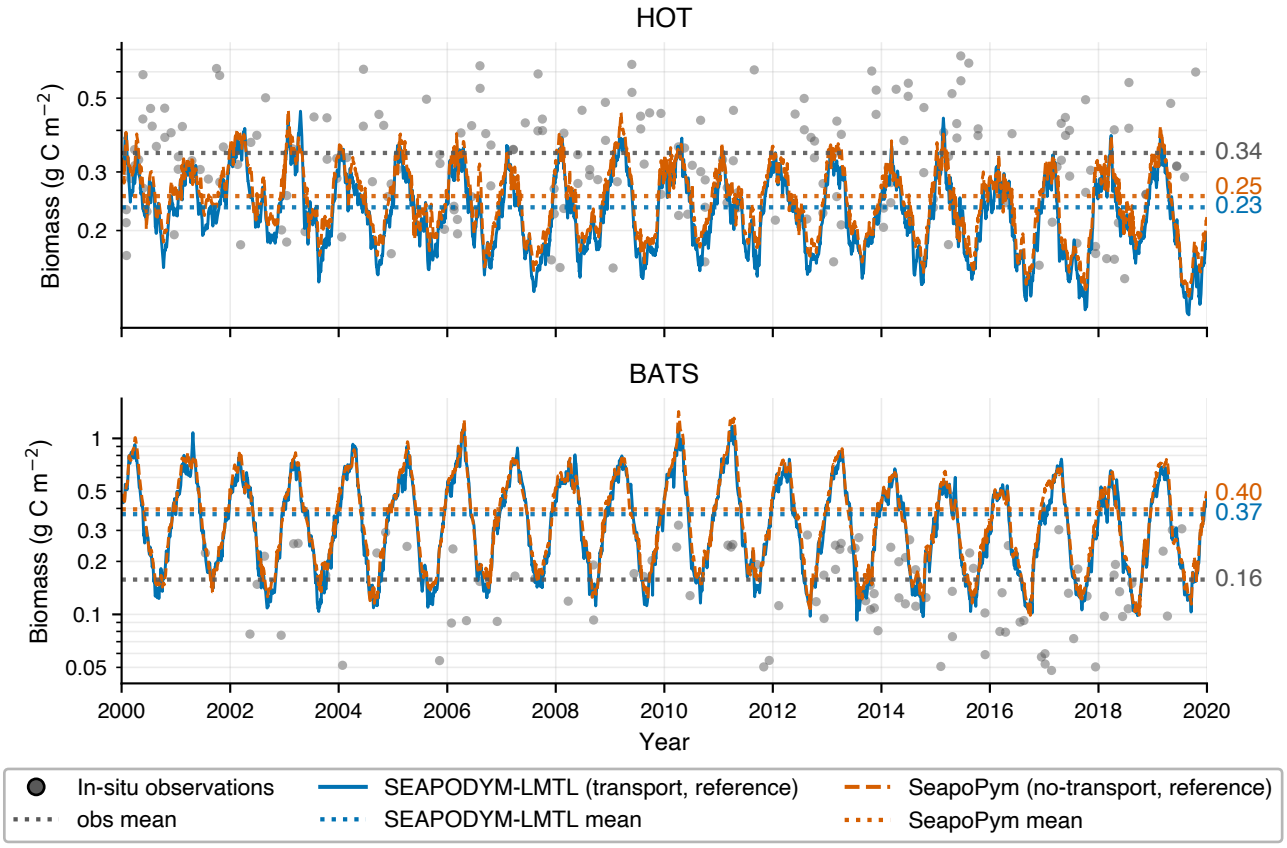


**Figure 5.** Model biomass against in-situ observations at the two stations with open long-term records, HOT (top) and BATS (bottom). Grey dots are the daily-mean in-situ mesozooplankton biomass, clipped to the 5th to 95th percentile. The solid blue line is the operational SEAPODYM-LMTL product and the dashed orange line is SeapoPym (without transport), both driven by the same forcing and parameters. The dotted horizontal lines mark the time-mean of each series, with their values printed at the right. The y-axis is logarithmic and biomass is in g C m$^{-2}$. The two model means nearly coincide while the observed mean sits apart, so the difference between the models, which is due to transport and a different implementation, is five to seven times smaller than the gap between the operational SEAPODYM-LMTL product and the observations.

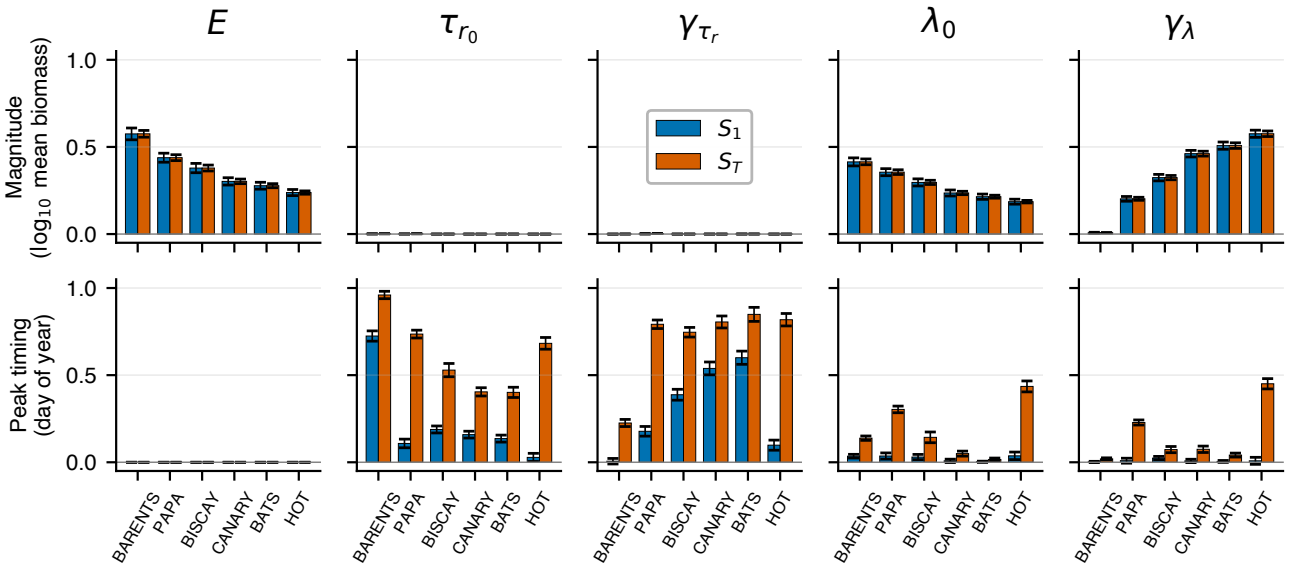


**Figure 6.** Sobol sensitivity indices of the biomass to the five biological parameters at the six stations, ordered from coldest to warmest. The rows are the two output descriptors, the magnitude (base-10 logarithm of the mean biomass, top) and the seasonal peak timing (day of year of the maximum, bottom). The columns are the five parameters $E$, $\tau_{r_0}$, $\gamma_{\tau_r}$, $\lambda_0$ and $\gamma_\lambda$. In each panel the blue bars are the first-order index $S_1$, the share of output variance due to that parameter alone, and the orange bars the total-order index $S_T$, which adds its interactions with the other parameters. Error bars are 95 percent bootstrap confidence intervals. The indices are computed from a Saltelli design of $N = 8192$ base samples, that is 57 344 model runs at each station. The magnitude is governed by the energy transfer and the two mortality parameters, the peak timing by the two recruitment parameters.

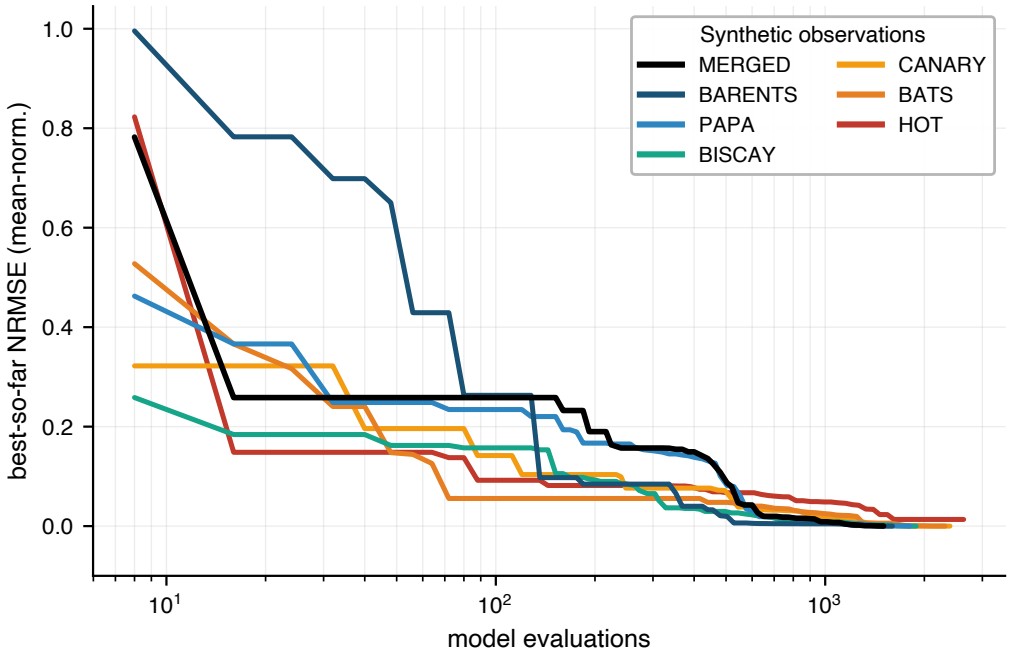


**Figure 7.** CMA-ES convergence for the seven twin experiments, each the best of twenty random restarts. Each line is the best-so-far cost against the number of model evaluations, for the six single-station experiments and the joint MERGED experiment. The cost is the NRMSE between the recovered and the synthetic biomass, normalized by the mean of the target series. The x-axis is logarithmic. Station color follows the cold-to-warm convention and MERGED is black. Every experiment drives the cost to near zero, about $10^{-4}$, except HOT, which floors near $10^{-2}$. Because the target is the model's own output a zero cost is attainable, so the HOT residual likely reflects an incomplete search rather than a structural floor.

**Table 2.** Parameter recovery in the twin experiments, where the optimizer recovers known parameters from synthetic observations that the model generated itself, taking the best of twenty CMA-ES restarts. Each entry is the signed relative divergence of a recovered parameter against its synthetic-observation reference, in percent. The last column is the achieved cost (mean-normalized NRMSE). The cold stations and MERGED recover all five parameters, CANARY and BATS miss the two recruitment parameters, and HOT does not reach the optimum. A large error next to a near-zero cost is equifinality.

| Synthetic observation | $E$ | $\tau_{r_0}$ | $\gamma_{\tau_r}$ | $\lambda_0$ | $\gamma_\lambda$ | NRMSE |
|---|---|---|---|---|---|---|
| (station) | $0.1668$ | $10.38$ | $-0.11$ | $0.006667$ | $0.15$ | |
| MERGED | $0$ | $0$ | $0$ | $0$ | $0$ | $1.1 \times 10^{-4}$ |
| BARENTS | $0$ | $0$ | $0$ | $0$ | $0$ | $4.9 \times 10^{-6}$ |
| PAPA | $0$ | $0$ | $0$ | $0$ | $0$ | $3.5 \times 10^{-5}$ |
| BISCAY | $0$ | $+0.1$ | $0$ | $0$ | $0$ | $1.4 \times 10^{-5}$ |
| CANARY | $0$ | $-77$ | $+78$ | $0$ | $0$ | $1.9 \times 10^{-8}$ |
| BATS | $0$ | $+158$ | $-41$ | $0$ | $0$ | $1.7 \times 10^{-6}$ |
| HOT | $+0.1$ | $-0.5$ | $+0.2$ | $+27$ | $-7.4$ | $1.3 \times 10^{-2}$ |

**Table B1.** Per-station summary, stations ordered by increasing mean temperature. For each station: position, mean temperature of the epipelagic layer, the recruitment age $\tau_r = \tau_{r_0} \exp(\gamma_{\tau_r} \bar{T})$ (Table 1, Appendix A) at the mean and at the warmest point of the forcing series, in days and equivalently in daily cohorts, the difference between the 0D SeapoPym and the operational SEAPODYM-LMTL product (RMSE and MAPE sampled from the maps of Fig. 4 over 2000 to 2019), and, at the two stations with in-situ records, the structural gap between the 2D reference and the observations (RMSE) with its ratio to the 0D-to-2D difference. RMSE in $\mathrm{g\,C\,m^{-2}}$, MAPE in percent. The ratio is computed from the daily series as in Fig. 5, with the observations clipped to their 5th–95th percentiles.

| | | $T$ | $\tau_r$ (d) | | 0D–2D | | obs–2D | |
|---|---|---|---|---|---|---|---|---|
| Station | Position | (°C) | mean | min | RMSE | MAPE | RMSE | ratio |
| BARENTS | 75°N, 40°E | 1.1 | 9.18 | 6.59 | 0.133 | 8.5 | – | – |
| PAPA | 50°N, 132°W | 9.4 | 3.90 | 2.50 | 0.402 | 11.4 | – | – |
| BISCAY | 45.5°N, 4°W | 13.9 | 2.45 | 1.72 | 0.176 | 6.6 | – | – |
| CANARY | 30°N, 13°W | 19.0 | 1.48 | 1.11 | 0.083 | 6.9 | – | – |
| BATS | 32°N, 64°W | 21.4 | 1.18 | 0.78 | 0.058 | 10.8 | 0.316 | 5.5 |
| HOT | 23°N, 158°W | 23.6 | 0.96 | 0.77 | 0.026 | 9.8 | 0.183 | 7.1 |

**Table C1.** Abbreviations and acronyms used in this paper.

| Abbreviation | Definition |
|---|---|
| 0D | Zero-dimensional, the local SeapoPym model with transport neglected |
| 2D | Two-dimensional, the operational SEAPODYM-LMTL product with transport resolved |
| CMA-ES | Covariance matrix adaptation evolution strategy |
| CMEMS | Copernicus Marine Environment Monitoring Service |
| COPEPOD | Coastal and Oceanic Plankton Ecology, Production and Observation Database |
| GLORYS12 | Global ocean reanalysis at 1/12° resolution, the source of the temperature forcing |
| LMTL | The component of SEAPODYM which models the Low and Mid-Trophic Levels |
| MAPE | Mean absolute percentage error |
| NPP | Net primary production |
| NRMSE | Normalized root mean square error, here normalized by the mean |
| RMSE | Root mean square error |
| SEAPODYM | Spatial Ecosystem and Populations Dynamics Model |
| SeapoPym | The open Python re-implementation of the SEAPODYM-LMTL biology presented here |
| VGPM | Vertically Generalized Production Model |

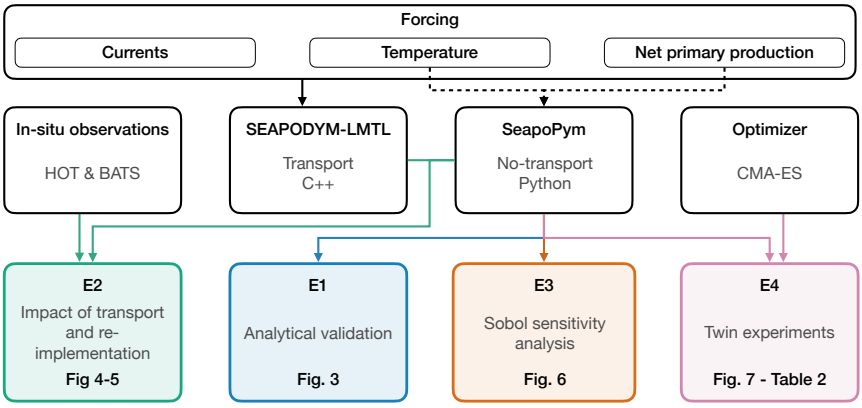


**Figure D1.** Overview of the SeapoPym modeling and experimental workflow. Temperature and NPP drive both models, whereas the currents drive only the SEAPODYM-LMTL model (with transport, C++). The dashed link marks that SeapoPym (without transport, Python) omits the currents. Each experiment is color-coded and linked to the components it uses, with the figures that report it. E1 (analytical validation, Fig. 3) and E3 (Sobol sensitivity analysis, Fig. 6) exercise the SeapoPym model. E2 (impact of transport and re-implementation, Figs. 4–5) compares SeapoPym with the operational SEAPODYM-LMTL product and with the in-situ observations at HOT and BATS. E4 (twin experiments, Fig. 7 and Table 2) recovers the SeapoPym parameters with the CMA-ES optimizer from synthetic observations generated by SeapoPym itself.