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

Jules V. 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), Nouméa, New Caledonia
[4]Mercator Ocean International, Toulouse, France
[5]Pacific Community (SPC), Noumea, New Caledonia

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

Abstract. SEAPODYM-LMTL~~is a global~~ , the low and mid trophic level component of SEAPODYM, simulates mesozooplankton and micronekton biomass globally as an advection-diffusion-reaction ~~model that simulates age-structured zooplankton and micronekton populations~~ system driven by physical and biogeochemical forcing. ~~This study introduces SeapoPym, a simplified version of this model that decouples biological dynamics from physical transportand incorporates a Genetic Algorithm (GA)for stochastic parameter estimation within the Python scientific ecosystem. Comparisons with~~ 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 ~~show~~ 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 ~~produces notable discrepancies in highly dynamic warm regions and cold environments with~~ matters most in strongly advective regions and in cold high-latitude waters, where the long zooplankton life ~~cycles. However, SeapoPym remains suitable for simulating mesozooplankton across most warm- and temperate-ocean regions. Sobol sensitivity analysis identifies mortality parameters as key drivers of biomass magnitude and variability, with strong parameter interactions~~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 ~~highlight challenges in estimating recruitment-timing parameters and emphasize the importance of data from cold, contrasting environments. SeapoPym provides a flexible, low-cost framework for exploring parameter estimation, designing observational strategies, and addressing challenges in zooplankton model assessment, with the potential to integrate with circulation models or machine-learning emulators~~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 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

~~Marine mesozooplankton, defined as organisms typically~~

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 ~~2 mm in size (Sieburth et al., 1978), represent a pivotal node in pelagic food webs. While primarily composed of copepods, this diverse functional group also encompasses various planktonic larval forms (meroplankton) (Moriarty and O'Brien, 2013; Drago et al., 2022). As a key conduit of energy from lower trophic levels, mesozooplankton contribute substantially to the sustenance of micronekton—a community of larger organisms~~ 20 mm (Sieburth et al., 1978), dominated by copepods, although the size class is taxonomically diverse (Moriarty and O'Brien, 2 . They provide a major food source for micronekton (2–~~20 cm)~~20 cm), including mesopelagic fish, euphausiids ~~, and gelatinous species (Brodeur et al., 2005; St. John et al., 2016). Given this critical role in supporting higher predators and regulating biogeochemical cycles, accurate modeling of zooplankton dynamics is essential to predict ecosystem responses to accelerating climate change, acidification, and deoxygenation (Rohr et al., 2023; Steinberg et al., 2000, 2002; Stukel et al., 2022; Siegel et al., 2023).~~ 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).

~~A major challenge for the scientific community is the lack of consistent time-series observations spanning large scales, particularly for zooplankton and micronekton (Holland et al., 2025). There are significant gaps in both spatial and temporal coverage, as well as shortcomings in taxonomic representation (Ratnarajah et al., 2023). While traditional net trawling remains the standard method for sampling, new technologies are poised to quickly increase the number of observations (Holland et al., 2025) . In the meantime, enhancing the modeling of diverse key functional groups and species, along with the development of efficient parameter-estimation methods, becomes crucial (Schartau et al., 2017). One approach to these issues~~ The SEAPODYM-LMTL model provides operational global simulations of mesozooplankton and micronekton biomass. It is the Low and Mid-Trophic Levels ~~modeling component of the Spatial Ecosystem and Population Dynamics Model (SEAPODYM-LMTL)~~(LMTL) component of SEAPODYM, a model developed to simulate tuna and tuna-like populations (Lehodey et al., 2008). SEAPODYM-LMTL ~~mimics the distribution patterns of marine zooplankton and micronekton over space and time on a global scale, offering Essential Ocean Variables useful for modeling fish populations (Lehodey et al., 2008, 2010b, 2015b; Conchon, 2016; Delpech et al., 2020) . The model employs a system of equations that accounts for advection , diffusion , and reaction processes, using physical data such as~~ represents their prey environment through one mesozooplankton and six micronekton functional groups (Lehodey et al., 2010a; Con

. It couples biological processes with advection and diffusion and is driven by currents, temperature, ~~and euphotic depth , as well as biogeochemical information, such as~~ euphotic depth and net primary production (NPP)~~, derived from satellite data or other models. The water column is divided into several layers, with functional groups separated by their vertical movement strategies at different times of day. Currently, SEAPODYM-LMTL includes one group of mesozooplankton and six groups of micronekton, all of which are accessible through high-resolution global outputs available from~~ . The resulting biomass fields are distributed through the Copernicus Marine Service ~~(https://data. marine.copernicus.eu).~~ for ecosystem and fisheries applications (Titaud et al., 2024).

~~Addressing these implementation challenges is crucial for fostering innovation in model development and for enabling more efficient exploration of biological hypotheses. The existing Advection-Diffusion-Reaction system within the~~ 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 . When compared to in-situ observations (direct mesozooplankton biomass measurements) from the COPEPOD database (O'Brien, 2014), the operational SEAPODYM-LMTL ~~model, implemented in C~~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++~~, presents challenges for rapid hypothesis testing. In the current C+ + implementation , biological reaction terms are tightly coupled with the transport scheme within the numerical solver; consequently, any modification to a biological mechanism requires significant changes to the resolution matrices. This structural rigidity makes code adaptation complex and time-consuming, especially as the number of simulated groups and species increases~~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.

~~To overcome these limitations, we introduce a new modeling framework. This study presents SeapoPym, a simplified version of~~ 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 re-implementation of the SEAPODYM-LMTL ~~that decouples biological mechanisms from the complex transport infrastructure and leverages the Python scientific ecosystem. SeapoPym is designed as a flexible platform to address the need for rapid model development and testing. Additionally, SeapoPym incorporates a parameter-estimation module based on a Genetic Algorithm (GA), enabling stochastic optimization and maintaining independence between model design and parameter estimation.~~

~~This paper details the methodological development and validation strategy of the SeapoPym framework, with a specific focus on mesozooplankton. Numerical accuracy is demonstrated through theoretical benchmarks and comparison with the~~

~~reference~~ 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 ~~simulation. The model's structural behavior is then examined using~~ 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~~(Sobol', 2001) to assess its response to varying environmental conditions. The optimization framework's ability to recover parameters is evaluated using a twin experiment, verifying convergence to a global solution across diverse scenarios. The final section discusses model performance, optimization efficiency, and implications for future research in marine ecosystem modeling~~, 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~~s~~and ~~Method~~s~~methods~~

2.1 ~~Seapo~~Pym ~~~~SEAPODYM-LMTL (the reference~~ model~~)~~

~~SeapoPym is implemented as an open-source Python library released under the GPLv3 license (Lehodey, 2026). For this study, which focuses on the model's development and validation, we limit the analysis to a single zooplankton group in~~

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 ~~epipelagic layer. However, the framework is generic and can be applied to micronekton groups by parameterizing their vertical partitioning between layers over the diel cycle(as in~~ 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., 2010a; Conchon, 2016). It simulates their biomass as a system of advection-diffusion-reaction equations driven by physical and biogeochemical forcing (Lehodey et al., 2010a). 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.

~~Conceptually, SeapoPym is formulated as a zero-dimensional (0D) formulation of the fundamental equations (described in Appendix A). It focuses exclusively on the biological reaction terms (i.e., recruitment, aging, and mortality~~SEAPODYM-LMTL represents the pelagic habitat with three vertical layers whose boundaries follow the euphotic depth Z_{eu}: an epipelagic layer (down to $1.5\,Z_{\mathrm{eu}}$), ~~driven by environmental forcing (i.e., temperature, euphotic depth, and NPP), while explicitly neglecting spatial transport. This simplification allows the model to function independently without worrying about the complex interdependencies with the transport code.~~

~~The model adopts a modular architecture where biological processes are implemented as independent functions, enabling straightforward extension and parallel computation using the Dask library (Dask Development Team, 2016).~~ 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., 2010a, 2015a). 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., 2010a). 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., 2010a) . 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.

~~All data management (including input/output and intermediate states) is performed using the Xarray library (Hoyer and Hamman, 2017) . This library enables the transparent use of standard data formats (e.g., NetCDF and Zarr) and deployment across heterogeneous infrastructures, from local workstations to cloud platforms, without modifying the core codebase.~~

~~A rigorous management of physical units is integrated directly into the parameterization layer using the Pint library (~~The biology of the mesozooplankton group rests on three coupled processes: energy intake from primary production, aging through an age-structured zooplankton production (p), ~~enforcing dimensional consistency and preventing semantic errors.~~

~~The framework is organized into three primary components: (i) a Configuration layer ensuring data integrity and physical units consistency; (ii) a Kernel library containing elementary biological processes, which leverage either native vectorization using the Numpy library (Harris et al., 2020) or Just-In-Time (JIT)compilation provided by the Numba library (Lam et al., 2015) to achieve execution speeds comparable to C++; and~~ and its recruitment ~~(iii) a Model orchestrator that manages the state and temporal evolution~~ 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 τ_r (Lehodey et al., 2010a; Conchon, 2016). The biomass pool then declines through a temperature-dependent mortality rate λ, 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., 2010a), 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 $T_{\mathrm{ref}} = 0°\mathrm{C}$). 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).

2.2 ~~Parameter optimization~~

~~GAs are evolutionary optimization methods that mimic natural selection to solve global optimization problems. They evolve a population of candidate solutions through successive generations using biologically inspired operators: selection, crossover, and mutation (Goldberg, 1989). Unlike deterministic methods, GAs are stochastic and particularly effective for exploring multimodal parameter spaces where traditional techniques might get trapped in local optima (Goldberg, 1989; Falls et al., 2022) . SeapoPym allows users to easily change the operators proposed by the DEAP library (Fortin et al., 2012)~~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., 2015a) and in a synthetic observing-system experiment (Delpech et al., 2020), while the zooplankton recruitment and mortality parameters were manually tuned (Conchon, 2016).

~~Within this framework, each solution corresponds to a specific set of model parameters to be calibrated, fully described in Appendix A and categorized here into three functional groups: (i)~~

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, ageing 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 ~~energy transfer coefficient E, which determines the transfer efficiency from NPP; (ii) the recruitment parameters τ_{r_0} and γ_{τ_r}, controlling the age of transition from production cohorts to recruited biomass; and (iii) the mortality parameters λ_0~~ 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 ~~γ_λ, which define the~~ 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 rate. Table **??** summarizes the model parameters, their reference values, and~~

~~optimization ranges.~~ 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 ~~first phase of the optimization initializes the population using a Sobol sequence to ensure comprehensive coverage of the search space.~~

~~Model parameters with reference values and optimization ranges. It then proceeds through successive generations, evaluating individuals using the objective function and tournament selection (tournament size = 3) to select promising parents.~~

~~We use standard NSGA-II meta-parameters (Deb et al., 2002) with two-point crossover ($CXPB = 0.9$)to promote exploration of the parameter space (see Discussion).~~

~~The optimization framework is organized into modules for modelinstantiation, parameter specification~~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 ~~evaluation, observation handling, and evolutionary algorithm orchestration.Additional modules support constraint handling (e.g., parameter bounds)and logging functionality for tracking optimization progress~~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 ~~reference zooplankton biomass data~~observations

~~The SeapoPym simulations are~~

SeapoPym is forced with the same temperature and NPP fields ~~used to produce the~~ as the operational SEAPODYM-LMTL product, the Copernicus Marine Service ~~product~~ "Global Ocean Low and Mid Trophic Levels Biomass Content Hindcast" (GLOBAL_~~MULTIYEAR_BGC_~~MULTIYEAR_BGC_001_033, ~~DOI: 10.48670/moi-00020) . These fields are provided alongside the LMTL variables (zooplankton and micronekton biomass) at a horizontal resolution of 1/12° (Figure 1)over the global domain. The model outputs include one zooplankton and six micronekton functional groups. The mean temperature in the epipelagic layer is computed from the global ocean model configuration~~ https://doi.org/10.48670/moi-00020) (Titaud et al., 2024) . The temperature is the layer average over the epipelagic layer (named T_{epy} in Appendix A) from the GLORYS12 ~~(Lellouche et al., 2021)~~

, and the vertically integrated NPP is derived from satellite ocean-colour observations using 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) .

To reduce computational costs associated with the global domain, a spatial subsampling is applied. The domain is discretized on a regular grid of 360° longitude × 170° latitude, spanning from 80°S to 90°N and 180°W to 179°E, with a spatial resolutionof 1 °. For each grid cell, the value from the nearest corresponding cell in the original dataset is selected. The temporal coverage spans 22 years, from January 1, 1998, to December 31, 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. The model is initialized with zero biomass and production; the These fields are provided at daily resolution, matching the model's one-day time step. The first two years serve as a are discarded as spin-upperiod for the system to reach equilibrium, and are excluded from subsequent analysis.Data were extracted from the reference simulation at six locations corresponding to regularly sampled sites (hereafter referred to as 'stations '), which were subsequently used to conduct the twin experiment and Sobol analysis. They cover a range of environmental conditions in temperature and NPP, spanning a gradient from cold to warm temperatures and from low to high productivity (Figure , 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). The station coordinates are provided in Table **??**.

Geographic coordinates of the six oceanographic stations used for data extraction.

2.4 Simulation experiments

A series of simulation experiments is conducted to validate the modeling approach. The first experiment verifies the correctness of In order of increasing temperature, these are the Barents Sea (BARENTS, 75°N, 40°E), a subarctic northeast Pacific location in the numerical implementation and assesses model behavior in a simplified, controlled scenario. The second experiment isolates the role of transport processes to determine whether model dynamics are realistic only when physical connectivity is included or , for zooplankton , remain valid in regions with weak transport. The third experiment systematically assesses structural uncertainties through a global sensitivity analysis, identifying parameters and processes that most strongly influence model outputs and those that are poorly constrained. This analysis enhances understanding of model behavior and informs the design of the optimization procedure.The fourth experiment evaluates the optimization strategy within a controlled and transparent framework using twin experiments. 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).

2.3.1 ~~Validation of the numerical implementation~~

~~To validate the numerical implementation against the theoretical analytical solution ($B_{eq} = \frac{R}{\lambda}$, see Appendix A and Lehodey (2001)) , we conducted a series of simulations under constant environmental forcing.~~ 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 μm). 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 ($1.5\,Z_{eu}$, 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.

~~Specifically, four simulations were performed at~~

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 $B_{eq} = R/\lambda$ (Sect. 2.1, Appendix A) under constant forcing. This experiment is conducted at four different temperatures (0, 10, 20 ~~,~~ and 30 ~~°C~~to span the broad physiological thermal range of marine zooplankton. ~~A mean value of 300 $mg \cdot m^{-2} \cdot d^{-1}$ was prescribed for NPP . In this theoretical setup,~~ °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 ~~magnitude of NPP does not affect the test outcome, as it serves solely as a scaling factor for the theoretical asymptote~~time to come within 1% of the analytical steady state.

2.4.1 ~~Impact of transport~~

~~A zooplankton simulation is performed with SeapoPym using the same forcing fields and the reference parameter values listed in Table ??, as employed in the reference LMTL simulation included in the CMEMS product (Ref.QUID, 2024). The comparison of mean zooplankton biomass distributions of LMTL and SeapoPym provides a quantitative assessment of the impact of transport removal on model predictions (see metrics below).~~ 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.

2.4.1 ~~Sobol sensitivity analysis~~

~~Following the validation and evaluation of SeapoPym 's implementation above,~~

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

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

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

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 ~~model's inherent structure is examined using a Sobol analysis (Sobol', 2001) to determine its identifiability limits (i.e., the ability to uniquely estimate parameter values from observations) before any optimization is performed. Sobol's global sensitivity analysis characterizes the influence of model parameters on the environment using three metrics we selected: the date of the seasonal biomass peak (argmax), the average annual biomass (mean), and the biomass variance (variance). The Python library used for this analysis is SALib (Iwanaga et al., 2022; Herman and Usher, 2017). The range of environmental conditions in temperature and NPP is represented by a set of six oceanographic stations spanning a gradient from cold to warm temperatures and from low to high productivity.~~ in-situ observations. This comparison focuses solely on the order of magnitude.

~~In~~

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, ~~the~~ it provides a first-order ~~(S_1) and total-order (~~index S_1, representing the fraction of output variance attributable to that parameter alone, as well as a total index S_T) ~~indices both measure how important the model parameters are — but they capture different kinds of effects. S_1 measures the direct, independent effect of a single parameteron the output, ignoring interactions with other parameters, i.e., how much of the output variance is explained by a given parameter alone. If $S_1 = 0.3$, then 30 % of the variance in the output is caused by that parameter by itself. S_T measures the overall contribution~~ , 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~~, including both its S_1 (direct) effect and all interaction effects with any other parameters.~~

~~The larger the number of simulations, the lower the uncertainty in the Sobol sensitivity analysis.Saltelli (2002) and Prieur and Tarantola (~~ ~~recommended that, for a model with P parameters, a Sobol analysis requires a minimum sample size of $(P + 2) \times N$, where N typically ranges from 10^3 to 10^6 for the S_1 indices.In our case, this corresponds to a total of 7×10^3 to 7×10^6 simulations. To keep computation costs reasonable, we set the number of simulations to 10^6.~~

~~The range of values explored for the model parameters covers up to five times the reference values for zooplankton. The SALib library is used to generate the ensemble of parameter sets. For each simulation, the three metrics are computed independently for each station and stored. SALib then uses all of the resulting metrics to compute the sensitivity indices , which are subsequently summarized in a synthetic plot for the analysis.~~'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).

2.5.1 ~~Parameter optimization using the GA and twin experiments~~

~~The GA optimization was applied to the five model parameters using a twin experiment. Starting from a quasi-randomly initialized parameter set, using the Sobol sampling function provided by SALib,~~ We run SeapoPym at the six station locations (Sect. 2.3), with the ~~twin experiment assesses the optimization algorithm's ability to recover the reference parameters by fitting the pseudo-observations. These pseudo-observations are~~ 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 ~~zooplankton biomass time series (2000-2019) created from the reference parameter set with SeapoPym at~~ 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 ~~six oceanographic stations used in the Sobol analysis (cf.section 2.4.2). The parameter search interval was restricted to a reasonable range for zooplankton (0-200 % of the reference value for each parameter). The stopping criterion for the optimization was set to 20 generations: an initial population of 14,336 individuals generated using Sobol sampling, followed by 19 generations of 10, 000 individuals each.~~ 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.

~~In total, 204, 336 simulations are run by experiment. Individual simulation times ranged from 250 ms to~~ 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 model itself and used without added noise, 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~~s depending on the computing environment, corresponding to approximately 14–56 hours on a single core or 1–2 hours with 30 cores in parallel. All experiments were conducted using Python 3.12.~~) 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.

2.7 ~~Metrics~~

~~Two classical and complementary metrics are used to quantify differences and structural divergences in space and time. The Mean Absolute Percentage Error (MAPE)is a measure of the average relative error of the model compared to the reference. The Root Mean Square Error (RMSE) quantifies the average magnitude of differences between the two simulated fields, offering a global estimate of model accuracy while allowing spatial and temporal mapping of discrepancies.~~

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

$$\text{MAPE} = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|$$

~~For the optimization, the~~ 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 ~~based on the Normalized Root Mean Square Error~~ the root mean square error normalized by the mean of the synthetic observations (NRMSE)~~between model predictions and pseudo-observations:~~

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

~~The use of NRMSE emphasizes the model's ability to reproduce system dynamics rather than absolute mean levels and , importantly, allows comparison across multiple~~ . Following Eq. (2), $\hat{y}$ refers to the simulated biomass and y refers to the synthetic observations, so that stations with different ~~variances.~~

~~The GA's ability to recover the initial parameters used to generate the observations is evaluated using the Mean Absolute Percentage Error (MAPE)~~ 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.

~~The cost function can be readily adapted to use an alternative metric (see Discussion).~~

3 Results

3.1 ~~Theoretical~~ Implementation validation ~~of~~ and the ~~numerical implementation~~impact of transport

~~Figure 3 shows the temporal evolution of the modeled biomass with SeapoPym and the theoretical convergence value. As expected from the temperature-dependent definition of mortality (Eq.A2), the time required to reach steady state from the initial condition increases as the temperature decreases. This time is only 10 days at 30°C and 30 days at 20°C, but extends to~~

~~4 and 16 months at 10°C and 0°C, respectively. Once steady state is achieved, the difference from the theoretical value is less than 0.01 % (Figure 3) and is attributed to computational accuracy~~

Under constant forcing, the simulated biomass converges to the analytical steady state $B_{eq} = R/\lambda$ 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°C and about two years at 0°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.

3.2 ~~Impact of transport~~

~~To estimate the error induced by neglecting transport , we compare the reference~~ The difference between SeapoPym and the operational SEAPODYM-LMTL ~~simulation (Fig. 1d) against the static SeapoPym output (Fig.4a), both run with identical forcing. The resulting spatial error (RMSE) reveals that the impact of transport~~ product measures the combined effect of omitting transport and the re-implementation. As Fig. 4 shows, it is non-uniform ~~and driven by the interplay of current intensity, biomass gradients, and temperature (Figure 4b).~~

~~The highest discrepancies are found in dynamically active regions characterized by strongfronts, such as~~ 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. ~~In these areas, the combination of intense transport and steep biomass gradients maximizes the divergence between the static and advective models.~~

~~Temperature also plays a critical role by modulating biological time scales (Figure 3). In colder, productive regions (e.g., the Labrador and Oyashio currents), slower growth and mortality rates extend the time required for cohort development (e.g., Daase et al. (2013)) . This increases the organisms' exposure to currents, thereby amplifying transport-induced errors.~~

~~Conversely, the RMSE remains minimal in subtropical gyres and tropical regions. Here, low biomass combined with rapid biological turnover (Figure 3) limits the accumulation of discrepancies, rendering the static assumption locally appropriate.~~

3.2 ~~Global sensitivity analysis (Sobol)~~

~~Sobol's analysis reveals that model variability is primarily driven by parameter interactions rather than independent effects. Across all stations, the S_T indices are substantially larger than the S_1 indices, often by a factor of two to three. This large discrepancy reveals that the direct effect of any single parameter is minor compared to~~ 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 under-estimate 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 ~~variance driven by parameter interactions.However, when considering S_T, all parameters exhibit significant influence on at least one output metric. This confirms that no parameter can be neglected and that the model behaves as a highly nonlinear system in which recruitment, energy intake, and mortality processes are tightly coupled.~~ observations are not used for calibration.

~~While all parameters are influential, they affect specific components of the zooplankton dynamics, revealing a clear functional division. First, the mortality parameters (λ_0,γ_λ) emerge as the primary direct drivers~~

3.2 Sobol sensitivity analysis

Sensitivity analysis of biomass magnitude ~~. Their S_1 indices dominate both mean biomass and variance metrics. This is physically consistent with the model structure: whereas NPP strictly bounds energy input (~~reveals a distribution of roles (Fig. 6). This magnitude is determined by energy transfer E ~~), a mortality rate approaching zero can, in principle, lead to unbounded biomass growth. In sharp contrast, recruitment parameters(τ_{r_0}, γ_{τ_r}) exert negligible influence on biomass quantity ($S_T \approx 0$ for mean and variance). Their impact is almost exclusively restricted to the timing of the seasonal peak, where they exhibit high S_T. Finally, the energy parameter (E) acts as an indirect modulator. It shows no direct effect on peak timing ($S_1 \approx 0$) and influences biomass magnitudeentirely through interactions with mortality parameters, effectively linking the energy source to the mortality sink~~and the two mortality parameters, with the balance between these factors shifting along the temperature gradient. First-order indices for E and λ_0 decrease from the coldest station to the warmest station, whereas the index for γ_λ rises from a value near zero at BARENTS to approximately 0.6 at HOT. The thermal term γ_λ 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 ~~sensitivity of these parameters varies across the environmental gradient. Crucially, the identifiability of recruitment parameters (τ_{r_0},γ_{τ_r}) depends on the clarity of the seasonal signal. In cold environments, slow development times lead to extended life cycles of 1–2 years. This temporal scale is captured by the model dynamics shown in Figure 3, where the time required to reach equilibrium at low temperatures (0°C) extends to similar durations. This is consistent with the phenology of high-latitude copepods (Tande and Båmstedt, 1985; Conover, 1988; Kosobokova, 1999). This phenology creates sharp seasonal peaks, enhancing the sensitivity of the "Peak timing " metric to recruitment parameters in these regions compared to continuous tropical production in warm waters~~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 τ_{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 γ_{τ_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 τ_{r_0} and γ_{τ_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.

~~In summary, accurately constraining the model requires a cost function that simultaneously captures errors in magnitude , variability, and seasonal timing~~ 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). ~~The Normalized Root Mean Square Error (NRMSE)is particularly well-suited for this purpose, as it inherently aggregates these three components (Taylor, 2001) . Consequently, minimizing the NRMSE across contrasting environmental regimes is the most robust strategy for parameter optimization.~~

3.3 ~~Twin experiment optimization using GA~~

~~Two types of experiments are conducted to evaluate the optimization approach's capabilities. The first experiment analyzes each station individually, while the second considers all stations jointly. This illustrates the approach's ability to estimate model parameters from observations collected at one or multiple stations.It also highlights its potential for conducting OSSEs to support the design of observation networks~~Therefore, its logarithm is expressed as a sum of distinct terms: $\log B = \log E - \log \lambda_0 - \gamma_\lambda \bar{T} +$ 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.

3.2.1 ~~Cost function and fit to biomass~~

~~When each station is analyzed independently, the best performance is obtained in the coldest waters (BARENTS station), where the NRMSE decreases steadily over successive generations and reaches a minimum of 0.015 after 20 generations (Figure 6a). The other stations do not exhibit the same level of continuous improvement and may reach their optimal score after only a few generations. Conversely, when all stations are used jointly, all cost functions improve over the generations and reach their best values between generations 15-20, with substantially better scores than in the single-station experiments, all below 0.1 (Figure 6b). The only exception is BARENTS, for which the solution remains good but is slightly worse than the single-station result (0.035). In terms of fit to observed biomass, an $NMRSE < 0.1$ means that the coefficient of determination R^2 $(1 - NRMSE^2) > 0.99$ (assuming no bias in errors: Müller-Plath and Lüdecke (2024)).~~

3.3 Twin experiments: convergence and parameter recovery

3.3.1 ~~Parameter estimates~~

The most challenging parameters to estimate are τ_{r_0} and 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 τ_{r_0} and γ_{τ_r} , which exhibit similar patterns of evolution throughout the optimisation process (Figure 7). When considering each observation station independently, only the BARENTS station achieves excellent performance for these two parameters, with MAPE values below 1 % and 10 %, respectively. However, combining all stations in the optimisation enables recovery of their initial parameter values with errors below 10 miss the reference by 40 to 160 %.

In the warm, low-productivity environment of the HOT station, λ_0 cannot be accurately estimated in separate optimisation experiments . In contrast, this parameter is estimated with an error below 10 % at the other stations 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, with the best performance at BARENTS (<1 %). When all stations are combined, λ_0 is also recovered with an error of about 10 %. The parameter E is estimated with errors generally below or slightly above10 %. The parameter γ_λ yields the best performance, with errors below 10 % when stations are considered separately and below 5 % when all stations are combinedso none of its twenty restarts converged and we leave HOT out of the recovery analysis.

The highest density interval (HDI) is a widely used measure of uncertainty in parameter estimation (e.g., Kruschke (2015)). It represents the narrowest range of parameter values containing a specified proportion of the posterior probability mass—commonly 95 %, as in this study—derived here from the 1, 000 best solutions (Figure **??**). Notably, for optimisations at individual stations, the HDI of parameter τ_{r_0} decreases only for the BARENTS stationThree 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 same time, for γ_{τ_r} the HDI remains maximal both at BARENTS and at the other stations. Even when data from all stations are combined, the HDI for γ_{τ_r} remains unchanged. For the other parameters, the reduction of the HDI over successive generations is clearly evident, with particularly narrow intervals achieved for λ_0 and γ_λ. 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.

This overall performance can be explained by the complementarity of the sensitivity profiles: cold waters constrain τ_{r_0}, λ_0, and E, whereas warm waters constrain γ_λ. The overall consistency of the optimized parameters validates the multi-station approach. The non-identifiability of γ_{τ_r} is not due to a failure of the GA—which converges well for the other parameters and obtains a solution very close to the initial result in terms of NRMSE—but rather to its weak structural influence on the model outputs.

4 Discussion

~~This study introduces SeapoPym, a simplified version of the~~

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~~that decouples biological mechanisms from transport, using the Python scientific ecosystem, and incorporates a GA for flexible, stochastic parameter estimation. The results validate SeapoPym's numerical accuracy, explore its behavior through Sobol sensitivity analysis, and evaluate its optimization framework using a twin experiment~~. 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 ~~been recognized for a long time~~ 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; **?**; Mackas and Coyle, 2005)~~(e.g., Banse, 1964; Edvardsen et al., 2006; Mackas and Coyle, 20 . Our comparison with the operational SEAPODYM-LMTL ~~confirms that SeapoPym remains generally suitable for simulating and estimating mesozooplankton parameters across most warm- and temperate-water regions, except in areas with strong dynamic activity. This finding~~ product confirms that transport cannot be ignored in such areas, especially when these conditions prevail in cold-water environments. This result aligns with theoretical ~~expectations~~ predictions stating that physical transport becomes ~~non-negligible~~ significant when biological timescales exceed physical retention timescales. This ~~condition~~ scenario is more likely in cold waters, where ~~the generation times of zooplankton~~ zooplankton generation times can exceed one year ~~(Tande and Båmstedt, 1985; Conover, 1988; Kosobokova, 1999; Daase et al., 2013).~~ (Conover, 1988; Kosobokova, 1999; Daase et al., 201 . 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.

~~However, since cold to temperate environments are characterized by strong seasonal cycles, they provide a clear and detectable signal for parameter estimation, as demonstrated by our results. Slow rates of development in cold-water environments exert a~~ 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 λ, 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 λ, 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.

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 ~~control on biomassstructure, which facilitates the estimation of certain parameter values, as shown by the twin experiments.~~ index of γ_{τ_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.

~~While these environments facilitate estimation, the S_T indices highlight indirect parameter interactions, which complicate optimization when relying on a single environment. The use of diverse observational data—especially from cold, high-contrast environments—can help constrain parameter values more effectively, as confirmed by the twin experiments. Given that these experiments were realized in an ideal framework without noise and using long, continuous time series, specific parameters will likely need to be fixed to their best-known values during model calibration in less favorable, real-world applications, especially τ_{r_0} and γ_{τ_r}, which govern recruitment timing. This strategy should help reduce the number of equally plausible solutions that fit the same observational dataset (equifinality).~~ 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.

~~The GA employed in this study is a basic, highly exploratory approach. While it entails a higher computational cost than gradient-based methods, it offers exceptional flexibility and facilitates parallel computing, enabling efficient distribution of the computational load (Cantú-Paz, 2001). As an evolutionary algorithm operating on discrete genotypes, the GA inherently~~

handles both continuous and discrete variables, and applies directly without requiring any modification to the model code. For example, although the current implementation assigns a fixed Diel Vertical Migration (DVM) behavior to each group, future iterations could optimize this trait by defining it as a discrete variable. Similarly, for species-specific applications, a discrete parameterization approach would allow the algorithm to dynamically select optimal prey fields from distinct Plankton Functional Types (PFTs). 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).

Several algorithmic enhancements could improve optimization performance: elitism to preserve the best solutions across generations (Angelova et al., 2011), Simulated Binary Crossover (SBX) to potentially accelerate convergence (Deb et al., 2002), or CMA-ES, which self-adapts to parameter correlations and search step size (Hansen and Ostermeier, 2001; Hansen, 2016).

Future model improvements will rely heavily on the availability of suitable datasets. While 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 standardized time series remain scarce, our results indicate that maintaining sampling efforts in environmentally contrasting regimes is particularly effective for parameter optimization. Beyond these fixed stations, extensive databases of discrete sampling (e.g., COPEPOD) and acoustic profiles offer immediate opportunities to refine biomass and DVM parameterizations at each grid cell (Stanton, 2000; Dmitrenko et al., 2020). Concurrently, emerging high-resolution technologies, such as automated imaging onboard autonomous profilers (UPV) and environmental DNA (eDNA), will be essential for resolving specific functional groups or species. SeapoPym is specifically designed to integrate this increasing dataheterogeneity—a flexibility that extends to micronekton modeling. 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.

Beyond parameter estimation, this framework offers a robust mechanism for observing system design. Analogous to Observing System Simulation Experiments (OSSEs) used in oceanography, our twin experiment approach allows us to quantify the

potential value of new sampling stations prior to their deployment. By generating synthetic observations and forcings with realistic noise, the modelcan inform cost-effective monitoring strategies that maximize parameter retrieval accuracy.

5 Conclusions

Several limitations should be acknowledged. First, We presented here SeapoPym, an open-source Python re-implementation of the omission of physical transport restricts the model's applicability to regions where biological timescales are shorter than physical retention timescales, as discussed above. Second, the twin experiments were conducted under idealized conditions—without observational noise and using long, continuous time series—meaning that real-world applications may face additional challenges in parameter retrieval. Third, the current framework focuses on a single functional group (mesozooplankton), though the modular architecture allows for straightforward extension to micronekton groups. Finally, the non-identifiability of the recruitment-timing parameter γ_{τ_r} suggests that this parameter may need to be fixed based on prior knowledge when observational constraints are limited. 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.

This study demonstrates that SeapoPym, a Python-based reformulation of SEAPODYM-LMTL, successfully reproduces zooplankton dynamics in most oceanic regions where transport effects are secondary to biological processes. Sobol sensitivity analysis reveals mortality parameters as The identifiability of the parameters depends on the primary drivers of biomass magnitude and variability, while twin experiments confirm that parameter identifiability requires observations spanning contrasting thermal environments—particularly cold, high-latitude regions with pronounced seasonal cycles that facilitate the estimation of recruitment-timing parameters 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.

~~SeapoPym provides an agile 0-D modeling framework that enables non-experts to test biological hypotheses and explore parameter estimation strategies. It offers a low-cost tool for investigating how observational design and dataset composition affect parameter retrieval, without the computational burden of a complete spatial model. Future developments aim to apply the model to real-world datasets, including acoustic observations and data from emerging sensors (Everett et al., 2017). The next logical step is to address the lack of transport by coupling with an ocean circulation model (e.g., NEMO)or by using machine-learning-based emulators. Importantly, this expansion must follow strict software engineering standards to prevent the 'complexity trap' of legacy models, ensuring the code stays modular, adaptable, and accessible.~~ The twin experiments use noise-free observations, complete at every time step, and the exact forcing. The identifiability they show is therefore a best case. 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~~

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. 2010a). 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 dynamics of the total~~

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

$$\frac{\partial B}{\partial t} = R(t) - \lambda(t) B(t)$$

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

~~where:~~

where R [$\mathrm{g\,C\,m^{-2}\,d^{-1}}$] is the recruitment flux [$\mathrm{g\,m^{-2}\,day^{-1}}$]; B is the total biomass [$\mathrm{g\,m^{-2}}$]; from the ageing production field (defined below) and λ [$\mathrm{d^{-1}}$] is the mortality rate[$\mathrm{day^{-1}}$]. . At steady state ($\partial B/\partial t = 0$), this yields this gives the equilibrium biomass $B_{eq} = R/\lambda$.

Mortality is parameterized as a temperature-dependent function following an exponential relationship:

$$\lambda(t) = \lambda_0 \exp\left(\gamma_\lambda (T(t) - T_{\mathrm{ref}})\right)$$

$B_{eq} = R/\lambda$. Mortality rises exponentially with temperature:

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

where:

where λ_0 is the mortality rate at the reference temperature T_{ref} [$\mathrm{day^{-1}}$]; and γ_λ is the thermal sensitivitycoefficient [$\mathrm{^\circ C^{-1}}$]. Within the LMTL framework, the effective temperature is constrained to positive values:

$$T(t) = \max(T_{\mathrm{env}}(t), T_{\mathrm{ref}}) \quad \text{with} \quad T_{\mathrm{ref}} = 0^\circ\mathrm{C}$$

[$\mathrm{^\circ C^{-1}}$] its thermal sensitivity. The temperature entering the metabolic terms is the normalized temperature of the SEAPODYM-LMTL model (Lehodey et al., 2010a), 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_{env} is the temperature forcing (Sect. 2.3) and T_{ref} the reference temperature, with T floored at T_{ref} so that the recruitment age cannot exceed the maximum age the model resolves (below).

A2 Production

The evolution of zooplankton production $p(t, \tau)$ [$\mathrm{g\,C\,m^{-2}\,day^{-1}}$] as a function of time t and

The production field $p(t, \tau)$ [$\mathrm{g\,C\,m^{-2}\,d^{-1}}$] describes how production is distributed across age τ is described by the McKendrick–Von Foerster equation (McKendrick, 1926; Von Foerster, 1959). The source term corresponds to the transfer of NPP to the first age class ($\tau = 0$), while the sink term represents the transfer from production cohorts to the biomass pool via a recruitment rate μ:

$$\frac{\partial p}{\partial t} + \frac{\partial p}{\partial \tau} = -\mu(t, \tau)\, p(t, \tau)$$

. 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):

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

subject to the boundary condition:

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

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):

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

where: NPP is the Net Primary Production of phytoplankton $[\mathrm{g\,m^{-2}\,day^{-1}}]$;

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

A3 Recruitment

The transfer rate μ activates only when the cohort age falls within a

Production is recruited into the biomass pool when it reaches the temperature-dependent recruitment window:

$$\mu(t, \tau) = \begin{cases} \alpha & \text{if } \tau \in [\tau_r(t), \tau_{r_0}] \\ 0 & \text{otherwise} \end{cases} \quad \text{with} \quad \tau_r(t) = \tau_{r_0} \exp\left(-\gamma_{\tau_r}(T(t) - T_{\mathrm{ref}})\right)$$

age $\tau_r(t)$:

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

where: $\tau_r(t)$ is the minimum

where τ_{r_0} is the recruitment age at temperature T [day]; τ_{r_0} is the maximum recruitment age at T_{ref} [day]; the reference temperature (hence the maximum age represented in the model) and γ_{τ_r} is the thermal sensitivity coefficient for recruitment age $[^\circ\mathrm{C}^{-1}]$; α is a constant recruitment rate $[\mathrm{day}^{-1}]$. Assuming $\mu(t, \tau)$ is integrable, the analytical solution along the characteristic lines (for $\tau \leq t$) is :

$$p(t, \tau) = E \cdot \mathrm{NPP}(t - \tau) \cdot \exp\left(-\int_0^{\tau} \mu(t - \tau + s, s)\,\mathrm{d}s\right)$$

The exponential term represents the proportion of production retained in the cohort relative to the initial generation at $\tau = 0$. In the limiting case where $\alpha \to \infty$, total absorption occurs: all production reaching the recruitment age τ_r is instantaneously transferred to the biomass. This is equivalent to an absorbing boundary condition $p(t, \tau_r) = 0$ $[^\circ\mathrm{C}^{-1}]$ 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:

~~Conceptual Note: Unlike standard formulations where μ represents mortality, here it signifies a state transition flux. Conceptually, p and B represent distinct life stages (e.g., developing cohorts vs. recruited population). However, the model structure treats them differently:~~

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

Recruitment moves production from the field p ~~is age-structured with implicit mortality (captured in the coefficient E), while~~ into the pool B~~is an age-aggregated pool subject to explicit thermal mortality.~~, a transfer and not a loss. Mortality acts only on B, through λ, and any loss during development is absorbed into the efficiency E.

A4 ~~Numerical Resolution~~

~~We adopt the notation X_a^n to denote the value of variable X at age $\tau = a\Delta\tau$ and time $t = n\Delta t$. To minimize numerical diffusion, we enforce the Courant-Friedrichs-Lewy (CFL) condition $\Delta t = \Delta\tau = \Delta$.~~

A4 Numerical solution

Age and time are discretized into steps $\Delta\tau$ and Δt, 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:

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

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 ~~equation is solved using an explicit Eulerscheme:~~

$$B^{n+1} = B^n + \Delta \cdot (R^n - \lambda B^n)$$

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

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

~~The analytical solution for p (Eq. A5)requires knowledge of the full thermal history of each cohort, which is computationally expensive in an Eulerian framework. Therefore, p is computed iteratively from the previous time stepusing a survival probability S:~~

$$p_a^{n+1} = p_{a-1}^n \cdot S_a^{n+1} \quad \text{with} \quad S_a^{n+1} = e^{-\mu_a^{n+1}\Delta}$$

~~The specific recruitment flux from age class a at time $n+1$, denoted R_a^{n+1}, corresponds to the production lost by the cohort during the time step:~~

$$R_a^{n+1} = p_{a-1}^n - p_a^{n+1} = p_{a-1}^n(1 - S_a^{n+1})$$

~~Finally, the total recruitment flux R^{n+1} feeding the biomass is the sum of contributions from all age classes within the recruitment window:~~

$$R^{n+1} = \sum_{a=0}^{\tau_{r_0}} R_a^{n+1}$$

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 ~~v0.1~~ model code is open-source~~and~~, distributed under the GPLv3 license, and available at https://github.com/SeapoPym/seapopym. The ~~current~~ exact version ~~of the model~~ described in this paper is SeapoPym v0.1.1, archived on Zenodo (https:

//doi.org/10.5281/zenodo.21838667, Lehodey 2026)~~and is available at~~ . The repository includes the source code, installation instructions, and a user manual. A dedicated reproducibility deposit accompanying this paper~~—~~, providing the ~~notebooks,~~ 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 ~~v1~~v2.0.0) and archived on Zenodo at https://doi.org/10.5281/zenodo.21849337. 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. ~~Physical~~ The temperature and ~~biogeochemical forcing~~ NPP fields ~~used in this study (GLORYS12~~forcing the model, ~~Net Primary Production) are available~~ together with the mesozooplankton biomass from the operational SEAPODYM-LMTL product and the epipelagic current field, are distributed by the Copernicus Marine Service (CMEMS) ~~at~~ 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 ~~reference~~ 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 ~~data~~ (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 (~~SEAPODYM-LMTL~~https://hahana.soest.hawaii.edu/hot/hot-dogs/) ~~used for validation~~ and ~~twin experiments are also distributed by CMEMS~~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 ~~original draft~~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

Angelova, M., Tzonkov, S., and Pencheva, T.: Genetic Algorithms Based Parameter Identification of Yeast Fed-Batch Cultivation, p. 224–231, Springer Berlin Heidelberg, ISBN 9783642184666, https://doi.org/10.1007/978-3-642-18466-6_26, 2011.

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.

Cantú-Paz, E.: Migration Policies, Selection Pressure, and Parallel Evolutionary Algorithms, Journal of Heuristics, 7, 311–334, 2001.

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.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE T. Evol. Comput., 6, 182–197, https://doi.org/10.1109/4235.996017, 2002.

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.

Dmitrenko, I. A., Petrusevich, V., Darnis, G., Kirillov, S. A., Komarov, A. S., Ehn, J. K., Forest, A., Fortier, L., Rysgaard, S., and Barber, D. G.: Sea-ice and water dynamics and moonlight impact the acoustic backscatter diurnal signal over the eastern Beaufort Sea continental slope, Ocean Sci., 16, 1261–1283, https://doi.org/10.5194/os-16-1261-2020, 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.

Everett, J. D., Baird, M. E., Buchanan, P., Bulman, C., Davies, C., Downie, R., Griffiths, C., Heneghan, R., Kloser, R. J., Laiolo, L., Lara-Lopez, A., Lozano-Montes, H., Matear, R. J., McEnnulty, F., Robson, B., Rochester, W., Skerratt, J., Smith, J. A., Strzelecki, J., Suthers,

I. M., Swadling, K. M., van Ruth, P., and Richardson, A. J.: Modeling What We Sample and Sampling What We Model: Challenges for Zooplankton Model Assessment, Front. Mar. Sci., 4, https://doi.org/10.3389/fmars.2017.00077, 2017.

Falls, M., Bernardello, R., Castrillo, M., Acosta, M., Llort, J., and Galí, M.: Use of genetic algorithms for ocean model parameter optimisation: a case study using PISCES-v2_RC for North Atlantic particulate organic carbon, Geosci. Model Dev., 15, 5713–5737, https://doi.org/10.5194/gmd-15-5713-2022, 2022.

Fortin, F.-A., De Rainville, F.-M., Gardner, M.-A., Parizeau, M., and Gagné, C.: DEAP: Evolutionary Algorithms Made Easy, Journal of Machine Learning Research, 13, 2171–2175, 2012.

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.

Goldberg, D. E.: Genetic Algorithms in Search, Optimization, and Machine Learning, Addison-Wesley, 1989.

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.

Holland, M. M., Artigas, L. F., Atkinson, A., Best, M., Bresnan, E., Devlin, M., Eerkes-Medrano, D., Johansen, M., Johns, D. G., Machairopoulou, M., Pitois, S., Scott, J., Schilder, J., Stern, R., Tait, K., Whyte, C., Widdicombe, C., and McQuatters-Gollop, A.: Mind the gap - The need to integrate novel plankton methods alongside ongoing long-term monitoring, Ocean & Coastal Management, 262, 107 542, https://doi.org/10.1016/j.ocecoaman.2025.107542, 2025.

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.

Kruschke, J. K.: JAGS, p. 193–219, Elsevier, ISBN 9780124058880, https://doi.org/10.1016/b978-0-12-405888-0.00008-8, 2015.

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, 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, 2010a.

Lehodey, P., Senina, I., Sibert, J., Bopp, L., Calmettes, B., Hampton, J., and Murtugudde, R.: Preliminary forecasts of Pacific bigeye tuna population trends under the A2 IPCC scenario, Prog. Oceanogr., 86, 302–315, https://doi.org/10.1016/j.pocean.2010.04.021, 2010b.

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, 2015a.

Lehodey, P., Senina, I., Nicol, S., and Hampton, J.: Modelling the impact of climate change on South Pacific albacore tuna, Deep-Sea Res. Pt. II, 113, 246–259, https://doi.org/10.1016/j.dsr2.2014.10.028, 2015b.

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.

Müller-Plath, G. and Lüdecke, H. J.: Normalized coefficients of prediction accuracy for comparative forecast verification and modeling, Res. Stat., 2, https://doi.org/10.1080/27684520.2024.2317172, 2024.

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.

Prieur, C. and Tarantola, S.: Variance-based sensitivity analysis: Theory and estimation algorithms, Handbook of Uncertainty Quantification, pp. 1217–1239, 2019.

Ratnarajah, L., Abu-Alhaija, R., Atkinson, A., Batten, S., Bax, N. J., Bernard, K. S., Canonico, G., Cornils, A., Everett, J. D., Grigoratou, M., Ishak, N. H. A., Johns, D., Lombard, F., Muxagata, E., Ostle, C., Pitois, S., Richardson, A. J., Schmidt, K., Stemmann, L., Swadling, K. M., Yang, G., and Yebra, L.: Monitoring and modelling marine zooplankton in a changing climate, Nat. Commun., 14, https://doi.org/10.1038/s41467-023-36241-5, 2023.

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.

Stanton, T.: Review and recommendations for the modelling of acoustic scattering by fluid-like elongated zooplankton: euphausiids and copepods, ICES J. Mar. Sci., 57, 793–807, https://doi.org/10.1006/jmsc.1999.0517, 2000.

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.

Tande, K. S. and Båmstedt, U.: Grazing rates of the copepods Calanus glacialis and C. finmarchicus in arctic waters of the Barents Sea, Mar. Biol., 87, 251–258, https://doi.org/10.1007/bf00397802, 1985.

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.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res. Atmos., 106, 7183–7192, https://doi.org/10.1029/2000jd900719, 2001.

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 τ_r and the mortality rate λ both vary with temperature as $\exp(\gamma \bar{T})$ (Appendix A), so λ rises and τ_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
τ_{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., 2010a)	A6
γ_{τ_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
λ_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
γ_λ	Temperature coefficient of the mortality rate	$0.15\,^{\circ}\mathrm{C}^{-1}$	0.001 to $0.25\,^{\circ}\mathrm{C}^{-1}$	As above	A2

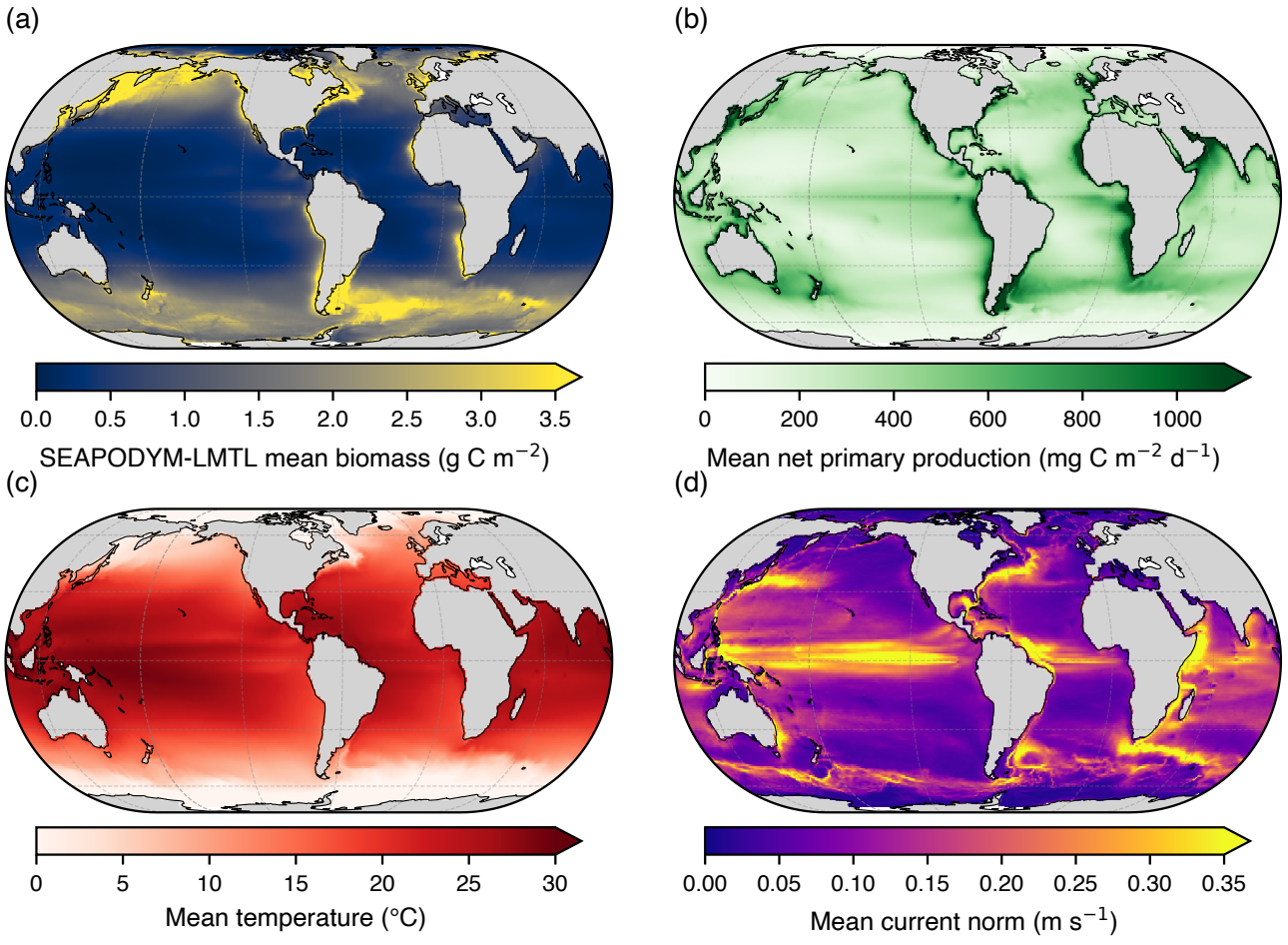

Figure 1. ~~Average map distributions (2000-2020)~~ Global ocean maps of the time-mean forcing and the reference biomass, averaged over 2000 to 2019. (a) ~~temperature in~~ Mean mesozooplankton biomass of the ~~epipelagic layer and~~ operational SEAPODYM-LMTL product, the two-dimensional reference that SeapoPym is compared against in Fig. 4. (b) Mean vertically integrated NPP~~used as forcing for the SeapoPym simulations~~. (c) Mean ~~current norm and~~ temperature of the epipelagic layer. (d) ~~zooplankton biomass modeled by LMTL are shown for further comparison and analysis~~Mean current norm of the epipelagic layer.

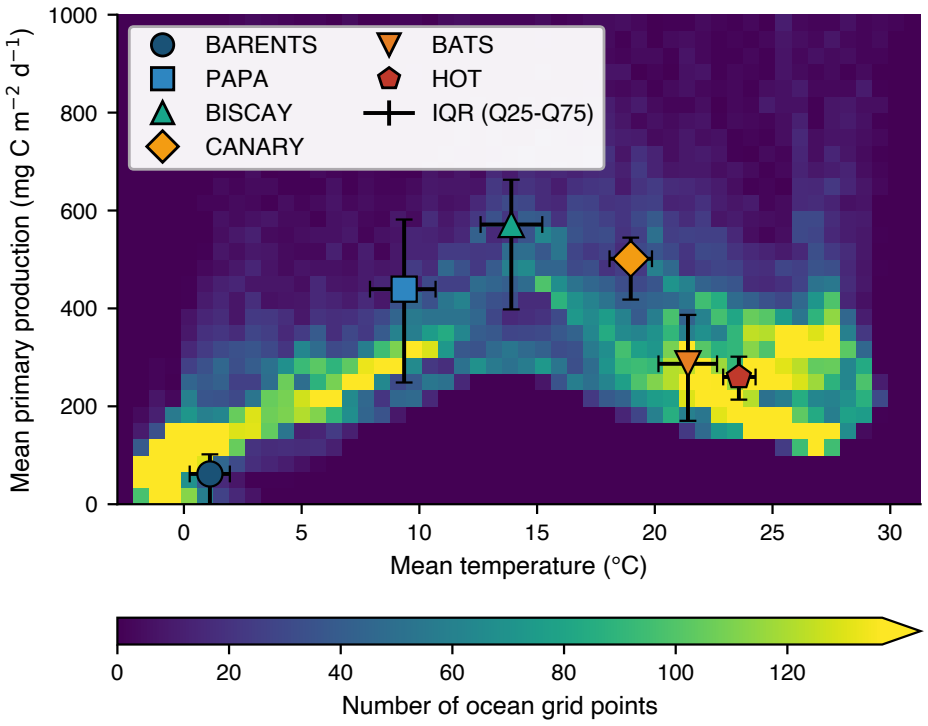

Figure 2. ~~Global distribution~~ The six stations placed in the global temperature and primary-production space. The color field is a two-dimensional histogram of the ocean grid cells, binned by their 2000 to 2019 mean temperature ~~versus~~ and mean ~~net primary production from 1998 to 2020. Light blue dots (39~~NPP~~, 853) represent~~ the ~~global forcing dataset~~color giving the number of cells in each bin. ~~The six selected oceanographic stations are highlighted with distinct symbols that indicate their~~ Each marker is the mean ~~positions within this environmental space. Error bars indicate~~ position of one station, and the bars span its interquartile range (25th ~~and~~ to 75th ~~quantiles~~ percentile) in temperature (horizontal) and in production (vertical) over the ~~time series~~same period. Station color encodes mean temperature from cold blue to warm red, a convention kept across every figure.

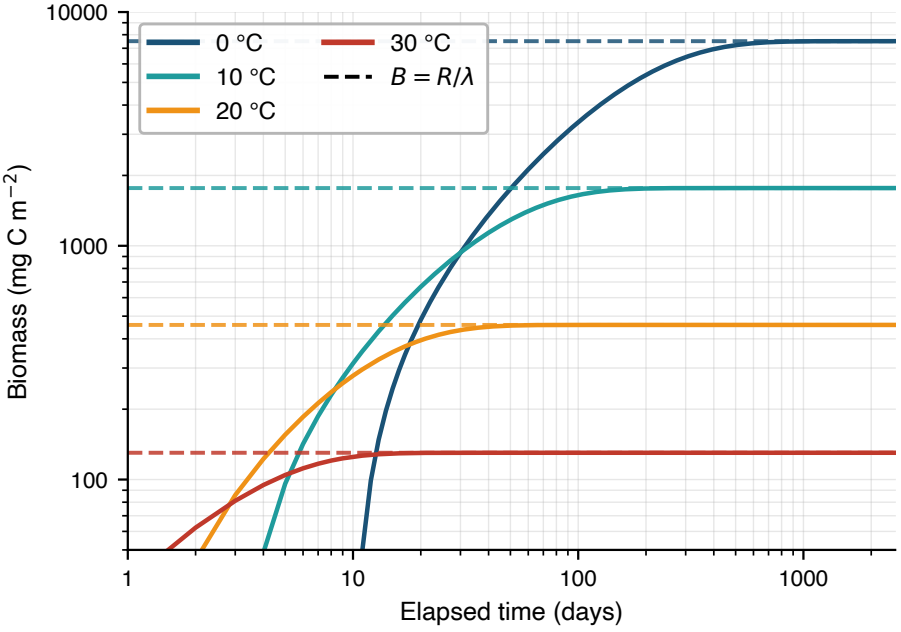

Figure 3. ~~Results from~~ Convergence of the ~~theoretical benchmark with~~ SeapoPym biomass to its analytical steady state under constant ~~mortality and recruitment rates~~ forcing. Each solid line is the simulated biomass at one of four ~~target~~ constant temperatures ~~for~~ (0, 10, 20 and 30 °C), with NPP held at 300 mg C m^{-2} d^{-1}, and the ~~SeapoPym model. The theoretical asymptote~~ matching dashed line is ~~shown~~ the analytical equilibrium $B_{eq} = R/\lambda$ (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 ~~a dashed line~~temperature falls, reaching within 1 percent in about fifteen days at 30 °C and about two years at 0 °C.

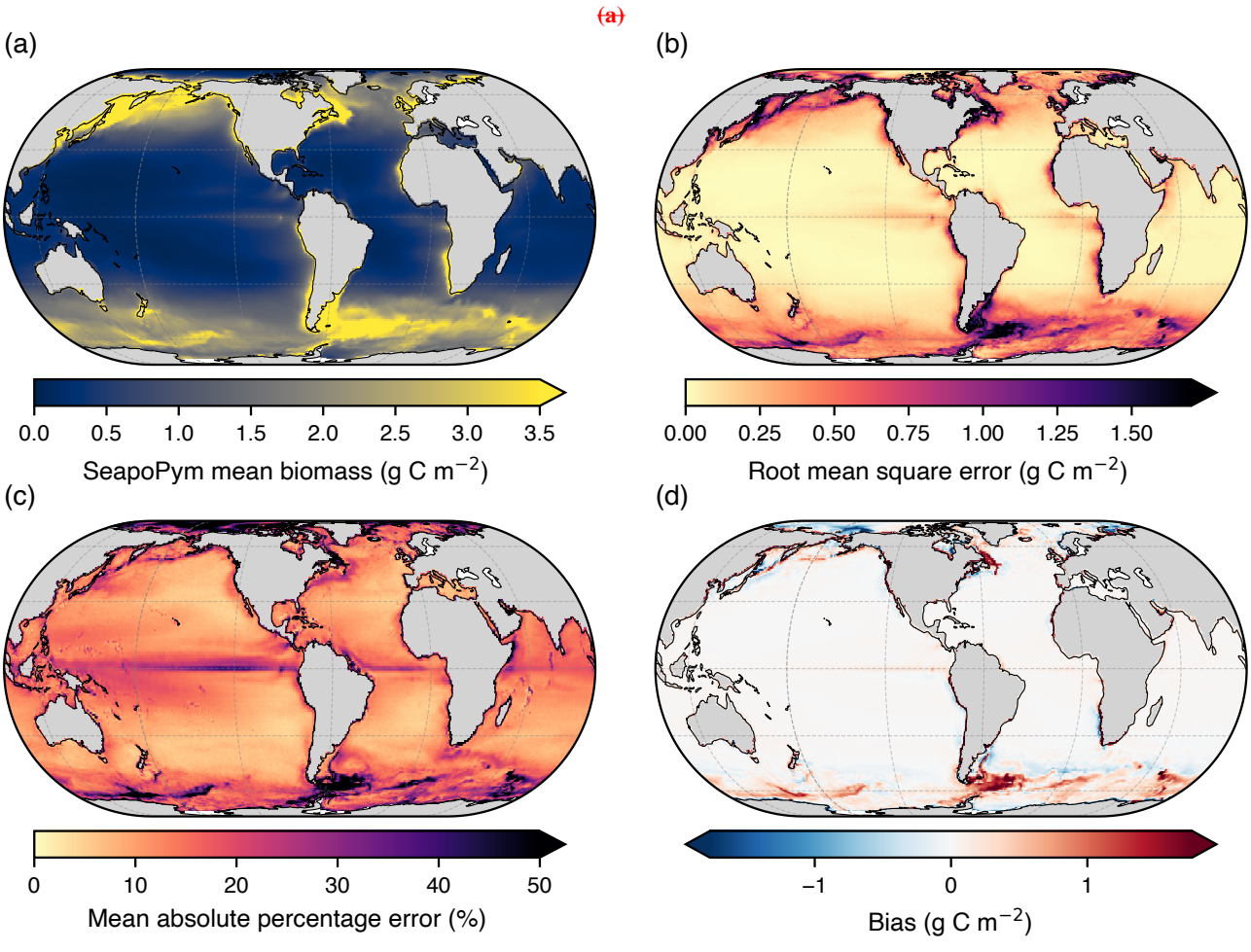

Figure 4. Effect of neglecting transport, shown as SeapoPym (0D) against the operational SEAPODYM-LMTL product (2D), as time means over 2000 to 2019. (a) Mean biomass distribution of zooplankton predicted by SeapoPym ; mean biomass, on the same scale as the reference biomass in Fig. 1a. (b) Root mean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

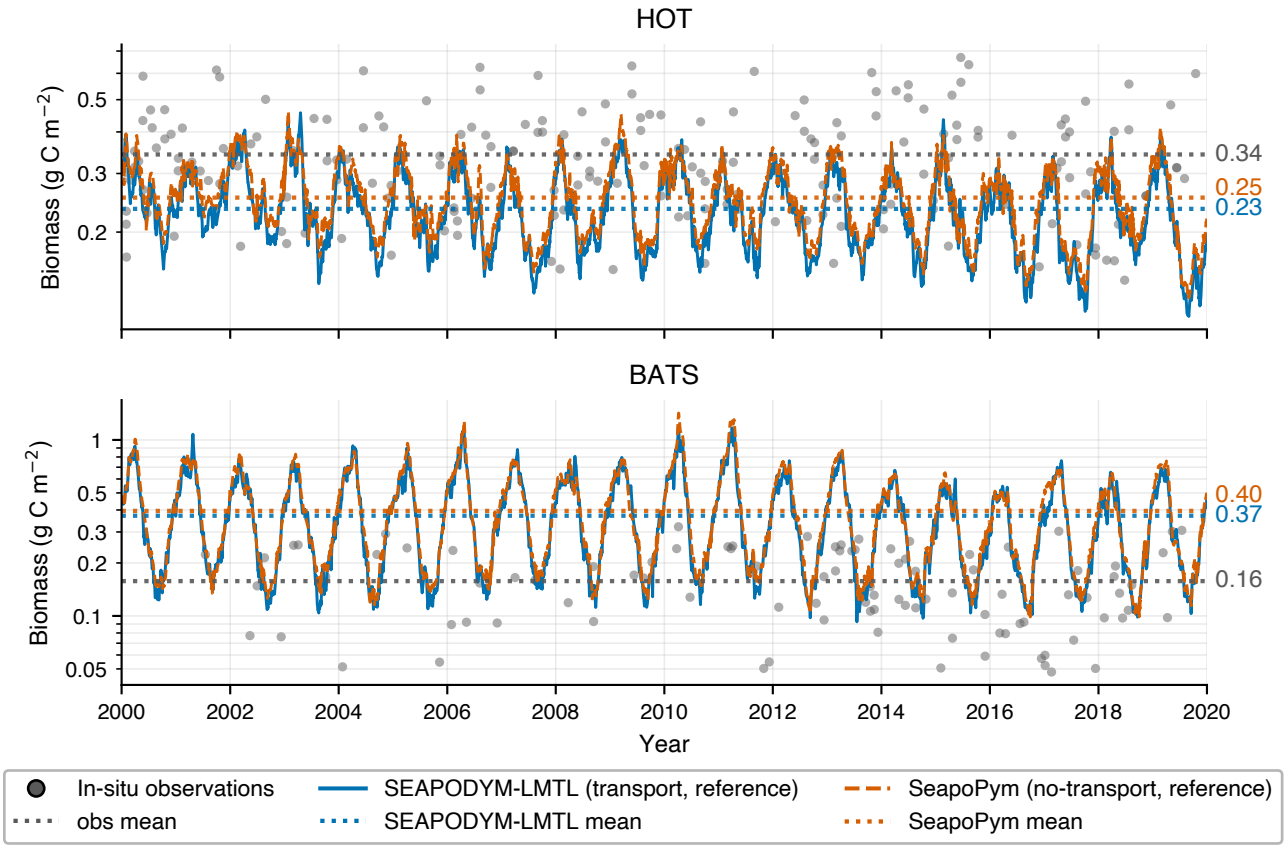

Figure 5. ~~Sobol analysis for three metrics (rows) and five parameters (columns)~~ Model biomass against in-situ observations at ~~six~~ the two stations with open long-term records, HOT (~~x-axis~~top) ~~, totaling 1,190,700 simulations~~ and BATS (~~N=99,225 samples~~bottom). Grey dots are the daily-mean in-situ mesozooplankton biomass, clipped to the 5th to 95th percentile. The ~~first-order index S_1 is shown with~~ solid blue ~~bars~~ line is the operational SEAPODYM-LMTL product and the ~~total order index S_T with~~ dashed orange ~~bars~~line is SeapoPym (without transport), both driven by the same forcing and parameters. The ~~error bars represent bootstrap 95 % confidence intervals~~dotted horizontal lines mark the time-mean of each series, with their values printed at the right. ~~Note that stations are sorted along~~ The y-axis is logarithmic and biomass is in g C m^{-2}. The two model means nearly coincide while the ~~x-axis from coldest~~ observed mean sits apart, so the difference between the models, which is due to ~~warmest conditions~~transport and a different implementation, is five to seven times smaller than the gap between the operational SEAPODYM-LMTL product and the observations.

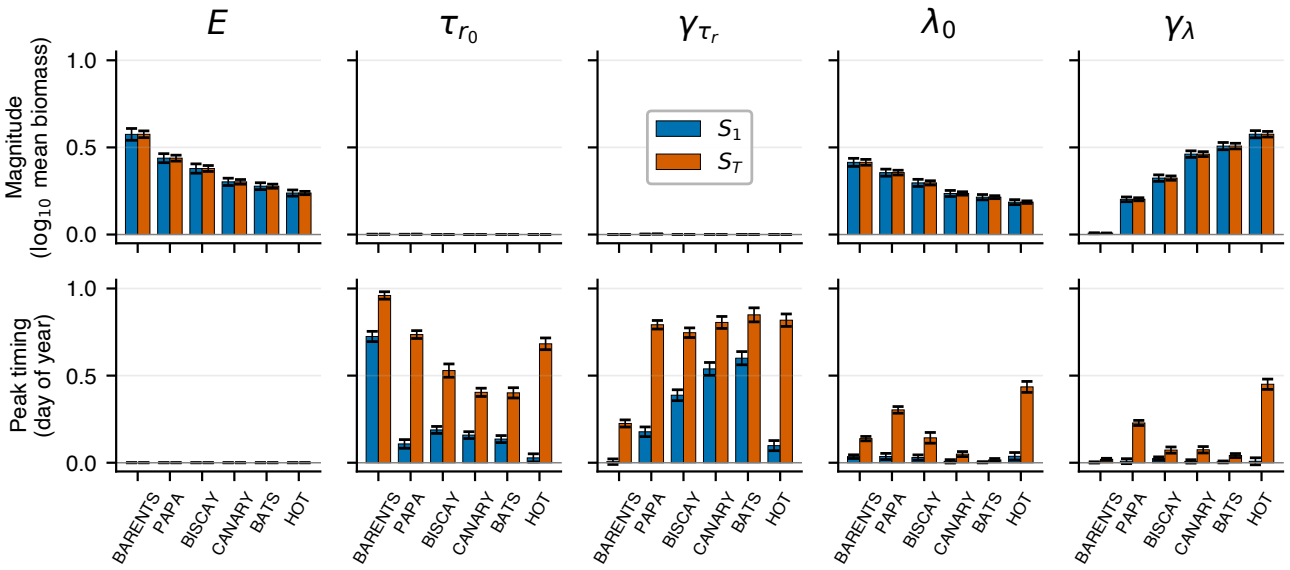

Figure 6. ~~Evolution~~ Sobol sensitivity indices of the ~~NRMSE for the predicted~~ biomass to the five biological parameters at ~~each station for optimization experiments considering each station separately~~ the six stations, ordered from coldest to warmest. The rows are the two output descriptors, the magnitude (base-10 logarithm of the mean biomass, top) ~~or together~~ and the seasonal peak timing (day of year of the maximum, bottom). The columns are the five parameters E, τ_{r_0}, γ_{τ_r}, λ_0 and γ_λ. 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.

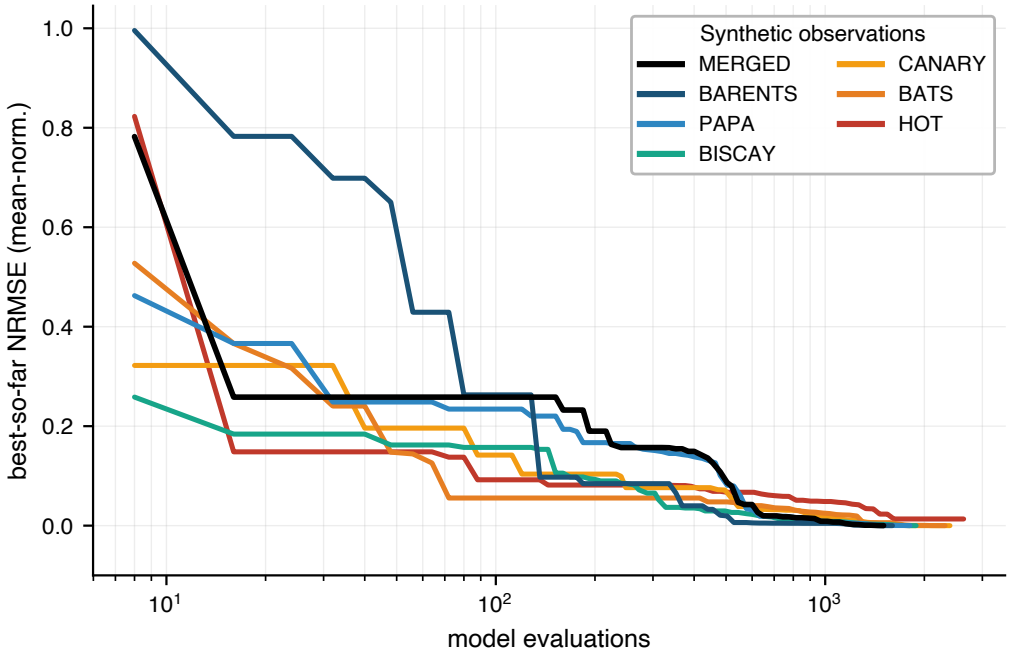

Figure 7. ~~Evolution through~~ CMA-ES convergence for the ~~optimization process~~ seven twin experiments, each the best of twenty random restarts. Each line is the ~~Mean Absolute Percentage Error (MAPE)~~ best-so-far cost against the number of model evaluations, for ~~each parameter for twin~~ the six single-station experiments and the joint MERGED experiment. The cost is the NRMSE between the recovered and the synthetic biomass, ~~with an independent optimization for each simulated observation time~~ normalized by the mean of the target series~~(BARENTS~~. 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, ~~BATS~~which floors near 10^{-2}. Because the target is the model's own output a zero cost is attainable, ~~BISCAY, CANARY, PAPA) or~~ so the HOT residual likely reflects an incomplete search rather than a ~~joint optimization (MERGED)~~structural floor.

Table 2. ~~Highest density interval (HDI) evolution of~~ Parameter recovery in the ~~1~~twin experiments, ~~000~~ where the optimizer recovers known parameters from synthetic observations that the model generated itself, taking the best ~~individuals per generation for each~~ of twenty CMA-ES restarts. Each entry is the signed relative divergence of a recovered parameter ~~estimate~~ against its synthetic-observation reference, in ~~twin simulation experiments using all~~ percent. The last column is the achieved cost (~~MERGED~~mean-normalized NRMSE)~~or separate~~. 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 (station)	E 0.1668	τ_{r_0} 10.38	γ_{τ_r} -0.11	λ_0 0.006667	γ_λ 0.15	NRMSE
MERGED	0	0	0	0	0	1.1×10^{-4}
BARENTS	0	0	0	0	0	4.9×10^{-6}
PAPA	0	0	0	0	0	3.5×10^{-5}
BISCAY	0	$+0.1$	0	0	0	1.4×10^{-5}
CANARY	0	-77	$+78$	0	0	1.9×10^{-8}
BATS	0	$+158$	-41	0	0	1.7×10^{-6}
HOT	$+0.1$	-0.5	$+0.2$	$+27$	-7.4	1.3×10^{-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 $\tau_r = \tau_{r0} \exp(\gamma_{\tau_r} \bar{T})$ (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	τ_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

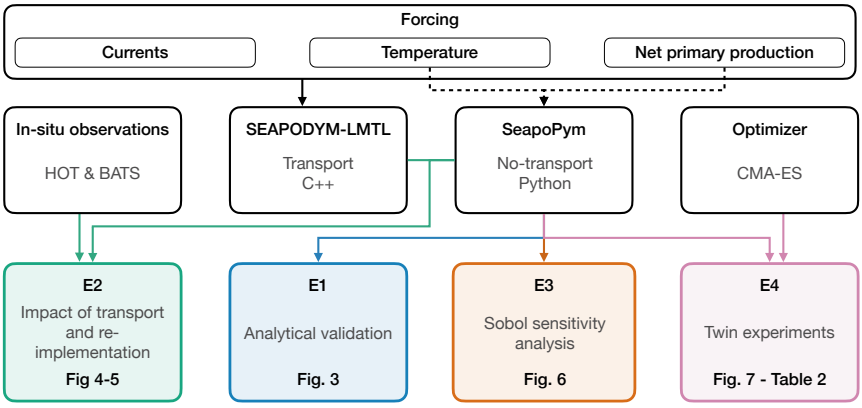

Figure D1. Overview of the SeapoPym modeling and experimental workflow. Temperature and NPP drive both models, whereas the currents drive only the SEAPODYM-LMTL model (with transport, C++). The dashed link marks that SeapoPym (without transport, Python) omits the currents. Each experiment is color-coded 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.