Supplement of Biogeosciences, 23, 4967–4998, 2026
https://doi.org/10.5194/bg-23-4967-2026-supplement

Biogeosciences
Open Access

Supplement of

High-dimensional parameter optimization of a biogeochemical model: a multi-variable BGC-Argo data assimilation approach

Quentin Hyvernat et al.

Correspondence to: Quentin Hyvernat (qhyvernat@mercator-ocean.fr)

Supplementary Information

This Supplementary Information accompanies the main text and is organised as follows. Section S1 provides the full mathematical definitions of the statistical quantities summarised verbally in the main text (Eqs. S1–S14). Section S2 describes the model–observation colocation procedure and the adaptation of the NRMSE metric for the three-dimensional validation. Section S3 contains the parameter prior ranges and constraints (Table S1). Section S4 presents the detailed parameter optimisation results for all three strategies (Table S2). Section S5 collects the supplementary figures referenced in the main text (Figs. S2–S25).

S1 Mathematical formulation

This section provides the full mathematical definitions of the statistical quantities used in the main text.

S1.1 Sobol Indices

The first-order index is calculated as

$$\boldsymbol{S}_i = \frac{\mathbb{V}ar_{\mathbf{X}_i}\left[\mathbb{E}_{\mathbf{X}_{\sim i}}\left(f(\boldsymbol{\theta})|(\boldsymbol{\theta}_i)\right)\right]}{\mathbb{V}ar[f(\boldsymbol{\theta})]}, i = 1, \dots, d \qquad (S1)$$

and the total-order index, capturing both direct and interaction effects, is calculated as:

$$\boldsymbol{T}_i = 1 - \frac{\mathbb{V}ar_{\mathbf{X}_i}\left[\mathbb{E}_{\mathbf{X}_i}\left(f(\boldsymbol{\theta})|(\boldsymbol{\theta}_{\sim i})\right)\right]}{\mathbb{V}ar[f(\boldsymbol{\theta})]} = \frac{\mathbb{E}_{\mathbf{X}_{\sim i}}\left[\mathbb{V}ar_{\mathbf{X}_i}\left(f(\boldsymbol{\theta})|(\boldsymbol{\theta}_i)\right)\right]}{\mathbb{V}ar[f(\boldsymbol{\theta})]} \qquad (S2)$$

S1.2 Quantities of interest for the sensitivity analysis

The RMSE for a metric M and for a sample j is computed as:

$$\boldsymbol{Y}_j^{RMSE,M} = \sqrt{\frac{1}{T}\sum_{t\in\boldsymbol{\mathcal{T}}}\left(\boldsymbol{s}_{j,t}^M - \boldsymbol{y}_t^M\right)^2} \qquad (S3)$$

Where $\boldsymbol{\mathcal{T}}$ is the set of observation time indices and $T = |\boldsymbol{\mathcal{T}}|$ is the number of available observations. Here, $\boldsymbol{s}_{j,t}^M$ and $\boldsymbol{y}_t^M$ denote the restrictions of the model state vector and the observation vector, respectively, to metric M at time index t.

The Correlation Coefficient (ρ) for each metric and sample is computed as follow:

$$\boldsymbol{Y}_j^{\rho,M} = \frac{\sum_{t\in\mathcal{T}}\left(\boldsymbol{s}_{j,t}^M - \overline{\boldsymbol{s}_j^M}\right)\left(y_{j,t}^M - \overline{\boldsymbol{y}_t^M}\right)}{\sqrt{\sum_{t\in\mathcal{T}}\left(\boldsymbol{s}_{j,t}^M - \overline{\boldsymbol{s}_j^M}\right)^2 \sum_{t\in\mathcal{T}}\left(\boldsymbol{y}_{j,t}^M - \overline{\boldsymbol{y}_t^M}\right)^2}} \qquad (S4)$$

Where the $\overline{\mathbf{s}_\mathrm{j}^\mathrm{M}}$ and $\overline{\mathbf{y}^\mathrm{M}}$ denote the time average of the model and observation metrics.

S1.3 Normalised Root Mean Square Error

We calculate two versions of the NRMSE for each metric M: one for the single best ensemble member $\boldsymbol{s}_{Best}^M$ and one for the full weighted-mean ensemble $\overline{\boldsymbol{s}_t^M}^{\omega}$. The NRMSE for the best member is:

$$\boldsymbol{NRMSE}_{Best}^M = \sqrt{\frac{1}{T}\sum_{t\in\mathcal{T}}\left(\frac{\boldsymbol{y}_t^M - \boldsymbol{s}_{Best,t}^M}{\boldsymbol{\varepsilon}_t^M}\right)^2} \qquad (S5)$$

Here, $\boldsymbol{y}_t^M$ is the observation of metric M at time t, $\boldsymbol{s}_{Best,t}^M$ is the corresponding value from the single best ensemble member, and $\boldsymbol{\varepsilon}_t^M$ is the total observational error (measurement and representativity errors combined in quadrature). For the depth-based metrics (H_{DCM}, $H_{nitracline}$, and H_{O2min}), the error $\boldsymbol{\varepsilon}_t^M$ is taken as the average of the asymmetric errors above and below the observed value . For computing the NRMSE for the weighted-mean of the new subset of 512 members we first define $\boldsymbol{p}_j$:

$$\boldsymbol{p}_j = \frac{\boldsymbol{\omega}_j}{\sum_{j\in\mathcal{N}}\boldsymbol{\omega}_j} \qquad (S6)$$

Where $\boldsymbol{p}_j$ is the normalized weight of each simulation for the jth ensemble member. With $\boldsymbol{\mathcal{N}}$, the set of simulation member indices and $N = |\boldsymbol{\mathcal{N}}|$ is the number of available simulations.

We also define the weighted average $\overline{\boldsymbol{s}_t^M}^{\omega}$:

$$\overline{\boldsymbol{s}_t^M}^{\omega} = \sum_{j\in\mathcal{N}} \boldsymbol{s}_{j,t}^M * \boldsymbol{p}_j \qquad (S7)$$

Where $\boldsymbol{s}_{j,t}^M$ is the value of metric M for the jth ensemble member at time , and with $\boldsymbol{p}_j$ the normalised weight of the jth member.

The NRMSE for the weighted-mean ensemble is given by:

$$\boldsymbol{NRMSE}_{Ensemble}^{M} = \sqrt{\frac{1}{T}\sum_{t\in\mathcal{T}}\left(\frac{\boldsymbol{y}_t^M - \overline{\boldsymbol{s}_t^M}^{\omega}}{\boldsymbol{\varepsilon}_t^M}\right)^2} \qquad (S8)$$

S1.4 Reduced Centred Random Variable

For computing the Reduced Centred Random Variable (RCRV) with the weighted standard deviation we first define $\boldsymbol{\sigma}_t^{M,\omega}$:

$$\boldsymbol{\sigma}_t^{M,\omega} = \sqrt{\sum_{j\in\boldsymbol{\mathcal{N}}}\left(\boldsymbol{s}_{j,t}^M - \overline{\boldsymbol{s}_t^M}^{\omega}\right)^2 * \boldsymbol{p}_j} \qquad (S9)$$

Where $\overline{\boldsymbol{s}_t^M}^{\omega}$ is the weighted ensemble mean simulation, $\boldsymbol{p}_j$ is the normalized weight of the jth member. With $\boldsymbol{\mathcal{N}}$, the set of simulation member indices and $N = |\boldsymbol{\mathcal{N}}|$ is the number of available simulations.

For each metric M at each time point t, the **RCRV** is calculated as:

$$\boldsymbol{RCRV}_t^M = \frac{\boldsymbol{y}_t^M - \overline{\boldsymbol{s}_t^M}^{\omega}}{\sqrt{{\boldsymbol{\varepsilon}_t^M}^2 + {\boldsymbol{\sigma}_t^{M,\omega}}^2}} \qquad (S10)$$

Where $\boldsymbol{y}_t^M$ is the observation, $\overline{\boldsymbol{s}_t^M}^{\omega}$ is the weighted ensemble mean simulation, $\boldsymbol{\varepsilon}_t^M$ is the total observational error (measurement + representativity), and $\boldsymbol{\sigma}_t^{M,\omega}$ is the weighted ensemble standard deviation at that time. For the depth-based metrics (H_{DCM}, $H_{nitracline}$, and H_{O2min}), the error $\boldsymbol{\varepsilon}_t^M$ used is the average of the asymmetric errors above and below the observed value.

We then summarize the full **RCRV** time series using two key statistics:

- Bias: Calculated as the mean of the **RCRV** time series. A value near zero indicates an unbiased ensemble.

$$\boldsymbol{RCRV}_{Bias}^{M} = \frac{1}{T}\sum_{t\in\mathcal{T}}\boldsymbol{RCRV}_t^M \qquad (S11)$$

- Dispersion: Calculated as the standard deviation of the **RCRV** time series. A value near one indicates a well-calibrated ensemble, where the model's predictive spread is consistent with the actual model-data errors. Values below one indicate a conservative (over-dispersive) ensemble, while values above one indicate an overconfident (under-dispersive) ensemble.

$$\boldsymbol{RCRV}_{std}^{M} = \sqrt{\frac{1}{T}\sum_{t\in\mathcal{T}}(\boldsymbol{RCRV}_{t}^{M} - \boldsymbol{RCRV}_{Bias}^{M})^2} \qquad (S12)$$

S1.5 Parameter uncertainty reduction

We quantify the reduction in parameter uncertainty by computing the percentage change in the width of the 67% HDI. The HDI represents the narrowest credible interval containing 67% of the probability mass, analogous to a one-standard-deviation range. The percentage reduction ($\Delta HDI_{67\%}$) is calculated as:

$$\Delta HDI_{67\%} = \left(1 - \frac{HDI_{67\%}^{posterior}}{HDI_{67\%}^{prior}}\right) * -1 \ * 100 \qquad (S13)$$

Here, $HDI_{67\%}^{prior}$ and $HDI_{67\%}^{posterior}$ are the widths of the 67% HDI for the initial and optimised distributions, respectively. A negative value indicates a reduction (a tightening) of parameter uncertainty.

S1.6 Observation likelihood function

The observations are assumed independent across metrics and time steps. For metrics with Gaussian errors, the likelihood for ensemble member j is the product over all metrics M and time steps t:

$$p(y|x_j) = \prod_{M\in\mathcal{G}}\prod_{t\in\mathcal{T}}\frac{1}{\sqrt{2\pi\varepsilon_t^M}} \exp\left[-\frac{1}{2}\left(\frac{y_t^M - s_{j,t}^M}{\varepsilon_t^M}\right)^2\right]$$
$$\times \prod_{M\in\mathcal{U}}\prod_{t\in\mathcal{T}}\mathbb{1}\left[|y_t^M - s_{j,t}^M| \le \varepsilon_t^M\right] \qquad (S14)$$

where $\mathcal{G}$ is the set of metrics with Gaussian errors (layer-mean concentrations), $\mathcal{U}$ is the set of metrics with uniform errors (H_{DCM}, $H_{nitracline}$, H_{O2min}), y^M_t is the observation of metric M at time t, $s^M_{j,t}$ is the corresponding model value for ensemble member j, ε^M_t is the total observation error (measurement and representativity errors combined in quadrature; Sects. 3.6–3.7 of the main text), and $\mathbb{1}[\cdot]$ is the indicator function equal to 1 when the condition is satisfied and 0 otherwise. For the uniform terms, a particle whose simulated depth-based metric falls outside the observation error bounds receives zero weight and is effectively excluded from the posterior.

S2. Three-dimensional validation: colocation procedure and metric adaptation

S2.1 Model–Observation Colocation

For the BGC-Argo validation, a matchup procedure colocates IBI model outputs with float observations in space, depth, and time over the 0–500 m range. This depth range is expected to encompass the maximum mixed-layer depth, and therefore the full productive layer, across all seasons and sub-regions. Below this horizon, biogeochemical concentrations in the 3D simulations remain largely stable over the three-year period and close to the initial conditions; including these deeper layers would therefore bias the comparison towards artificially good agreement that reflects the quality of the initialisation rather than the performance of the optimized parameters. Model values are matched to each float position using an inverse distance weighting of the four nearest grid cells, then linearly interpolated vertically to the observation depths. Temporally, daily model outputs are linearly interpolated to the exact observation time. Model concentrations are converted from model units (in mmol m^{-3}) to observation units (in µmol kg^{-1}) using seawater density derived from the model's own temperature and salinity profiles, so that all comparisons are performed in the native unit of the BGC-Argo observations. Outliers are removed via model-to-observation ratio filtering and interquartile range filtering (with a factor of 2.5 × IQR).

Model surface chlorophyll-a is compared to the MOB-TAC gridded satellite-derived product. Model Chl-a outputs from the REF and OPTI simulations are projected onto the MOBTAC 0.25° grid by nearest-neighbour interpolation. Daily model outputs and weekly MOB-TAC product are then monthly averaged, and meantime series are computed for each box defined in Sect. 3.10. Statistics are computed in $\log_{10}$-transformed space, consistent with the log-normal distribution of Chl-a concentrations. A Taylor diagram summarizes the performances in each box, in terms of the normalized standard deviation ($\sigma_{model}/\sigma_{obs}$), Pearson correlation coefficient, and centred root-mean-square difference. The MOBTAC product is represented by the black filled circle at unit normalised standard deviation, and the black dashed semicircle indicates where model and observed variability are equal ($\sigma_{model}/\sigma_{obs} = 1$); the green dotted contours show isolines of centred RMSE.

S2.2 Validation Metric for the 3D Assessment

Model performance in the 3D validation is assessed using the same NRMSE framework defined in Sect. 3.9, applied here to individual observation–model matchup points rather than to time-averaged metrics. Two methodological differences with respect to the 1D implementation are noted and justified below.

First, the representativity error term (Sect. 3.7) is omitted from the denominator. This term was introduced in the 1D framework to account for three-dimensional physical processes that the 1D model cannot resolve. Since the IBI-configuration explicitly simulates these processes, the representativity correction is no longer needed.

Second, the observation uncertainty σ_{obs}, is used to normalise each model–observation residual differs from the percentages reported in Table 2. In the 1D framework (Sect. 3.6.4), the measurement errors for CANYON-B and CONTENT variables were computed by propagating the O_2 input uncertainty through the neural networks, combining the result in quadrature with the intrinsic network error, and averaging into a single percentage per variable. Here, the normalisation uses instead the point-by-point error field associated with each BGC-Argo measurement. These fields reflect the intrinsic uncertainty of the neural-network derivation at each location and depth, but do not include the propagated O_2 input term. The normalising uncertainty is therefore somewhat smaller, making the NRMSE threshold of unity a slightly more demanding criterion. This choice is motivated by the fact that repeating the Monte Carlo propagation for approximately 470,000 matchup points would be computationally prohibitive.

The interpretation of the NRMSE remains identical to Sect. 3.9: values below unity indicate that model errors fall, on average, within observation uncertainty. The Δ%NRMSE is computed as defined in Sect. 3.9. When both NRMSE values fall below unity, a positive Δ%NRMSE does not constitute a meaningful degradation, since both simulations already reproduce observations within their associated uncertainties.

S3. Parameter constraints

Table S1. Ranges of values for the biogeochemical model parameters and their additional constraints.

Descriptions (unit)	Symbol	Range	Additional constraints
POC sinking speed (m d^{-1})	ω_{POC}	0.02-4	Must be less than ω_{GOC}
Big particles sinking speed (m d^{-1})	ω_{GOC}	0.5-100	

NO_3^- HS of Nanophytoplankton (mol N L^{-1})	$K_{NO_3}^{Nano}$	0.01-2	Must be equal to $K_{NH_4}^{Nano}$
NO_3^- HS of diatoms (mol N L^{-1})	$K_{NO_3}^{Diatoms}$	0.03-6	Must be equal to $K_{NH_4}^{Diatoms}$
NO_3^- HS for DOC remin (mol N L^{-1})	$K_{NO_3}^{Bact}$	0.003-0.6	Must be equal to $K_{NH_4}^{Bact}$
NO_3^- HS for DOC remin (mol N L^{-1})	$K_{NO_3}^{Bact}$	0.003-0.6	Must be equal to $K_{NH_4}^{Bact}$
Quadatric mortality of phytoplankton (d^{-1} mol C^{-1})	ω^P	0.0001-0.02	Must be less than ω_{max}^{Diat}
Maximum quadratic mortality of diatoms (d^{-1} mol C^{-1})	ω_{max}^{Diat}	0.0003-0.06	
Part of calcite not dissolved in MesoZoo guts (/)	ν^M	0.0075-1	Must be between 0 and 1
Part of calcite not dissolved in MicroZoo guts (/)	ν^Z	0.0075-1	Must be between 0 and 1
Maximal MesoZoo grazing rate (d^{-1})	g_m^M	0.005-1	Must be between 0 and 1
Maximal MicroZoo grazing rate (d^{-1})	g_m^Z	0.02-4	Must be 2 times greater than or equal to g_m^M
MesoZoo preference for nanophytoplankton (/)	P_p^M	0.01-1	Must be less than P_D^M and P_Z^M
MesoZoo preference for diatoms (/)	P_D^M	0.003-0.6	Must be greater than P_P^M and P_{POC}^M
MesoZoo preference for POC(/)	P_{POC}^M	0.003-0.6	Must be less than P_D^M and P_Z^M
MesoZoo preference for microzooplankton (/)	P_Z^M	0.01-1	Must be greater than P_P^M and P_{POC}^M
MicroZoo preference for nanophytoplankton (/)	P_p^Z	0.01-1	Must be greater than P_D^Z and P_{POC}^Z
MicroZoo preference for diatoms (/)	P_D^Z	0.008-1	Must be less than P_P^Z
MicroZoo preference for POC (/)	P_{POC}^Z	0.0015-0.3	Must be less than P_P^Z
Growth efficiency of MesoZoo (/)	e^M	0.004-0.8	Sum of e^M and σ^M must be less than or equal to 1
Non-assimilated fraction by MesoZoo (/)	σ^M	0.003-0.6	Sum of e^M and σ^M must be less than or equal to 1.
Growth efficiency of MicroZoo (/)	e^Z	0.004-0.8	Sum of e^Z and σ^Z must be less than or equal to 1
Non-assimilated fraction by MicroZoo (/)	σ^Z	0.003-0.6	Sum of e^Z and σ^Z must be less than or equal to 1.
Fraction of MesoZoo excretion as DOM	γ^M	0.006-1	Must be between 0 and 1
Fraction of MicroZoo excretion as DOM	γ^Z	0.006-1	Must be between 0 and 1
Linear mortality rate of Zoo. (d^{-1})	r^Z	0.0002-0.04	Must be 2 times greater than r^M
Exsudation rate of MesoZoo. (d^{-1})	r^M	0.00005-0.01	

S4. Parameter optimisation results

Table S2. Parameter optimization results. (a) Model parameters for microzooplankton. (b) Model parameters for mesozooplankton. (c) Model parameters for phytoplankton growth. (d) Model parameters for nutrient limitations. (e) Model parameters biological. (f) Model parameters for organic particles. (g) Model parameters for phytoplankton sinks. (h) Model parameters for remineralization. (i) Model parameters for iron chemistry. (j) Model parameters for calcite chemistry. (k) Model parameters for input deposition. (l) Model parameters for sediment mobilization. The 'Ref' column indicates the reference values used in the standard configuration of the PISCES model. The 'Best' column reports the parameter value implemented in the member with the lowest cost function during the optimization procedure. A dash in any column indicates that the corresponding parameter was not modified as part of the optimization of the parameters selected for this sensitivity test. An estimate of the uncertainty associated with the "Best" values obtained when optimizing all parameters simultaneously is provided by the lower and upper bounds of the 67% HDI.

(a)

Symbol	Description (Units)	Ref	Best Main effects	Best Total effects	Best All parameters (HDI)
$P_{Diat}^{MicroZoo}$	MicroZoo. preference for Diatoms	0.8	0.13 (8e-3 – 8.58e-2)	8.44e-2 (8e-3 – 9.57e-2)	0.29 (8e-3 – 0.12)
$P_{Nano}^{MicroZoo}$	MicroZoo. Preference for NanoPython	1	0.32 (0.17 – 0.41)	0.97 (0.16 – 0.64)	0.6 (0.17 – 0.58)
$K_{G}^{MicroZoo}$	H-Sat. constant for MicroZoo. grazing (µmol C L^{-1})	2e-5	3.49e-5 (2.78e-5 – 3.98e-5)	2.26e-5 (1.45e-5 – 3.16e-5)	2.65e-5 (2.21e-5 – 3.74e-5)
$e^{MicroZoo}$	Efficiency of MicroZoo. growth	0.4	0.5 (0.15 – 0.46)	0.23 (8.76e-2 – 0.46)	0.37 (0.23 – 0.59)
$e_{Min}^{MicroZoo}$	Min efficiency of MicroZoo. growth	0.4	0.5 (0.15 – 0.46)	0.23 (8.76e-2 – 0.46)	0.37 (0.23 – 0.59)
$G_{m}^{MicroZoo}$	Max. MicroZoo. grazing rate (d^{-1})	2	2.19 (1.69– 2.81)	1.96 (1.7 – 3.25)	1.81 (1.69 – 3.06)
$\gamma^{MicroZoo}$	Fraction of MicroZoo excretion as DOM	0.6	-	0.27 (6e-3 – 0.56)	0.4 (6e-3 – 0.47)
$\upsilon^{MicroZoo}$	Part of calcite not dissolved in MicroZoo	0.75	0.12 (7.5e-3 – 0.48)	0.4 (7.5e-3 – 0.51)	0.5 (7.5e-3 – 0.55)
$m^{MicroZoo}$	MicroZoo. mortality rate ((mol L^{-1})$^{-1}$ d^{-1})	0.005	-	8.61e-3 (4.4e-3 – 8.93e-3)	9.4e-3 (4.42e-3 – 9.46e-3)
$P_{POM}^{MicroZoo}$	MicroZoo. preference for POM	0.15	6.44e-2 (1.5e-3 – 8.51e-2)	0.19 (1.5e-3 – 0.11)	0.12 (1.5e-3 – 0.11)
$\phi_{Scalling}^{MicroZoo}$	Predation window size scalling MicroZoo	1	-	-	0.54 (1e-2 – 0.94)
$F_{Tresh}^{MicroZoo}$	Food threshold for feeding (µmol C L^{-1})	3e-7	-	-	2.01e-7 (8.12e-8 – 3.8e-7)
$\phi^{MicroZoo}$	Predation window size MicroZoo	0.5	-	0.23 (5e-3 – 0.63)	0.76 (0.21 – 0.86)
$\sigma^{MicroZoo}$	Non assimilated fraction of NanoPhyto. by MicroZoo	0.3	6.31e-3 (3e-3 – 0.29)	7.6e-2 (0.14 – 0.41)	0.25 (0.13 – 0.47)
$r^{MicroZoo}$	Linear mortality rate of MicroZoo (d^{-1})	0.02	-	3.29e-2 (1.89e-2 – 3.41e-2)	2.95e-2 (1.69e-2 – 3.36e-2)

Table S2. Continued

(b)

Symbol	Description (Units)	Ref	Best Main effects	Best Total effects	Best All parameters

$P_{POC}^{MesoZoo}$	MesoZoo. preference for POC	0.3	-	5.5e-2 (3e-3 – 8.83e-2)	0.32 (3e-3 – 0.1)
$P_{Nano}^{MesoZoo}$	MesZoo. preference for NanoPhyto	0.3	-	4.24e-2 (3e-3 – 9.5e-2)	3.96e-2 (3e-3 – 6.03e-2)
$e^{MesoZoo}$	Efficiency of MesoZoo growth	0.4	-	0.55 (0.38 – 0.63)	0.48 (0.36 – 0.67)
$e_{Min}^{MesoZoo}$	Min Efficiency of MesoZoo growth	0.4	-	0.55 (0.38 – 0.63)	0.48 (0.36 – 0.67)
$G_{m}^{MesoZoo}$	Max. MesoZoo. grazing rate (d^{-1})	0.5	0.4 (0.57– 0.95)	0.83 (0.49– 0.89)	0.53 (0.54– 0.9)
$\sigma^{MesoZoo}$	Non assimilated fraction of NanoPhyto. by MesoZoo.	0.3	-	0.24 (0.13– 0.35)	0.32 (7.83e-2– 0.32)
$K_{G}^{MesoZoo}$	H-Sat. constant for MesoZoo. grazing (µmol C L^{-1})	2e-5	8.91e-6 (1.46e-5 – 2.73e-5)	3.13e-5 (5.18e-6 – 2.26e-5)	1.38e-5 (5.89e-6 – 2.32e-5)
$P_{MicroZoo}^{MesoZoo}$	MesoZoo. preference for MicroZoo	1	-	0.13 (0.43 – 0.88)	0.52 (0.64 – 1.)
$P_{Diat}^{MesoZoo}$	MesoZoo. Preference for diatoms	1	-	0.41 (0.22 – 0.77)	0.88 (0.33 – 0.95)
$g_{FF}^{MicroZoo.}$	Flux-feeding rate ((mmol $L^{-1})^{-1}$)	3e3	-	5.04e3 (1.57e3 – 5.37e3)	3.46e3 (1.58e3 – 4.63e3)
$\phi_{Scalling}^{MesoZoo}$	Predation window size scalling MesoZoo	1	-	-	1.83 (0.62 – 1.85)
$F_{Tresh}^{MesooZoo}$	Food threshold for grazing (µmol C L^{-1})	3e-7	-	-	4.89e-8 (3.e-9 – 2.93e-7)
$\upsilon^{MesoZoo}$	Part of calcite not dissolved in MesoZoo	0.75	-	0.41 (7.5e-3 – 0.55)	0.71 (0.31 – 0.95)
$m^{MesoZoo}$	MesoZoo. mortality rate ((mol $L^{-1})^{-1}$ d^{-1})	0.01	-	1.61e-2 (6.13e-3 – 1.89e-2)	6.84e-3 (2.17e-3 – 1.23e-2)
$\phi^{MesooZoo}$	Predation window size MesoZoo	0.5	-	-	0.9 (5e-3 – 0.56)
$\gamma^{MesoZoo}$	Fraction of MesoZoo excretion as DOM	0.6	-	-	0.93 (6.e-3 – 0.49)
$r^{MesoZoo}$	Exudation rate of MesoZoo (d^{-1})	0.005	-	1.85e-3 (5.67e-4 – 6.78e-2)	8.02e-3 (3.2e-3 – 9.4e-3)

Table S2. Continued

(c)

Symbol	Description (Units)	Ref	Best Main effects	Best Total effects	Best All parameters
$\theta_{Max}^{Chl\ Nano}$	Max. Chl-a/C in NanoPhyto (mg Chl-a (mg C$)^{-1}$)	0.033	1.1e-2 (7.23e-3 – 1.26e-2)	6.15e-3 (5.52e-3 – 1.42e-2)	1.11e-2 (6.77e-3 – 1.42e-2)
α^{Nano}	P-I slope NanoPhyto ((W $m^{-2})^{-1}$ d^{-1})	2	1.2 (0.86 – 1.39)	1.87 (0.92 – 1.68)	1.34 (1.07 – 1.88)
$\theta_{Max}^{Chl\ Diat}$	Max. Chl-a /C in Diatoms (mg Chl-a (mg C$)^{-1}$)	0.05	3.65e-2 (2.66e-2 – 6.4e-2)	2.88e-2 (1.55e-2 – 5.14e-2)	9.13e-2 (1.61e-2 – 5.8e-2)
$\theta_{Min}^{Chl\ Phyto}$	Min. Chl-a /C in Phyto (mg Chl-a (mg C$)^{-1}$)	0.003	-	-	3.53e-3 (3.15e-3 – 5.8e-3)
$\theta_{Max}^{Fe\ Diat}$	Max. Fe/C in Diatoms	6e-5	5.16e-5	1.09e-4	4.16e-5

	(µmol Fe mol C^{-1})		(2.88e-5– 8.28e-5)	(1.29e-5– 7.99e-5)	(2.59e-5– 9.03e-5)
$\theta_{Max}^{Fe\ Nano}$	Max. Fe/C in NanoPhyto (µmol Fe mol C^{-1})	6e-5	3.14e-5 (1.31e-5– 6.9e-5)	6.02e-5 (4.92e-5– 1.07e-4)	4.27e-5 (4.31e-5– 1.07e-4)
$\gamma^{Diatoms}$	Excretion ration of Diatoms	0.05	-	-	3.38e-2 (1.63e-2 – 7.22e-2)
α^{Diat}	P-I slope for Diatoms ((W m^{-2})$^{-1}$ d^{-1})	2	-	1.69 (1.69 – 3.78)	3.3 (1.27– 3.71)
K_{Si}	Mean Si/C ratio (µmol Si mol C^{-1})	0.13	0.25 (0.14 – 0.25)	3.25e-2 (5.51e-2 – 0.22)	0.21 (7.67e-2 – 0.25)
γ^{Nano}	Excretion ration of NanoPhyto	0.05	-	-	9.96e-2 (3.76e-2 – 9.47e-2)
b_{resp}	Basal respiration rate (d^{-1})	0.033	-	-	5.41e-2 (2.2e-2 – 5.9e-2)

Table S2. Continued

(d)

Symbol	Description (Units)	Ref	Best Main effects	Best Total effects	Best All parameters
$\theta_{Opt}^{Fe\ Nano}$	Optimal quota of NanoPhyto. (µmol Fe (mol C)$^{-1}$)	10e-6	2.63e-6 (1e-7 – 1.65e-6)	2.12e-6 (1.e-7 – 4.22e-6)	6.39e-7 (1e-7 – 6.23e-6)
K_{DOC}	H-Sat. constant of DOC remineralization (µmol C L^{-1})	417e-6	7.77e-4 (4.54e-4 – 7.87e-4)	6.32e-4 (4.89e-4 – 7.62e-4)	5.39e-4 (5.46e-4 – 7.89e-4)
R_{CaCO_3}	Mean rain ratio	0.2	5.34e-2 (2e-3 – 9.84e-2)	0.12 (2e-3 – 0.21)	5.54e-2 (2e-3 – 0.19)
K_{Fe}^{Nano}	Iron H-Sat. for NanoPhyto (nmol Fe L^{-1})	1.7e-9	-	3.27e-9 (1.68e-11 – 1.94e-9)	2.01e-9 (7.14e-10 – 2.33e-9)
K_{Si}^{Uptake}	H-Sat. constant for Si uptake (µmol Si L^{-1})	8e-6	-	5.96e-6 (2.77e-6 – 1.2e-5)	6.43e-6 (8e-8 – 9.82e-6)
$\theta_{Opt}^{Fe\ Diat}$	Optimal quota of Diatoms (µmol Fe (mol C)$^{-1}$)	10e-6	-	1.19e-5 (1.e-7 – 9.36e-6)	1.01e-5 (6.07e-6 – 1.88e-5)
S_{Min}^{Diat}	Min. size criteria for Diatoms (m)	1e-6	-	1.73e-6 (1.11e-7 – 1.35e-6)	9.e-7 (7.47e-7 – 1.89e-6)
K_{Fe}^{Bact}	Iron H-Sat. for for DOC remin (nmol Fe L^{-1})	3e-11	-	-	5.91e-11 (1.6e-11 – 5.22e-11)
K_{Si}	H-Sat. constant for Si/C (µmol Si L^{-1})	20e-6	-	9.92e-6 (6.37e-6 – 3.11e-5)	4.39e-6 (2e-7 – 2.6e-5)
S_{rat}^{Nano}	Size ratio for NanoPhyto	3	-	5.58 (1.6 – 5.04)	3.91 (2.52 – 5.68)
S_{Min}^{Nano}	Min. size criteria for NanoPhyto (m)	1e-6	-	-	6.45.e-7 (1.65e-7 – 1.37e-6)
O_2^{min}	H-Sat. constant for anoxia (µmol O_2 L^{-1})	1e-6	-	-	7.72.e-7 (1e-8 – 1.24e-6)

$K_{Fe}^{Diatoms}$	Iron H-Sat. for Diatoms (nmol Fe L^{-1})	5e-9	-	8.91e-9 (2.09e-9 – 7.39e-9)	4.17e-10 (2.38e-9 – 7.9e-9)
$S_{rat.}^{Diat.}$	Size ratio for Diatoms	4	-	7.84 (3.58 – 8.)	1.37 (3.39 – 7.48)
$K_{NO_3}^{Nano}$	NO_3^- HS of NanoPhyto (mol N L^{-1})	1e-6	-	1.75e-6 (8.42e-7 – 1.83e-6)	4.79e-7 (8.86e-7 – 2e-6)
$K_{NH_4}^{Nano}$	NH_4 HS of NanoPhyto (mol N L^{-1})	1e-6	-	1.75e-6 (8.42e-7 – 1.83e-6)	4.79e-7 (8.86e-7 – 2e-6)
$K_{NO_3}^{Bact}$	NO_3^- HS for DOC remin (mol N L^{-1})	3e-7	-	5.73e-7 (3.39e-7 – 5.68e-7)	4.52e-7 (3.8e-7 – 5.6e-7)
$K_{NH_4}^{Bact}$	NH_4 HS for DOC remin (mol N L^{-1})	3e-7	-	5.73e-7 (3.39e-7 – 5.68e-7)	4.52e-7 (3.8e-7 – 5.6e-7)
$K_{NH_4}^{Diatoms}$	NH_4 HS of Diatoms (mol N L^{-1})	3e-6	-	3.54e-6 (2.97e-6 – 5.68e-6)	1.88e-6 (2.92e-6 – 5.99e-6)
$K_{NO_3}^{Diatoms}$	NO_3^- HS of Diatoms (mol N L^{-1})	3e-6	-	3.54e-6 (2.97e-6 – 5.68e-6)	1.88e-6 (2.92e-6 – 5.99e-6)

Table S2. Continued

(e)

Symbol	Description (Units)	Ref	Best Main effects	Best Total effects	Best All parameters
$\theta^{Fe\ MesoZoo}$	Fe/C in MesoZoo (µmol Fe (mol C)$^{-1}$)	15e-6	-	2.22e-5 (1.5e-8 – 1.25e-5)	6.16e-6 (1.5e-8 – 1.27e-5)
K_m	H-Sat. constant for mortality (µmol C L^{-1})	1e-7	-	1.04e-7 (3.32e-8 – 1.36e-7)	1.09e-7 (9.22e-8 – 2e-7)
W_{POC}	POC seeking speed (m d^{-1})	2	0.28 (0.02 – 1.19)	0.61 (0.02 – 1.87)	0.3 (0.02 – 1.33)
$\theta^{Fe\ MicroZoo}$	Fe/C in MicroZoo (µmol Fe mol C^{-1})	10e-6	-	1.39e-5 (9.54e-6 – 1.9e-5)	1.48e-5 (8.17e-6 – 1.99e-5)
W_{GOC}	Big particles sinking speed (m d^{-1})	50	87.88 (5.57 – 67.72)	84.03 (11.25 – 68.22)	18.86 (0.5 – 61.45)
$W_{GOC\ Max}$	Max big particles sinking speed (m d^{-1})	50	87.88 (5.57 – 67.72)	84.03 (11.25 – 68.22)	18.86 (0.5 – 61.45)
W_{GOC}^{length}	Big particles length scale of sinking (m)	5e3	-	-	6.77e3 (2.11e3 – 8.31e3)

Table S2. Continued

(f)

Symbol	Description (Units)	Ref	Best Main effects	Best Total effects	Best All parameters
γ_{shape}	Shape of the gamma function	1	0.55 (0.22 – 0.85)	0.84 (0.29 – 1.37)	0.82 (0.53 – 1.7)

λ_{POC}	Remineralisation rate of POC (d^{-1})	0.035	5.52e-2 (1.13e-2 – 4.89e-2)	2.16e-2 (1.14e-2 – 4.79e-2)	3.78e-2 (1.51e-2 – 5.1e-2)

Table S2. Continued

(g)

Symbol	Description (Units)	Ref	Best Main effects	Best Total effects	Best All parameters
m^{Nano}	NanoPhyto mortality rate (d^{-1})	0.01	-	1.4e-2 (7.19e-3 – 1.89e-2)	1.44e-2 (1.01e-2 – 1.91e-2)
W^{Nano}	Quadratic mortality of NanoPhyto ((µmol C L^{-1})$^{-1}$ L^{-1})	0.01	-	1.9e-2 (1.e-4 – 1.21e-2)	1.05e-2 (1.e-4 – 9.3e-3)
$W_{Max}^{Diatoms}$	Maximum quadratic mortality of Diatoms ((µmol C L^{-1})$^{-1}$ L^{-1})	0.03	-	2.11e-2 (1.92e-2 – 4.74e-2)	2.01e-2 (2.22e-2 – 5.13e-2)
m^{Diat}	Diatoms mortality rate (d^{-1})	0.01	-	7.32e-3 (7.63e-3 – 1.89e-2)	1.39e-2 (5.18e-3 – 1.78e-2)

Table S2. Continued

(h)

Symbol	Description (Units)	Ref	Best Main effects	Best Total effects	Best All parameters
λ_{NH_4}	NH_4 nitrification rate (d^{-1})	0.05	4.78e-2 (5e-4 – 5.06e-2)	6.5e-3 (5e-4 – 5.89e-2)	2.11e-2 (5e-4 – 5.11e-2)
λ_{Si}	Remineralization rate of Si (d^{-1})	0.003	3.17e-3 (3e-5 – 3.13e-3)	5.05e-3 (3e-5 – 2.81e-3)	1.59e-3 (3e-5 – 3.14e-3)
$\theta^{Fe,Bact}$	Fe/C quota in Bacteria (µmol Fe mol C^{-1})	60e-6	-	-	9.25e-6 (6e-7 – 6.7e-5)
λ_{Si}^{fast}	Fast remineralization rate of Si (d^{-1})	0.03	9.77e-3 (3e-4 – 3.16e-2)	1.86e-2 (9.62e-3 – 4.67e-2)	1.04e-2 (6.63e-3 – 3.94e-2)
K_{Fe}^{Bact}	H-Sat. constant for bacteria Fe/C (µmol Fe mol C^{-1})	4e-10	-	-	3.42e-10 (7.81e-11 – 5.07e-10)
χ_{lab}^{O}	Fraction of labile biogenic silica	0.5	0.98 (5e-3 – 0.67)	0.25 (0.35 – 0.89)	0.85 (0.37 – 0.9)

Table S2. Continued

(i)

Symbol	Description (Units)	Ref	Best Main effects	Best Total effects	Best All parameters
λ_{Fe}^{dust}	Savenging rate of iron by dust (d^{-1} mg^{-1} L)	150	-	-	2.14e2 (64.2 – 2.21e2)

λ_{Fe}^{POFe}	Fraction of scavenged Fe that goes to POFe	1	-	-	0.34 (0.16 – 0.79)
λ_{Fe}	Scavenging rate of iron by biogenic particles (d^{-1})	0.02	-	-	2.83e-3 (2e-4– 2.27e-2)
$\lambda_{Nanopart}$	Nanoparticle formation rate constant (s^{-1})	0.01	-	-	4.31e-3 (1e-4–1.21e-2)

Table S2. Continued

(j)

Symbol	Description (Units)	Ref	Best Main effects	Best Total effects	Best All parameters
λ_{CaCO_3}	Calcite dissolution rate constant (d^{-1})	100	-	-	1.46e2 (1. – 1.32e2)

Table S2. Continued

(k)

Symbol	Description (Units)	Ref	Best Main effects	Best Total effects	Best All parameters
W_{dust}	Dust sinking speed (m s^{-1})	2	-	-	2.27 (2e-2 – 2.39)
υ_{dust}^{Fe}	Fe mineral fraction of dust	0.035	-	-	6.32e-2 (1.16e-2 – 5.17e-2)
$\upsilon_{^3He}^{Fe}$	Fe to ^{3}He ratio assumed for vent iron supply (mol Fe mol $^3He^{-1}$)	1e7	-	-	9.92e6 (7.99e6 – 1.89e7)
$Fe_{Fe,min}^{sed}$	Coastal release of Iron (mol Fe m^{-2} s^{-1})	2e-9	-	-	7.73e-10 (2e-11 – 2.32e-9)
Fe_{ice}	Iron concentration in sea ice (mol Fe L^{-1})	15e-9	-	-	4.78e-9 (9.53e-9 – 2.89e-8)

Table S2. Continued

(l)

Symbol	Description (Units)	Ref	Best Main effects	Best Total effects	Best All parameters
N_{fix}	Nitrogen fixation rate (µmol N L^{-1} d^{-1})	2e-7	-	-	1.75e-7 (2e-9 – 2.74e-7)
$K_{Fe}^{Diazotrophs}$	Diazotrophs H-Sat constant for iron (nmol Fe L^{-1})	1e-10	-	-	4.6e-11 (1.53e-11 – 1.26e-10)

$b_{light}^{Diazotrophs}$	Diazotrophs sensitivity to light (W m^{-2})	30	-	-	36.1 (0.3 – 37.6)

S5. Supplementary Figures

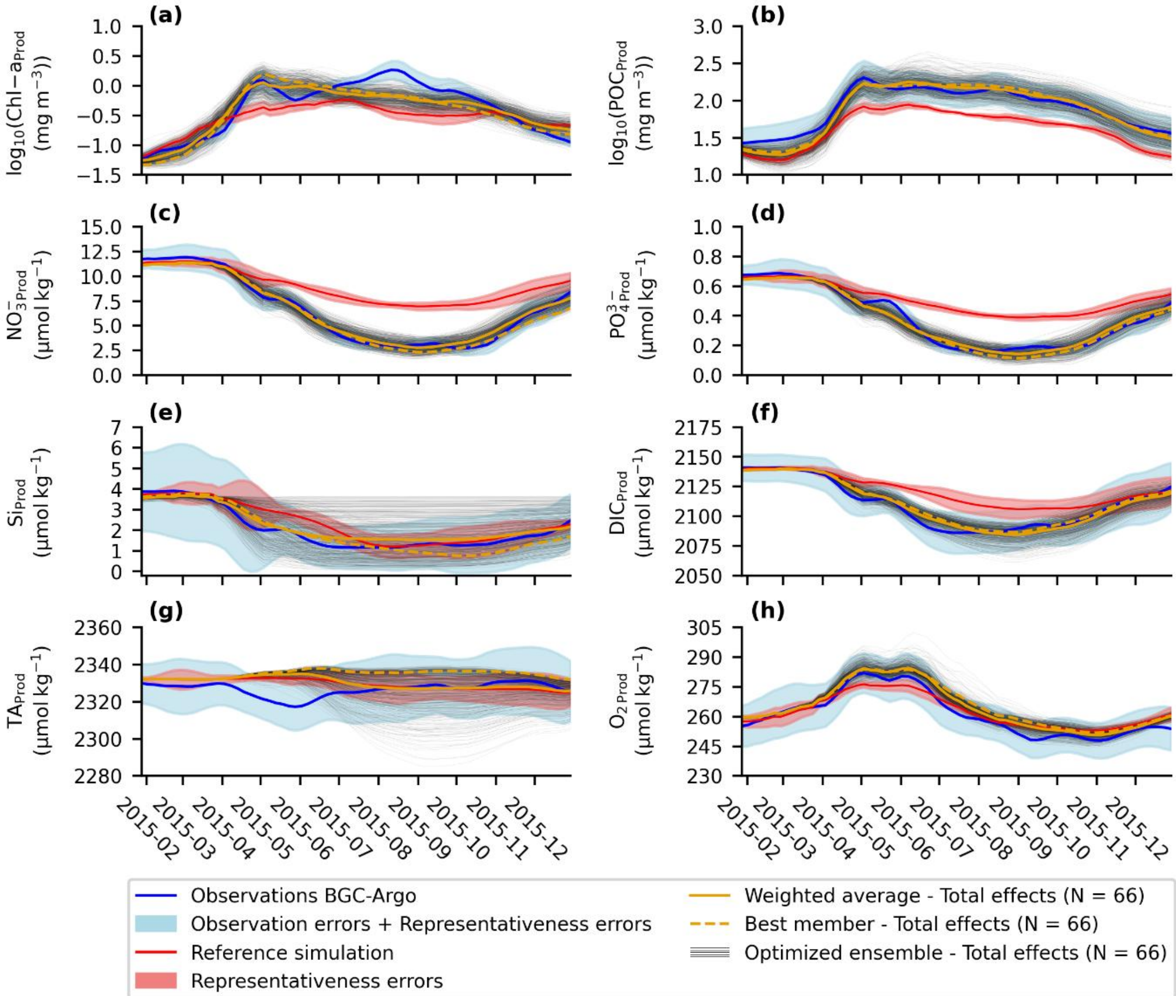

Figure S1. Seasonal cycle of assimilated metrics in the productive layer. Panels show : (**a**) $\log_{10}(\mathrm{Chl-a_{Prod}})$, (**b**) $\log_{10}(\mathrm{POC_{Prod}})$, (**c**) $\mathrm{NO_3^-}_{\mathrm{Prod}}$, (**d**) $PO_4^{3-}{}_{\mathrm{Prod}}$, (**e**) $\mathrm{Si_{Prod}}$, (**f**) $\mathrm{DIC_{Prod}}$, (**g**) $\mathrm{TA_{Prod}}$, and (**h**) $\mathrm{O_{2}}_{\mathrm{Prod}}$. The blue curve represents observations from BGC-Argo float #5904479, with the blue shading indicating the combined observations and representativity errors. The red curve corresponds to the reference simulation from PISCES-1D, with the red shading representing representativity errors. Orange line indicate the weighted means of the ensemble optimized using the Total effects parameters. The black curves represent the simulations of this ensemble. A six-point moving average was applied to all time series to smooth short-term fluctuations.

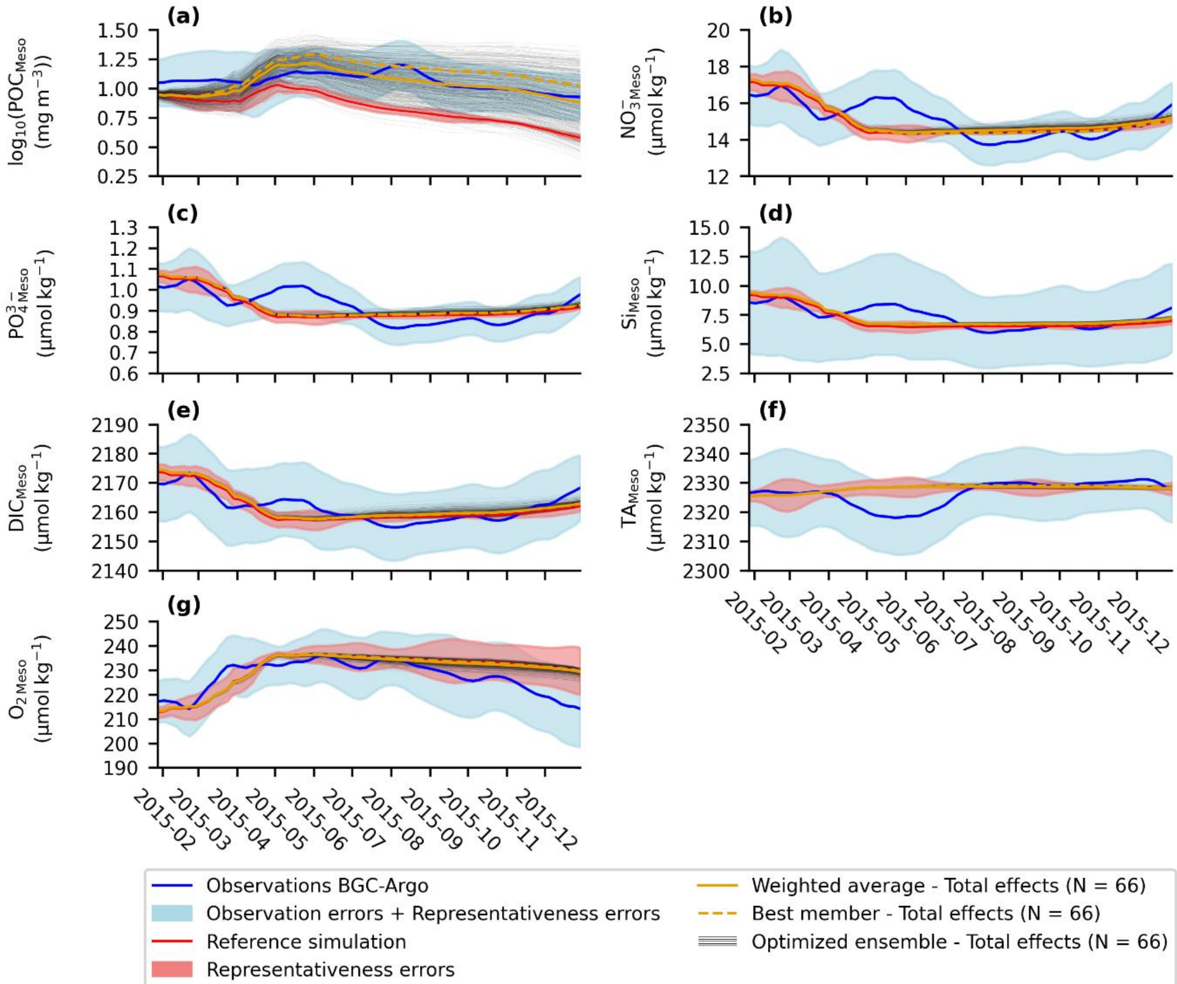

Figure S2. Seasonal cycle of assimilated metrics in the mesopelagic layer. Panels show : (**a**) $\log_{10}(\mathrm{POC_{Meso}})$ (**b**) $\mathrm{NO_3^-}_{\mathrm{Meso}}$, (**c**) $PO_4^{3-}{}_{\mathrm{Meso}}$, (**d**) $\mathrm{Si_{Meso}}$, (**e**) $\mathrm{DIC_{Meso}}$, (**f**) $\mathrm{TA_{Meso}}$, and (**g**) $\mathrm{O_2}_{\mathrm{Meso}}$. The blue curve represents observations from BGC-Argo float #5904479, with the blue shading indicating the combined observations and representativity errors. The red curve corresponds to the reference simulation from PISCES-1D, with the red shading representing representativity errors. Orange line indicate the weighted means of the ensemble optimized using the Total effects parameters. The black curves represent the simulations of this ensemble. A six-point moving average was applied to all time series to smooth short-term fluctuations.

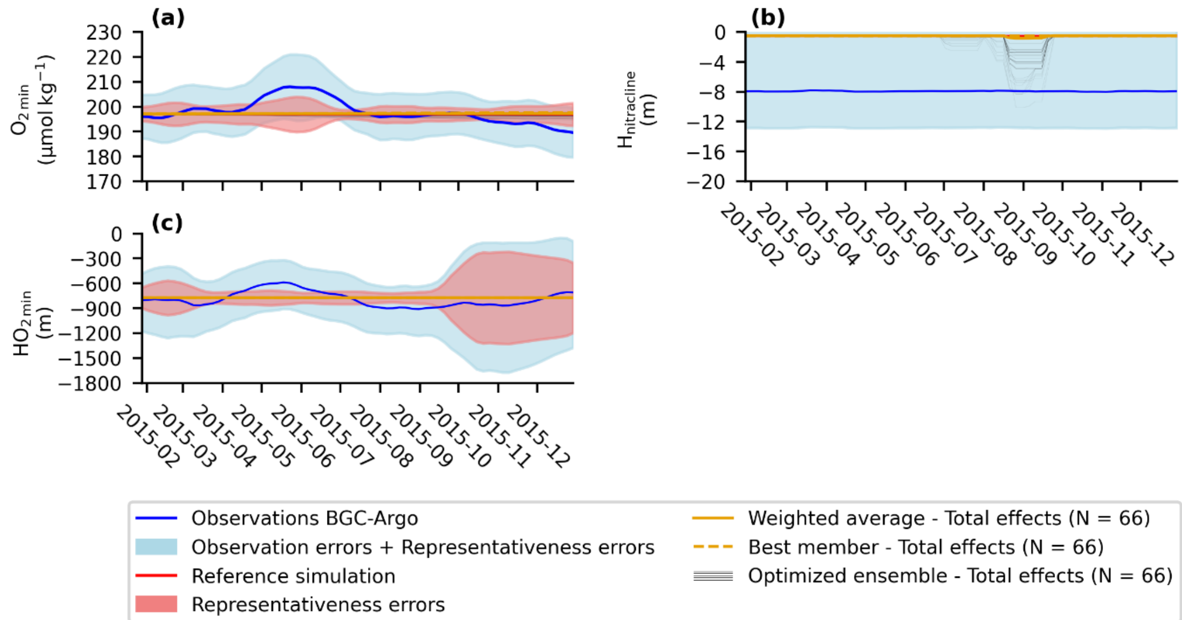

Figure S3. Seasonal cycle of assimilated emerging metrics. Panels show : (**a**) $O_{2\,min}$, (**b**) $H_{Nitracline}$, (**c**) $HO_{2\,min}$. Chl_{DCM} and H_{DCM} are not shown, as there were not enough DCM observations to reconstruct these metrics. The blue curve represents observations from BGC-Argo float #5904479, with the blue shading indicating the combined observations and representativity errors. The red curve corresponds to the reference simulation from PISCES-1D, with the red shading representing representativity errors. Orange line indicate the weighted means of the ensemble optimized using the Total effects parameters. The black curves represent the simulations of this ensemble. A six-point moving average was applied to all time series to smooth short-term fluctuations.

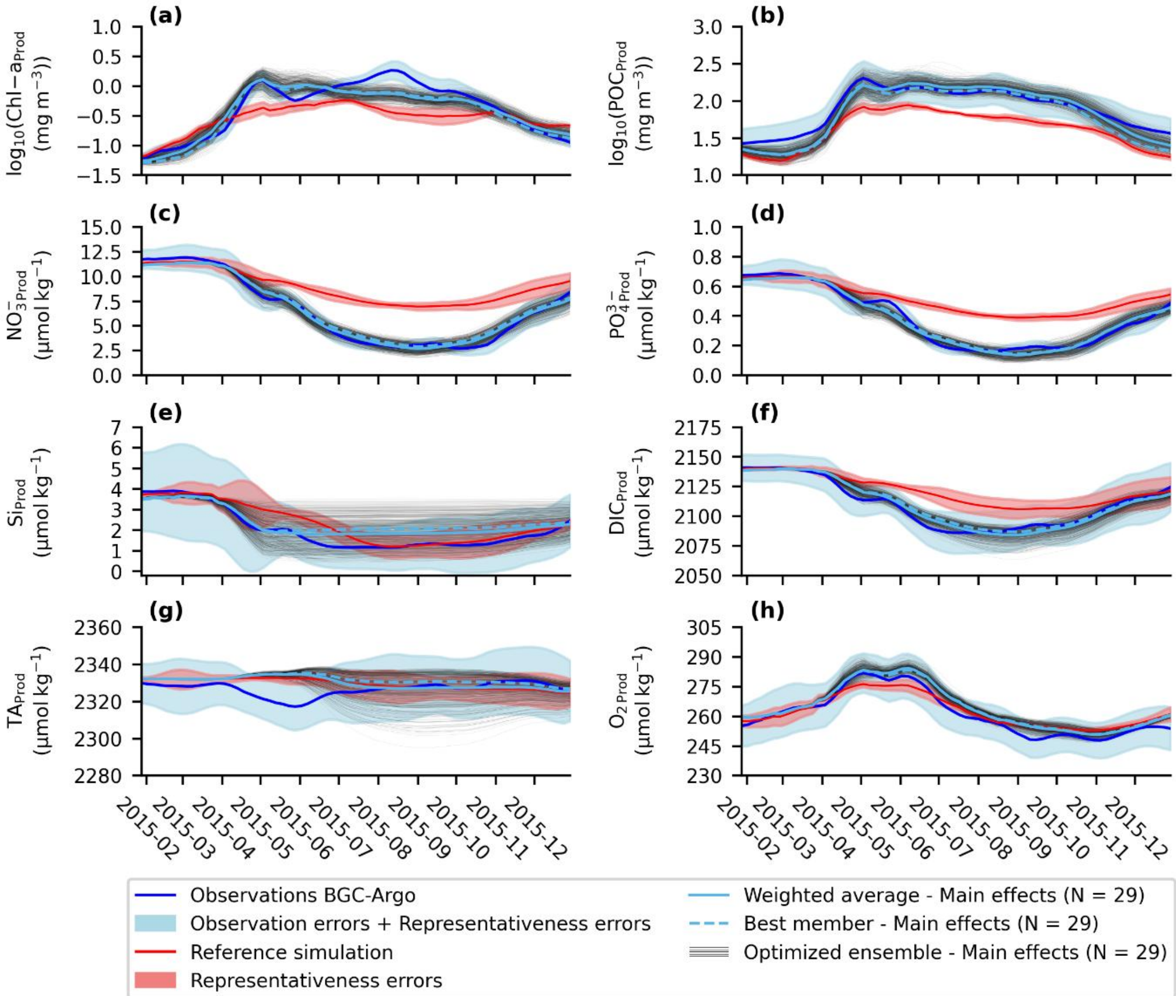

Figure S4. Seasonal cycle of assimilated metrics in the productive layer. Panels show : (**a**) $\log_{10}(\mathrm{Chl-a_{Prod}})$, (**b**) $\log_{10}(\mathrm{POC_{Prod}})$, (**c**) $\mathrm{NO_3^-}_{\mathrm{Prod}}$, (**d**) $PO_4^{3-}{}_{\mathrm{Prod}}$, (**e**) $\mathrm{Si_{Prod}}$, (**f**) $\mathrm{DIC_{Prod}}$, (**g**) $\mathrm{TA_{Prod}}$, and (**h**) $\mathrm{O_2}{}_{\mathrm{Prod}}$. The blue curve represents observations from BGC-Argo float #5904479, with the blue shading indicating the combined observations and representativity errors. The red curve corresponds to the reference simulation from PISCES-1D, with the red shading representing representativity errors. Blue line indicate the weighted means of the ensemble optimized using the Main effects parameters. The black curves represent the simulations of this ensemble. A six-point moving average was applied to all time series to smooth short-term fluctuations.

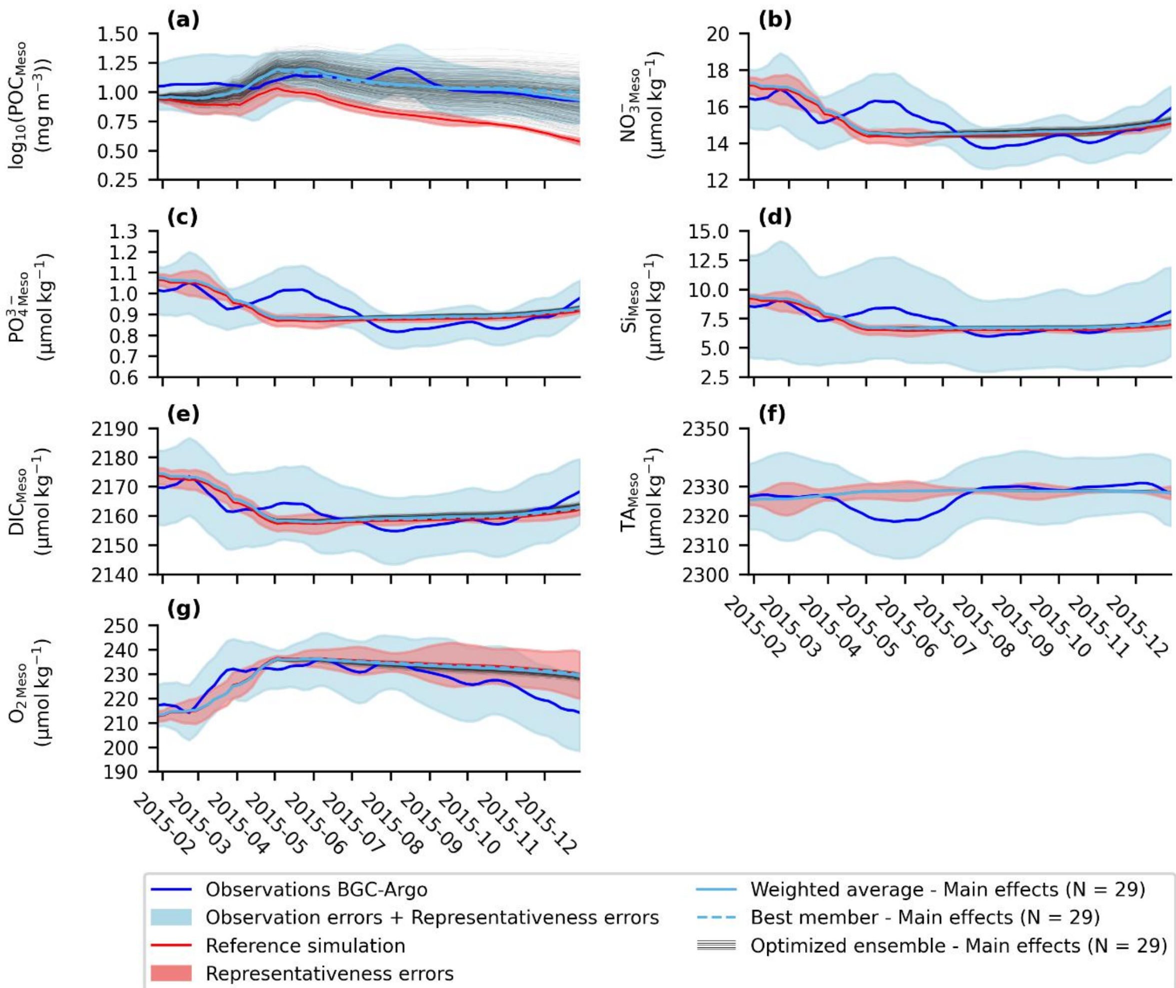

Figure S5. Seasonal cycle of assimilated metrics in the mesopelagic layer. Panels show : (**a**) $\log_{10}(\mathrm{POC_{Meso}})$ (**b**) $\mathrm{NO_3^-}_{\mathrm{Meso}}$, (**c**) $PO_4^{3-}{}_{\mathrm{Meso}}$, (**d**) $\mathrm{Si_{Meso}}$, (**e**) $\mathrm{DIC_{Meso}}$, (**f**) $\mathrm{TA_{Meso}}$, and (**g**) $\mathrm{O_2}_{\mathrm{Meso}}$. The blue curve represents observations from BGC-Argo float #5904479, with the blue shading indicating the combined observations and representativity errors. The red curve corresponds to the reference simulation from PISCES-1D, with the red shading representing representativity errors. Blue line indicate the weighted means of the ensemble optimized using the Main effects parameters. The black curves represent the simulations of this ensemble. A six-point moving average was applied to all time series to smooth short-term fluctuations.

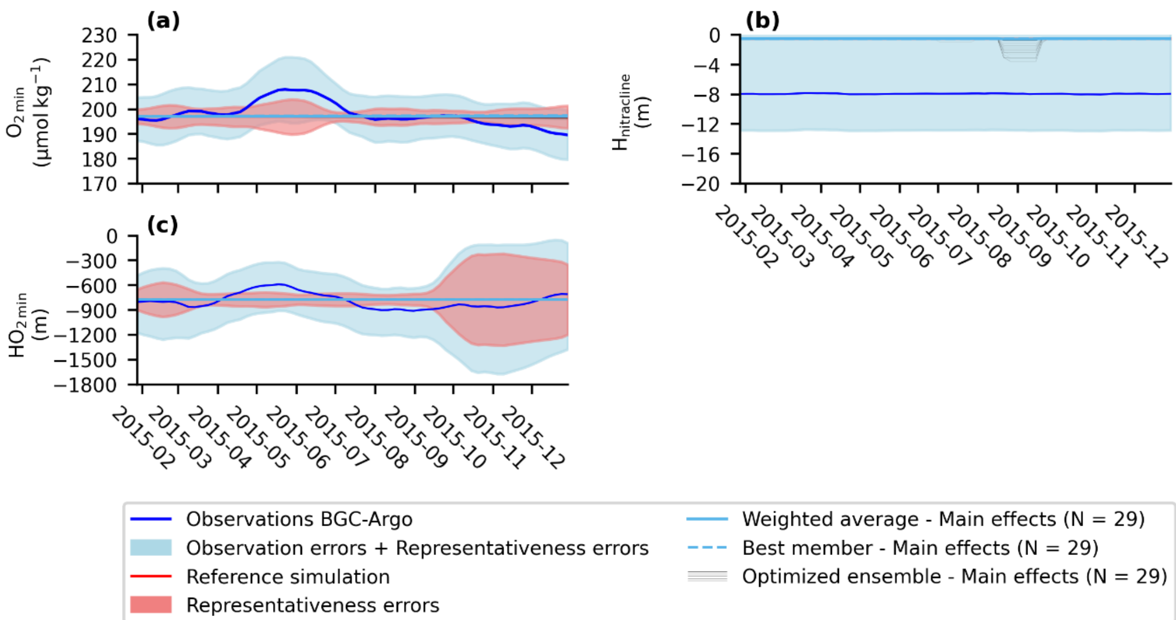

Figure S6. Seasonal cycle of assimilated emerging metrics. Panels show : (**a**) $O_{2_{min}}$, (**b**) $H_{Nitracline}$, (**c**) $HO_{2_{min}}$. Chl_{DCM} and H_{DCM} are not shown, as there were not enough DCM observations to reconstruct these metrics. The blue curve represents observations from BGC-Argo float #5904479, with the blue shading indicating the combined observations and representativity errors. The red curve corresponds to the reference simulation from PISCES-1D, with the red shading representing representativity errors. Blue line indicate the weighted means of the ensemble optimized using the Main effects parameters. The black curves represent the simulations of this ensemble. A six-point moving average was applied to all time series to smooth short-term fluctuations.

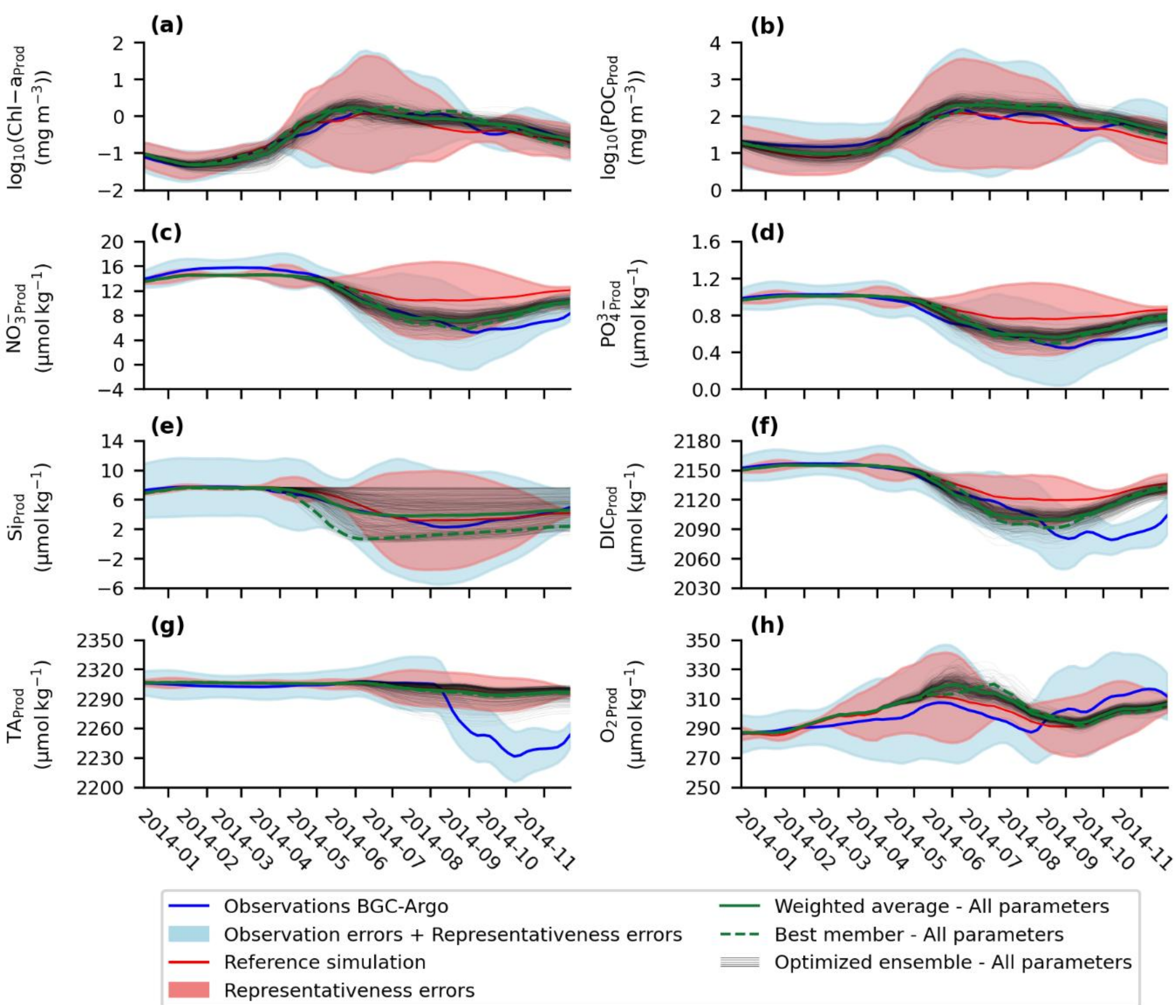
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
log10(Chl−aProd) (mg m−3))
log10(POCProd (mg m−3))
NO3−Prod (μmol kg−1)
PO43−Prod (μmol kg−1)
SiProd (μmol kg−1)
DICProd (μmol kg−1)
TAProd (μmol kg−1)
O2 Prod (μmol kg−1)
Observations BGC-Argo
Observation errors + Representativeness errors
Reference simulation
Representativeness errors
Weighted average - All parameters
Best member - All parameters
Optimized ensemble - All parameters

Figure S7. Seasonal cycle of assimilated metrics in the productive layer. Panels show : (**a**) $\log_{10}(\mathrm{Chl-a_{Prod}})$, (**b**) $\log_{10}(\mathrm{POC_{Prod}})$, (**c**) $\mathrm{NO_3^-}_{\mathrm{Prod}}$, (**d**) $PO_4^{3-}{}_{\mathrm{Prod}}$, (**e**) $\mathrm{Si_{Prod}}$, (**f**) $\mathrm{DIC_{Prod}}$, (**g**) $\mathrm{TA_{Prod}}$, and (**h**) $\mathrm{O_2}_{\mathrm{Prod}}$. The blue curve represents observations from BGC-Argo float #6901485, with the blue shading indicating the combined observations and representativity errors. The red curve corresponds to the reference simulation from PISCES-1D, with the red shading representing representativity errors. Green line indicate the weighted means of the ensemble optimized using all parameters of the PISCES model. The black curves represent the ensemble of selected members obtained by optimizing all parameters of the PISCES model. A six-point moving average was applied to all time series to smooth short-term fluctuations.

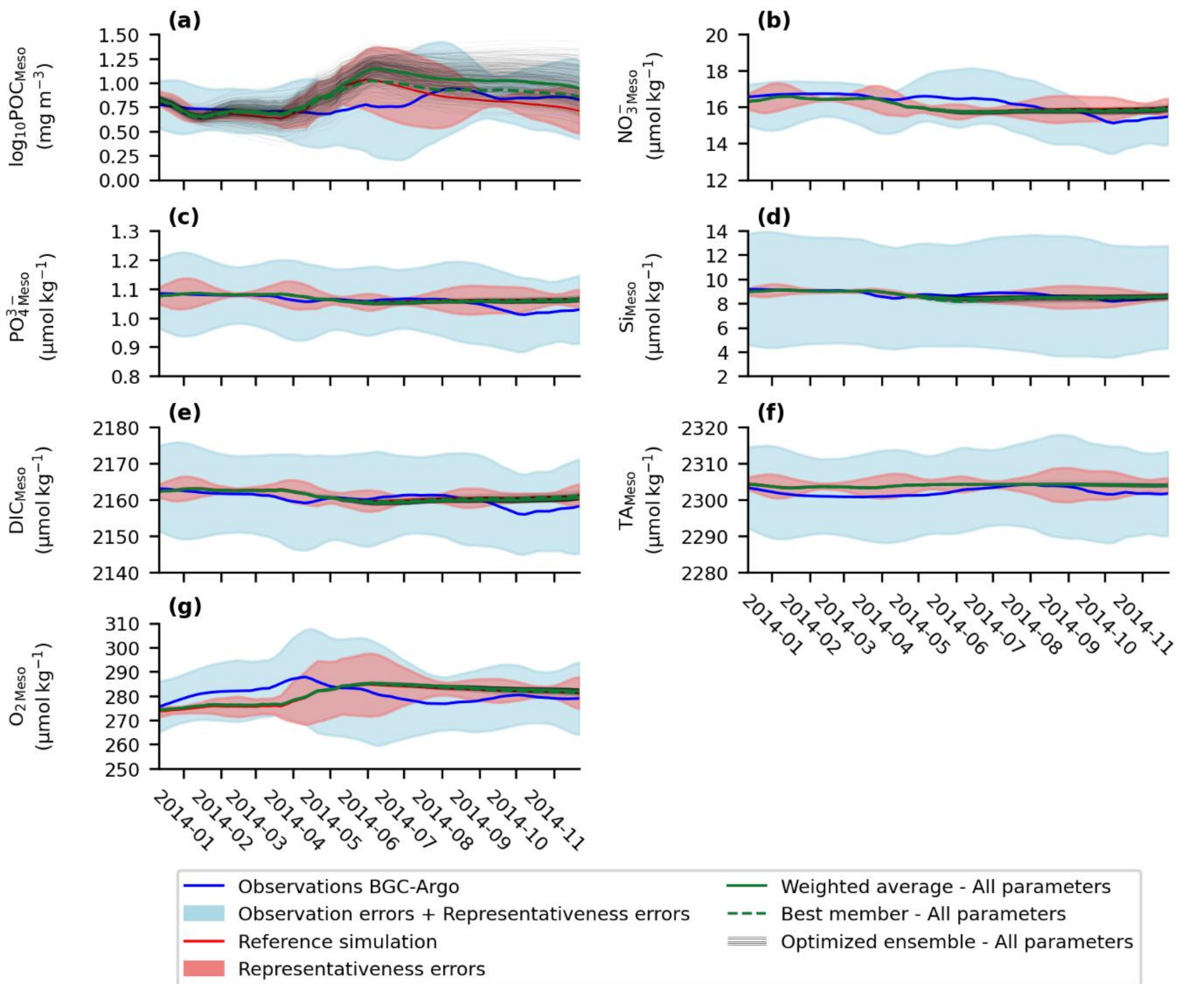

Figure S8. Seasonal cycle of assimilated metrics in the mesopelagic layer. Panels show : (**a**) $\log_{10}(\mathrm{POC_{Meso}})$ (**b**) $\mathrm{NO_3^-}_{\mathrm{Meso}}$, (**c**) $PO_4^{3-}{}_{\mathrm{Meso}}$, (**d**) $\mathrm{Si_{Meso}}$, (**e**) $\mathrm{DIC_{Meso}}$, (**f**) $\mathrm{TA_{Meso}}$, and (**g**) $\mathrm{O_2}_{\mathrm{Meso}}$. The blue curve represents observations from BGC-Argo float #6901485, with the blue shading indicating the combined observations and representativity errors. The red curve corresponds to the reference simulation from PISCES-1D, with the red shading representing representativity errors. Green line indicate the weighted means of the ensemble optimized using all parameters of the PISCES model. The black curves represent the ensemble of selected members obtained by optimizing all parameters of the PISCES model. A six-point moving average was applied to all time series to smooth short-term fluctuations.

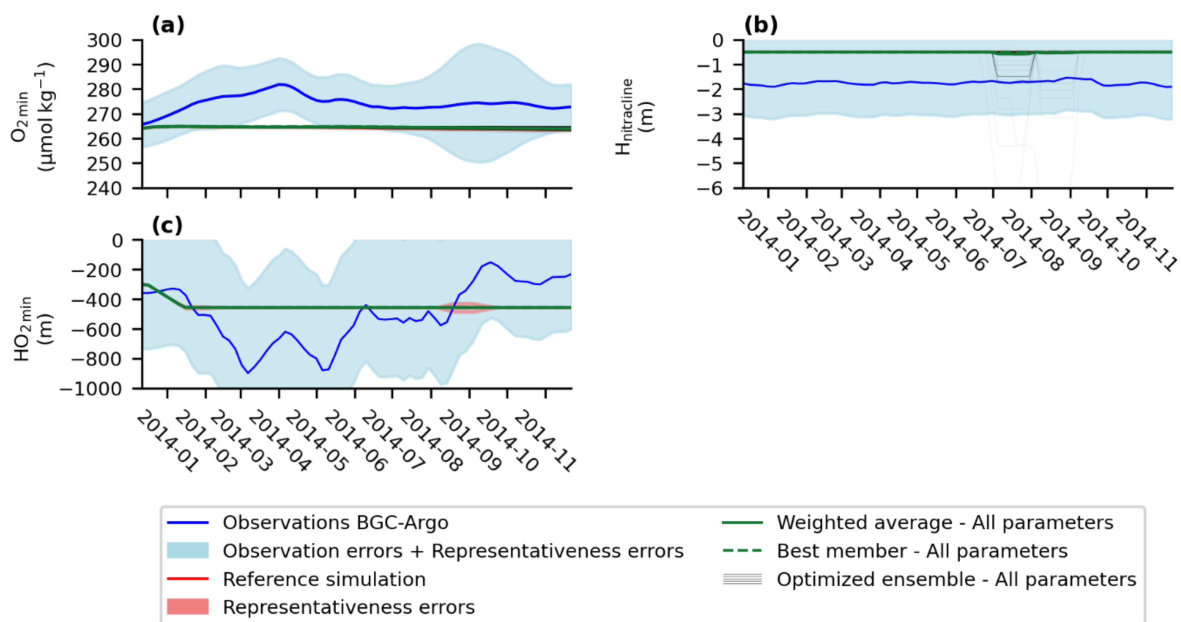

Figure S9. Seasonal cycle of assimilated emerging metrics. Panels show : (**a**) $\mathrm{O_2}_{\mathrm{min}}$, (**b**) $\mathrm{H_{Nitracline}}$, (**c**) $\mathrm{HO_2}_{\mathrm{min}}$. $\mathrm{Chl_{DCM}}$ and $\mathrm{H_{DCM}}$ are not shown, as there were not enough DCM observations to reconstruct these metrics. The blue curve represents observations from BGC-Argo float #6901485, with the blue shading indicating the combined observations and representativity errors. The red curve corresponds to the reference simulation from PISCES-1D, with the red shading representing representativity errors. Green line indicate the weighted means of the ensemble optimized using all parameters of the PISCES model. The black curves represent the ensemble of selected members obtained by optimizing all parameters of the PISCES model. A six-point moving average was applied to all time series to smooth short-term fluctuations.

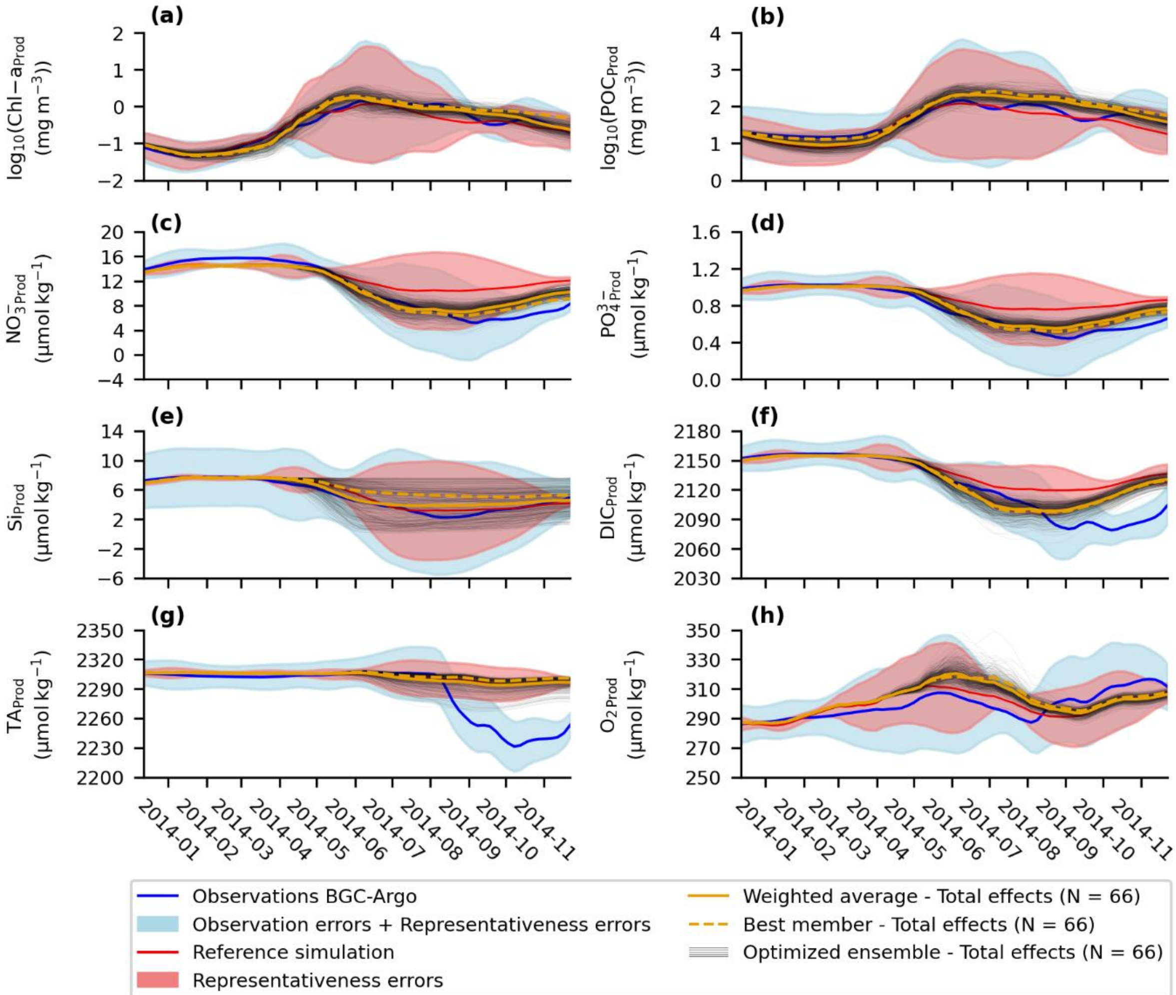

Figure S10. Seasonal cycle of assimilated metrics in the productive layer. Panels show : (**a**) $\log_{10}(\mathrm{Chl-a_{Prod}})$, (**b**) $\log_{10}(\mathrm{POC_{Prod}})$, (**c**) $\mathrm{NO_3^-}_{\mathrm{Prod}}$, (**d**) $PO_4^{3-}{}_{\mathrm{Prod}}$, (**e**) $\mathrm{Si_{Prod}}$, (**f**) $\mathrm{DIC_{Prod}}$, (**g**) $\mathrm{TA_{Prod}}$, and (**h**) $\mathrm{O_2}_{\mathrm{Prod}}$. The blue curve represents observations from BGC-Argo float #6901485, with the blue shading indicating the combined observations and representativity errors. The red curve corresponds to the reference simulation from PISCES-1D, with the red shading representing representativity errors. Orange line indicate the weighted means of the ensemble optimized using the Total effects parameters. The black curves represent the simulations of this ensemble. A six-point moving average was applied to all time series to smooth short-term fluctuations.

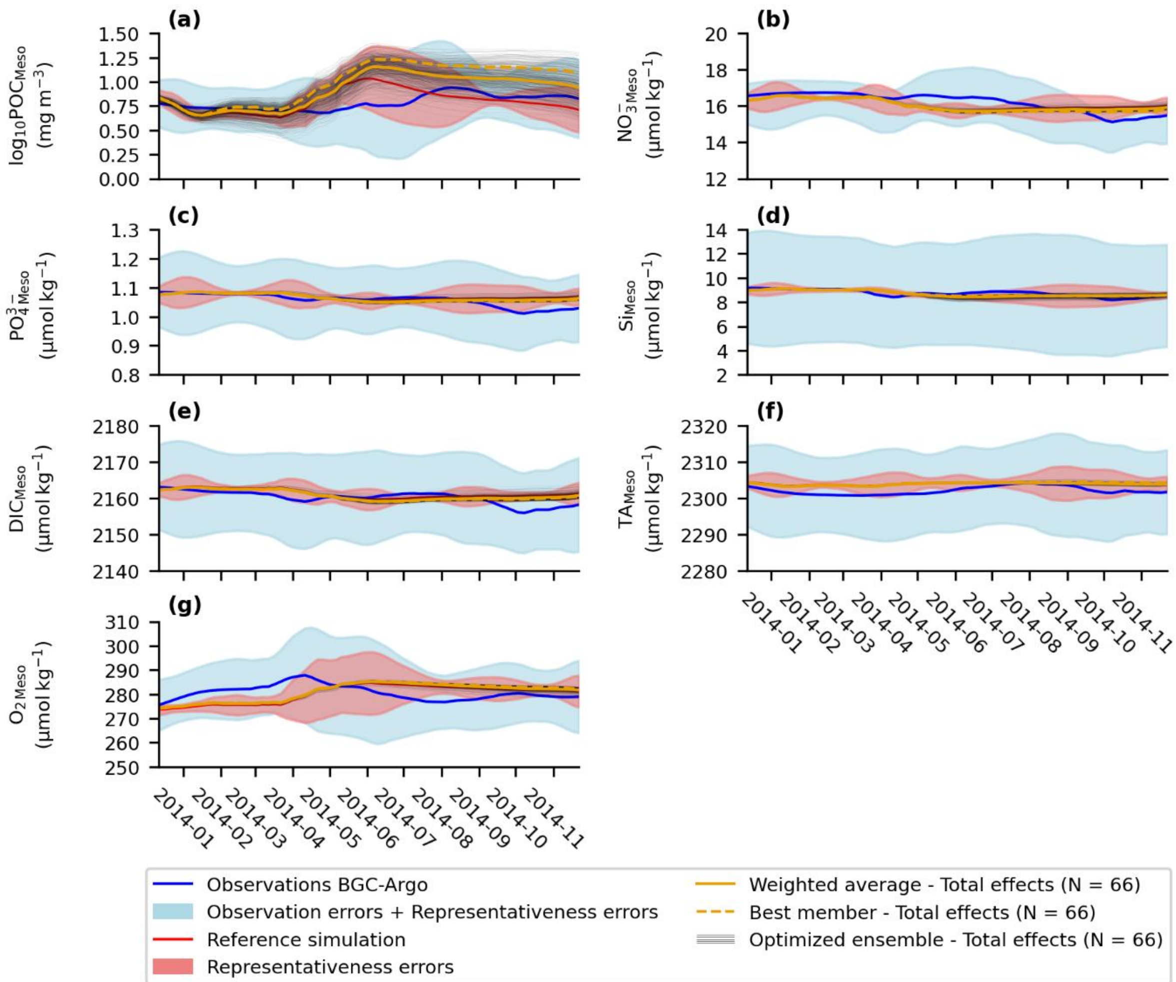

Figure S11. Seasonal cycle of assimilated metrics in the mesopelagic layer. Panels show : (**a**) $\log_{10}(\mathrm{POC_{Meso}})$ (**b**) $\mathrm{NO_3^-}_{\mathrm{Meso}}$, (**c**) $PO_4^{3-}{}_{\mathrm{Meso}}$, (**d**) $\mathrm{Si_{Meso}}$, (**e**) $\mathrm{DIC_{Meso}}$, (**f**) $\mathrm{TA_{Meso}}$, and (**g**) $\mathrm{O_2}_{\mathrm{Meso}}$. The blue curve represents observations from BGC-Argo float #6901485, with the blue shading indicating the combined observations and representativity errors. The red curve corresponds to the reference simulation from PISCES-1D, with the red shading representing representativity errors. Orange line indicate the weighted means of the ensemble optimized using the Total effects parameters. The black curves represent the simulations of this ensemble. A six-point moving average was applied to all time series to smooth short-term fluctuations.

Figure S12. Seasonal cycle of assimilated emerging metrics. Panels show : (**a**) $O_{2_{min}}$, (**b**) $H_{Nitracline}$, (**c**) $HO_{2_{min}}$. Chl_{DCM} and H_{DCM} are not shown, as there were not enough DCM observations to reconstruct these metrics. The blue curve represents observations from BGC-Argo float #6901485, with the blue shading indicating the combined observations and representativity errors. The red curve corresponds to the reference simulation from PISCES-1D, with the red shading representing representativity errors. Orange line indicate the weighted means of the ensemble optimized using the Total effects parameters. The black curves represent the simulations of this ensemble. A six-point moving average was applied to all time series to smooth short-term fluctuations.

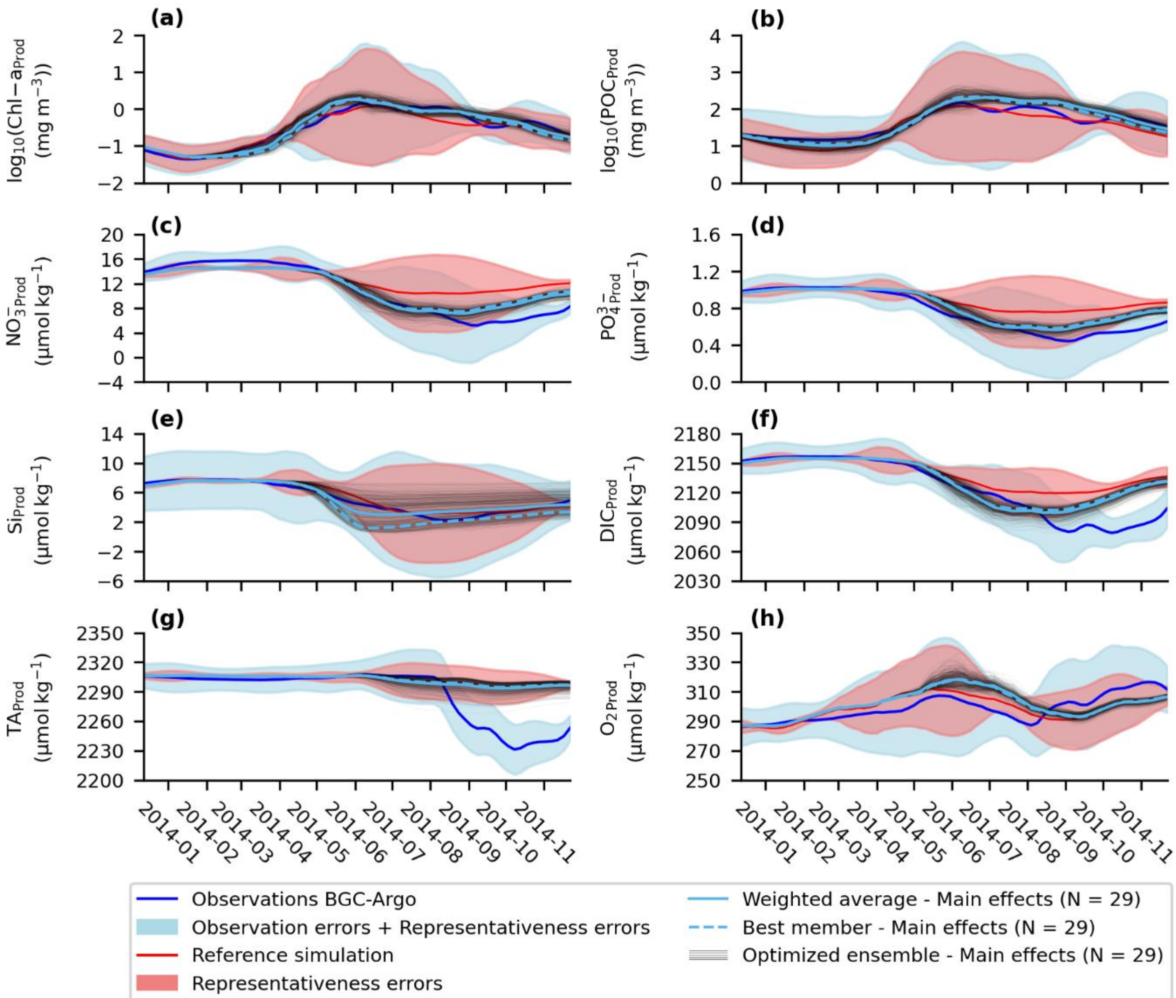

Figure S13. Seasonal cycle of assimilated metrics in the productive layer. Panels show : (**a**) $\log_{10}(\mathrm{Chl-a_{Prod}})$, (**b**) $\log_{10}(\mathrm{POC_{Prod}})$, (**c**) $\mathrm{NO_3^-}_{\mathrm{Prod}}$, (**d**) $PO_4^{3-}{}_{\mathrm{Prod}}$, (**e**) $\mathrm{Si_{Prod}}$, (**f**) $\mathrm{DIC_{Prod}}$, (**g**) $\mathrm{TA_{Prod}}$, and (**h**) $\mathrm{O_2}_{\mathrm{Prod}}$. The blue curve represents observations from BGC-Argo float #6901485, with the blue shading indicating the combined observations and representativity errors. The red curve corresponds to the reference simulation from PISCES-1D, with the red shading representing representativity errors. Blue line indicate the weighted means of the ensemble optimized using the Main effects parameters. The black curves represent the simulations of this ensemble. A six-point moving average was applied to all time series to smooth short-term fluctuations.

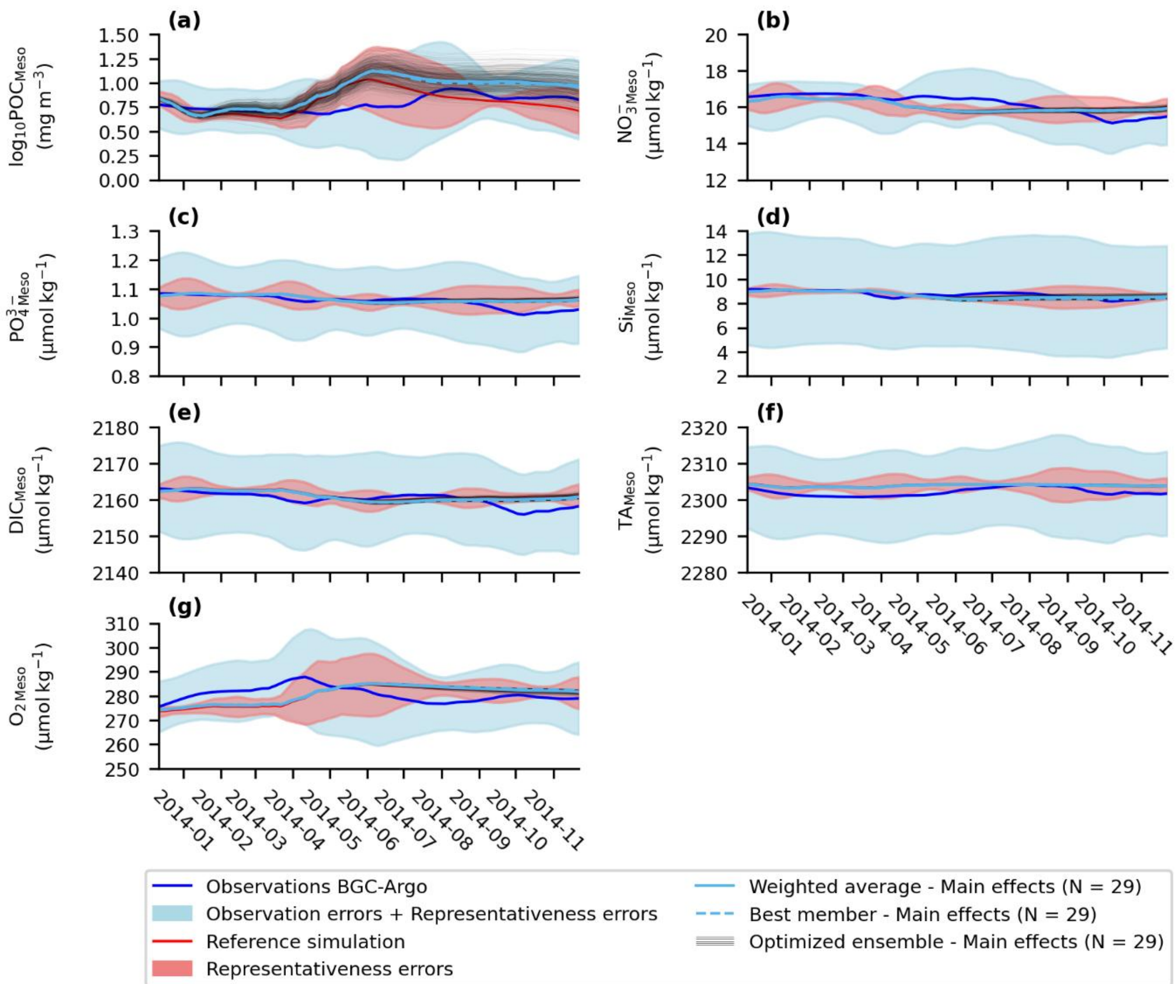

Figure S14. Seasonal cycle of assimilated metrics in the mesopelagic layer. Panels show : (**a**) $\log_{10}(\mathrm{POC_{Meso}})$ (**b**) $\mathrm{NO_3^-}_{\mathrm{Meso}}$, (**c**) $PO_4^{3-}{}_{\mathrm{Meso}}$, (**d**) $\mathrm{Si_{Meso}}$, (**e**) $\mathrm{DIC_{Meso}}$, (**f**) $\mathrm{TA_{Meso}}$, and (**g**) $\mathrm{O_2}_{\mathrm{Meso}}$. The blue curve represents observations from BGC-Argo float #6901485, with the blue shading indicating the combined observations and representativity errors. The red curve corresponds to the reference simulation from PISCES-1D, with the red shading representing representativity errors. Blue line indicate the weighted means of the ensemble optimized using the Main effects parameters. The black curves represent the simulations of this ensemble. A six-point moving average was applied to all time series to smooth short-term fluctuations.

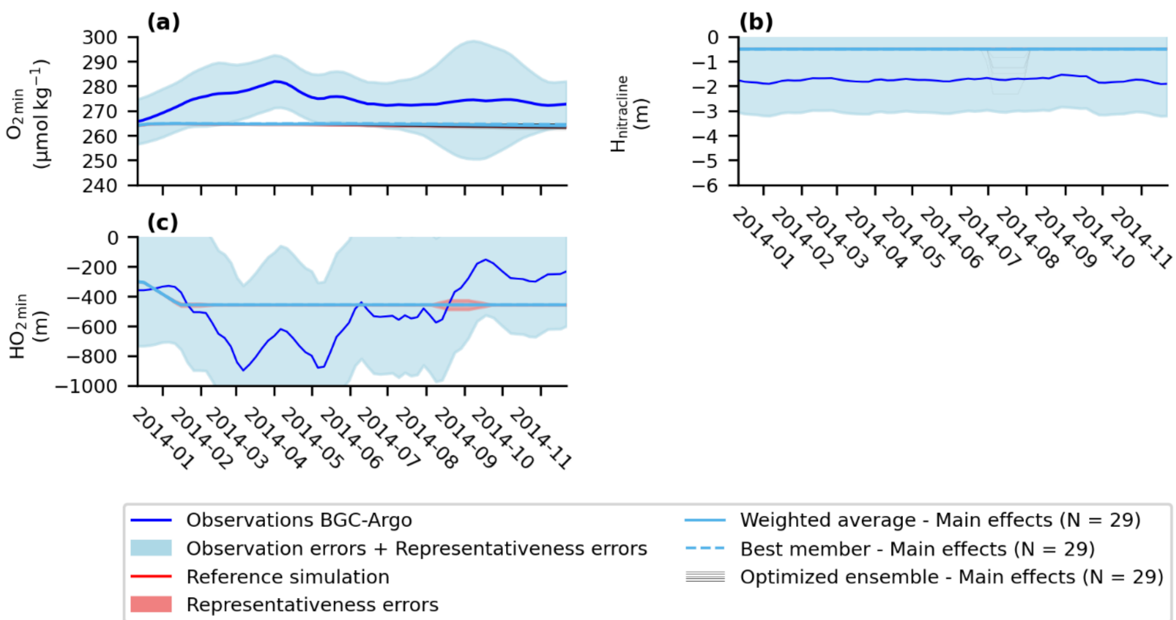

Figure S15. Seasonal cycle of assimilated emerging metrics. Panels show : (**a**) $O_{2_{min}}$, (**b**) $H_{Nitracline}$, (**c**) $HO_{2_{min}}$. Chl_{DCM} and H_{DCM} are not shown, as there were not enough DCM observations to reconstruct these metrics. The blue curve represents observations from BGC-Argo float #6901485, with the blue shading indicating the combined observations and representativity errors. The red curve corresponds to the reference simulation from PISCES-1D, with the red shading representing representativity errors. Blue line indicate the weighted means of the ensemble optimized using the Main effects parameters. The black curves represent the simulations of this ensemble. A six-point moving average was applied to all time series to smooth short-term fluctuations.

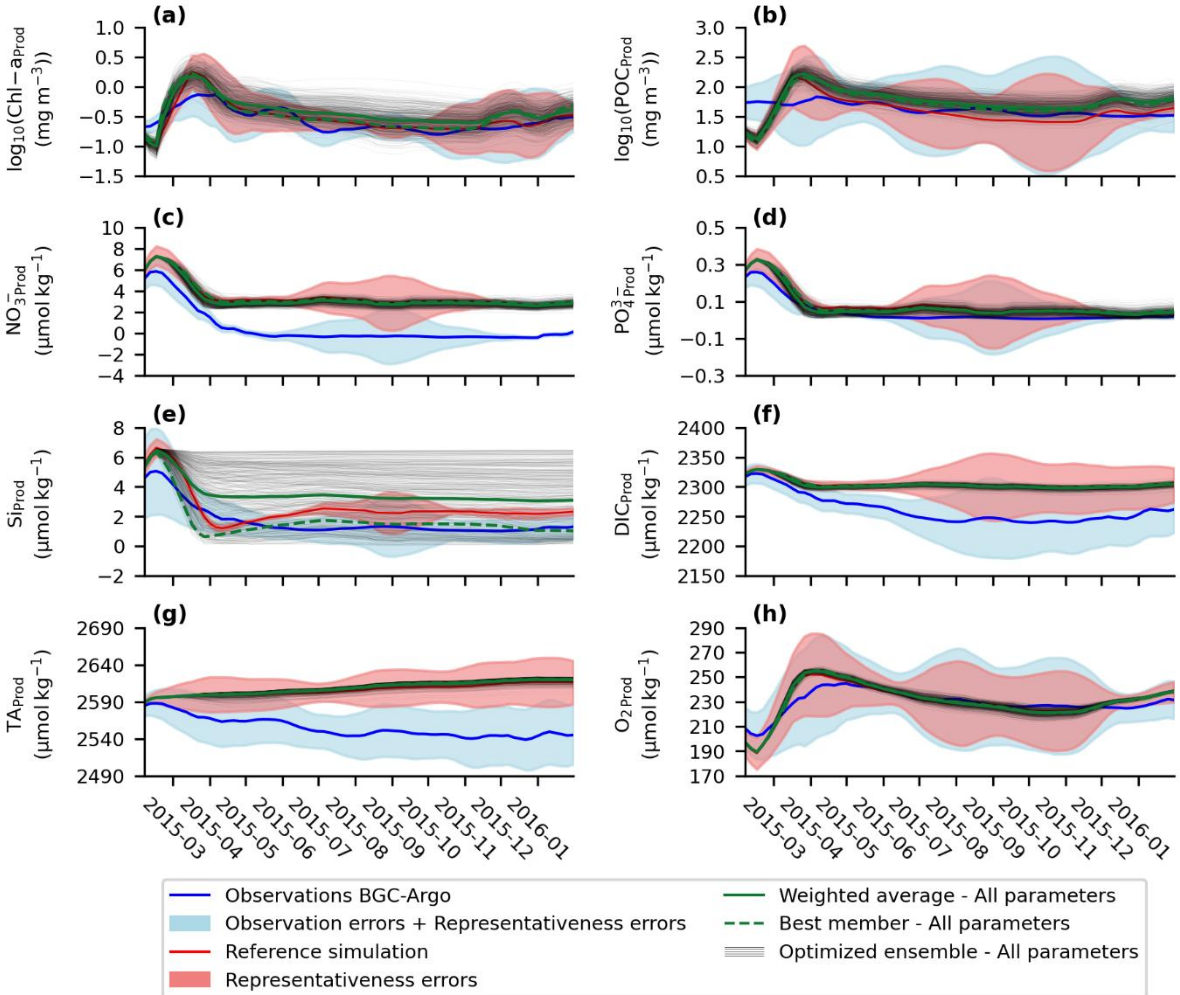

Figure S16. Seasonal cycle of assimilated metrics in the productive layer. Panels show : (**a**) $\log_{10}(\mathrm{Chl-a_{Prod}})$, (**b**) $\log_{10}(\mathrm{POC_{Prod}})$, (**c**) $\mathrm{NO_3^-}_{\mathrm{Prod}}$, (**d**) $PO_4^{3-}{}_{\mathrm{Prod}}$, (**e**) $\mathrm{Si_{Prod}}$, (**f**) $\mathrm{DIC_{Prod}}$, (**g**) $\mathrm{TA_{Prod}}$, and (**h**) $\mathrm{O_2}_{\mathrm{Prod}}$. The blue curve represents observations from BGC-Argo float #6901648, with the blue shading indicating the combined observations and representativity errors. The red curve corresponds to the reference simulation from PISCES-1D, with the red shading representing representativity errors. Green line indicate the weighted means of the ensemble optimized using all parameters of the PISCES model. The black curves represent the ensemble of selected members obtained by optimizing all parameters of the PISCES model. A six-point moving average was applied to all time series to smooth short-term fluctuations.

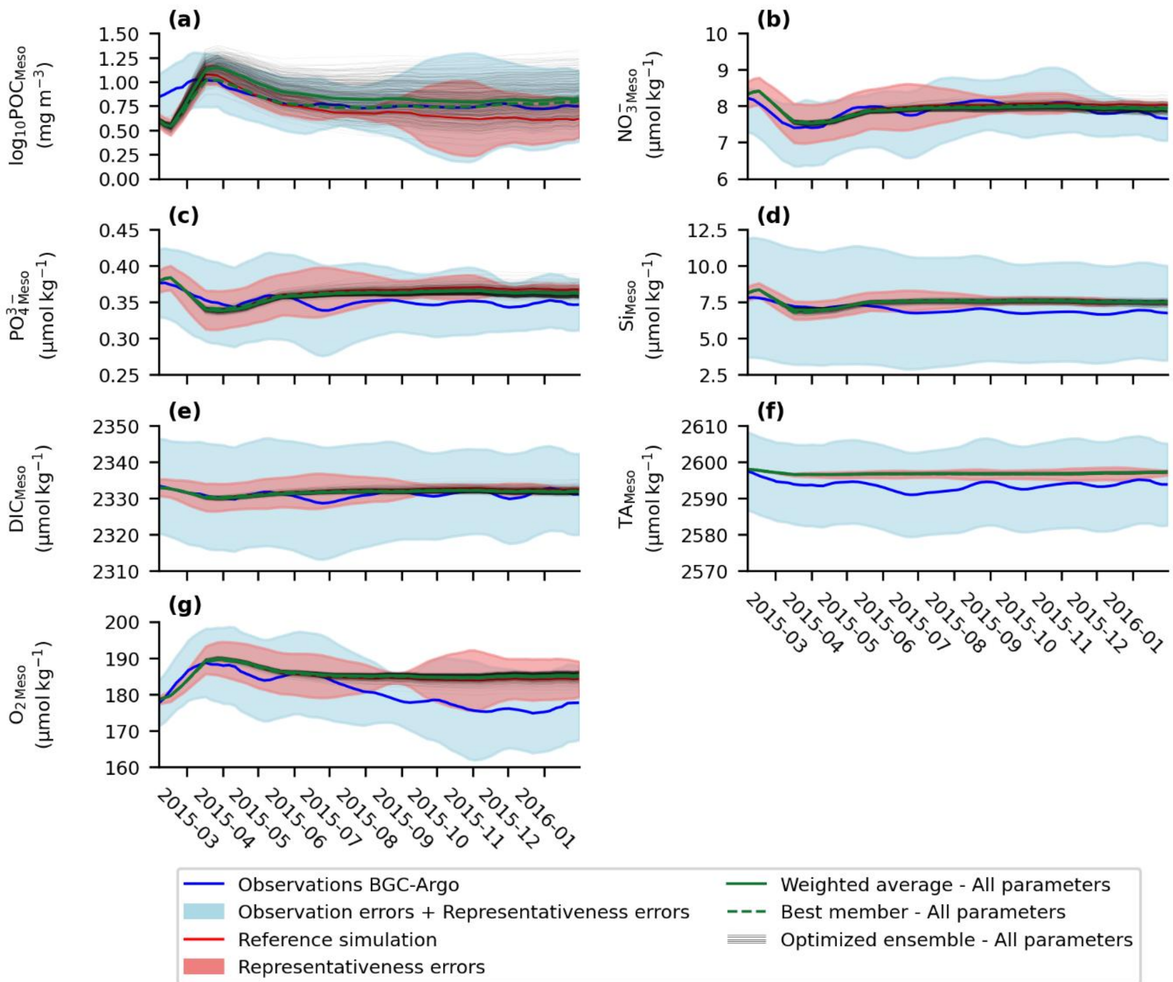

Figure S17. Seasonal cycle of assimilated metrics in the mesopelagic layer. Panels show : (**a**) $\log_{10}(\mathrm{POC_{Meso}})$ (**b**) $\mathrm{NO_3^-}_{\mathrm{Meso}}$, (**c**) $PO_4^{3-}{}_{\mathrm{Meso}}$, (**d**) $\mathrm{Si_{Meso}}$, (**e**) $\mathrm{DIC_{Meso}}$, (**f**) $\mathrm{TA_{Meso}}$, and (**g**) $\mathrm{O_2}_{\mathrm{Meso}}$. The blue curve represents observations from BGC-Argo float #6901648, with the blue shading indicating the combined observations and representativity errors. The red curve corresponds to the reference simulation from PISCES-1D, with the red shading representing representativity errors. Green line indicate the weighted means of the ensemble optimized using all parameters of the PISCES model. The black curves represent the ensemble of selected members obtained by optimizing all parameters of the PISCES model. A six-point moving average was applied to all time series to smooth short-term fluctuations.

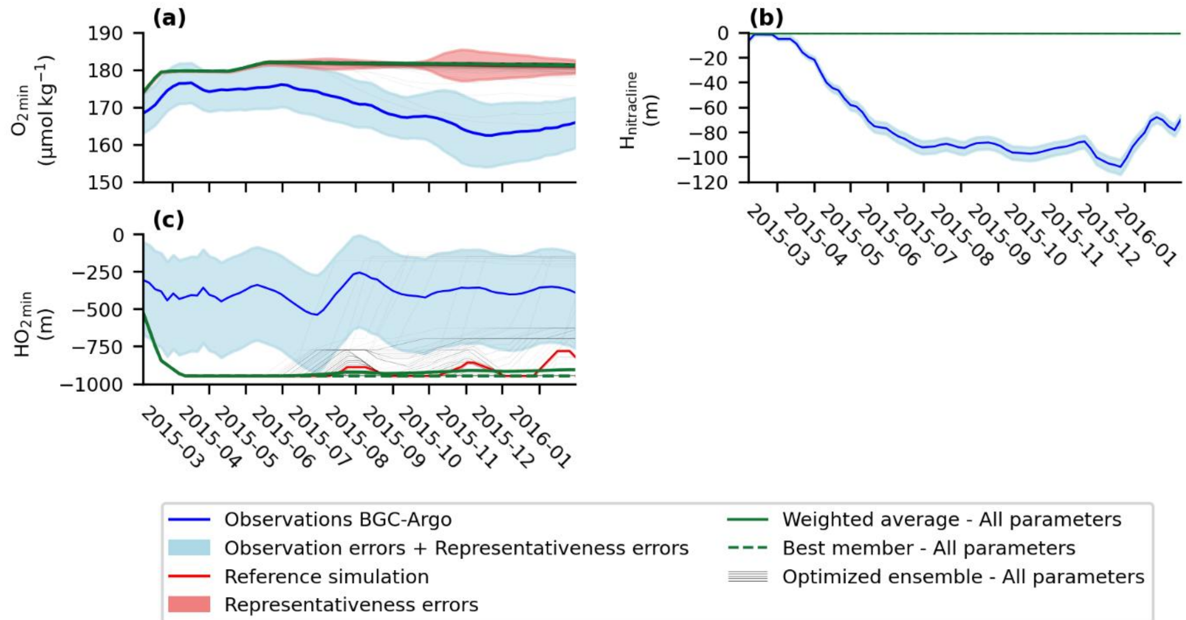

Figure S18. Seasonal cycle of assimilated emerging metrics. Panels show : (**a**) $O_{2_{\min}}$, (**b**) $H_{\mathrm{Nitracline}}$, (**c**) $HO_{2_{\min}}$. $\mathrm{Chl}_{\mathrm{DCM}}$ and H_{DCM} are not shown, as there were not enough DCM observations to reconstruct these metrics. The blue curve represents observations from BGC-Argo float #6901648, with the blue shading indicating the combined observations and representativity errors. The red curve corresponds to the reference simulation from PISCES-1D, with the red shading representing representativity errors. Green line indicate the weighted means of the ensemble optimized using all parameters of the PISCES model. The black curves represent the ensemble of selected members obtained by optimizing all parameters of the PISCES model. A six-point moving average was applied to all time series to smooth short-term fluctuations.

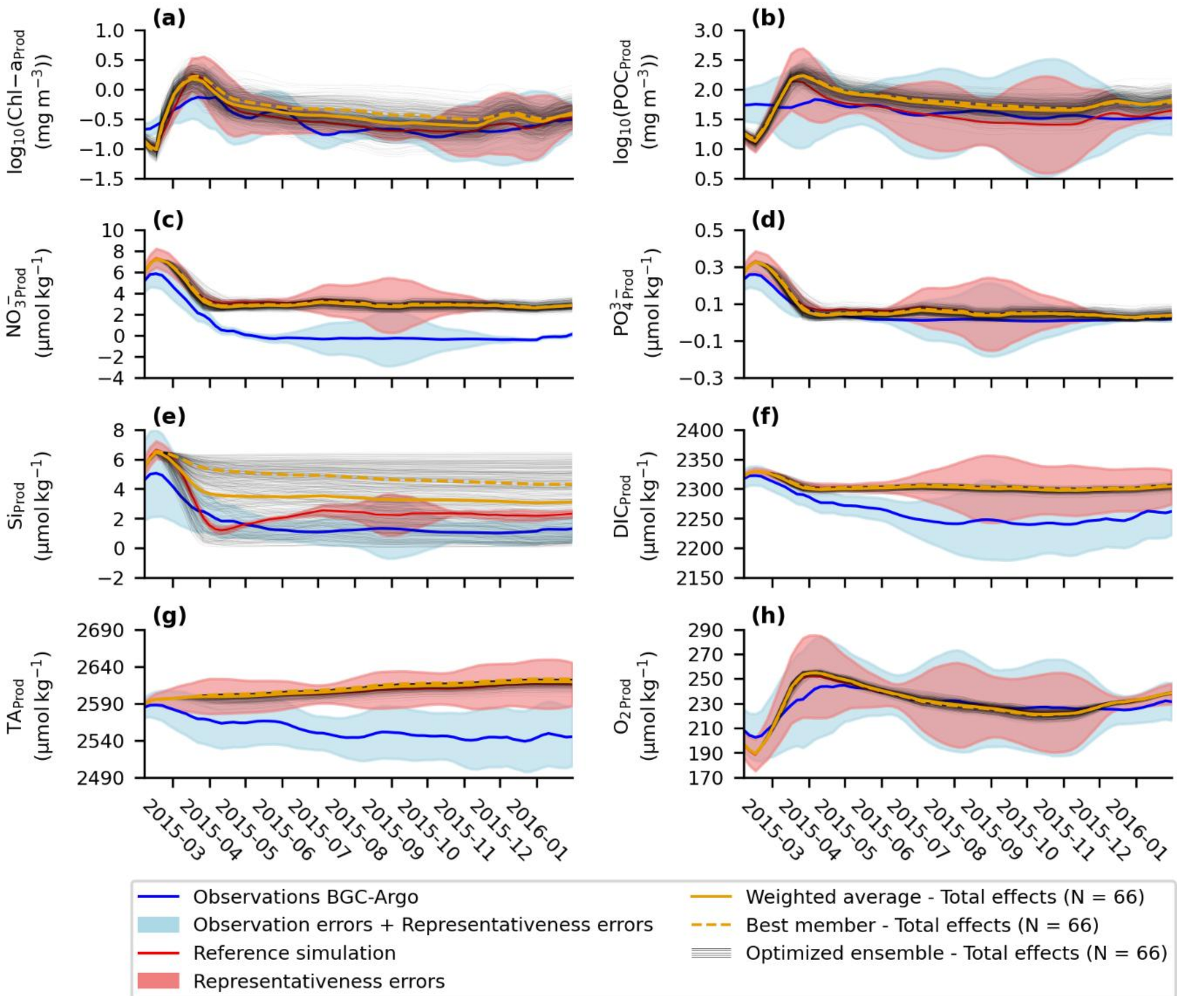

Figure S19. Seasonal cycle of assimilated metrics in the productive layer. Panels show : (**a**) $\log_{10}(\mathrm{Chl-a_{Prod}})$, (**b**) $\log_{10}(\mathrm{POC_{Prod}})$, (**c**) $\mathrm{NO_3^-}_{\mathrm{Prod}}$, (**d**) $PO_4^{3-}{}_{\mathrm{Prod}}$, (**e**) $\mathrm{Si_{Prod}}$, (**f**) $\mathrm{DIC_{Prod}}$, (**g**) $\mathrm{TA_{Prod}}$, and (**h**) $\mathrm{O_2}_{\mathrm{Prod}}$. The blue curve represents observations from BGC-Argo float #6901648, with the blue shading indicating the combined observations and representativity errors. The red curve corresponds to the reference simulation from PISCES-1D, with the red shading representing representativity errors. Orange line indicate the weighted means of the ensemble optimized using the Total effects parameters. The black curves represent the simulations of this ensemble. A six-point moving average was applied to all time series to smooth short-term fluctuations.

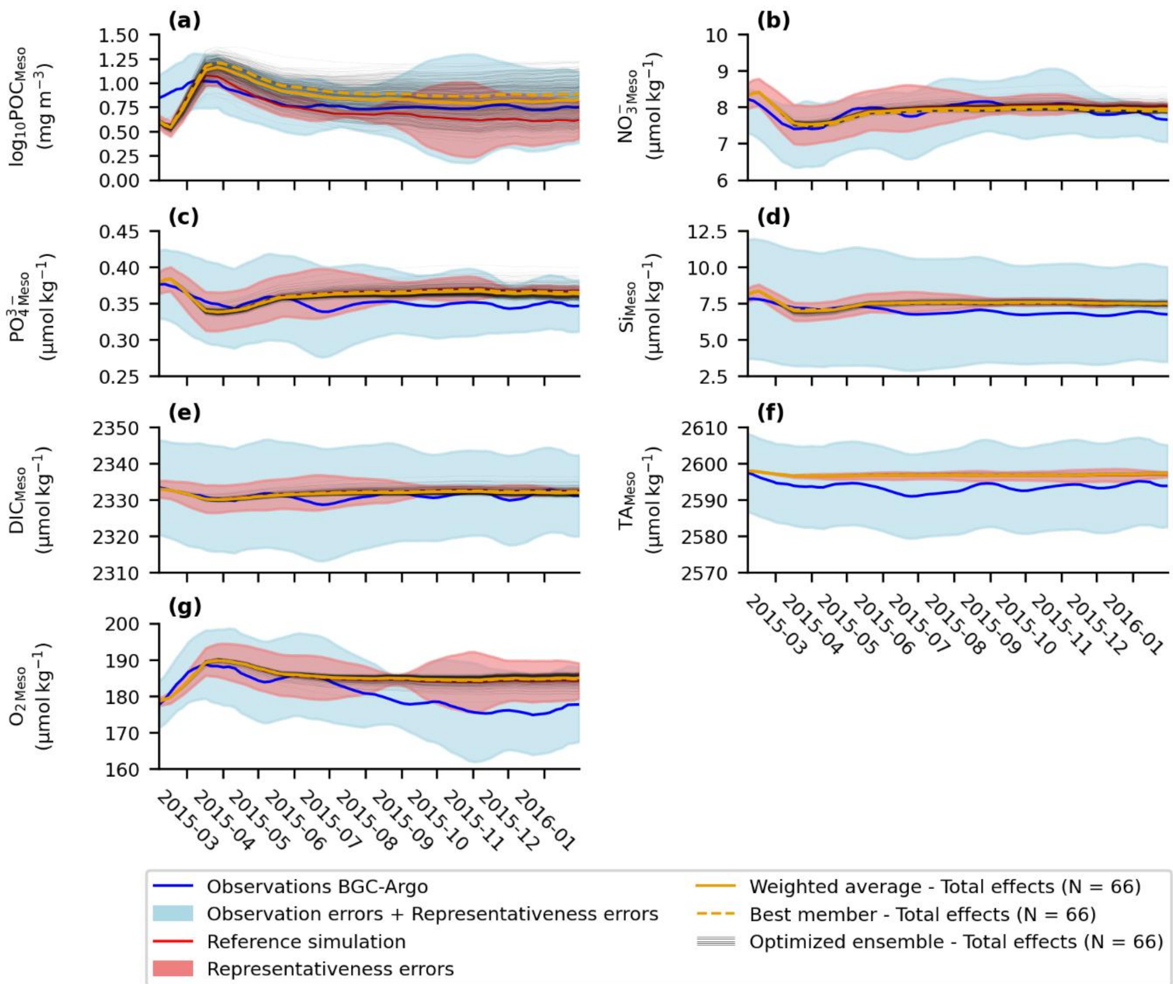

Figure S20. Seasonal cycle of assimilated metrics in the mesopelagic layer. Panels show : (**a**) $\log_{10}(\mathrm{POC_{Meso}})$ (**b**) $\mathrm{NO_3^-}_{\mathrm{Meso}}$, (**c**) $PO_4^{3-}{}_{\mathrm{Meso}}$, (**d**) $\mathrm{Si_{Meso}}$, (**e**) $\mathrm{DIC_{Meso}}$, (**f**) $\mathrm{TA_{Meso}}$, and (**g**) $\mathrm{O_2}_{\mathrm{Meso}}$. The blue curve represents observations from BGC-Argo float #6901648, with the blue shading indicating the combined observations and representativity errors. The red curve corresponds to the reference simulation from PISCES-1D, with the red shading representing representativity errors. Orange line indicate the weighted means of the ensemble optimized using the Total effects parameters. The black curves represent the simulations of this ensemble. A six-point moving average was applied to all time series to smooth short-term fluctuations.

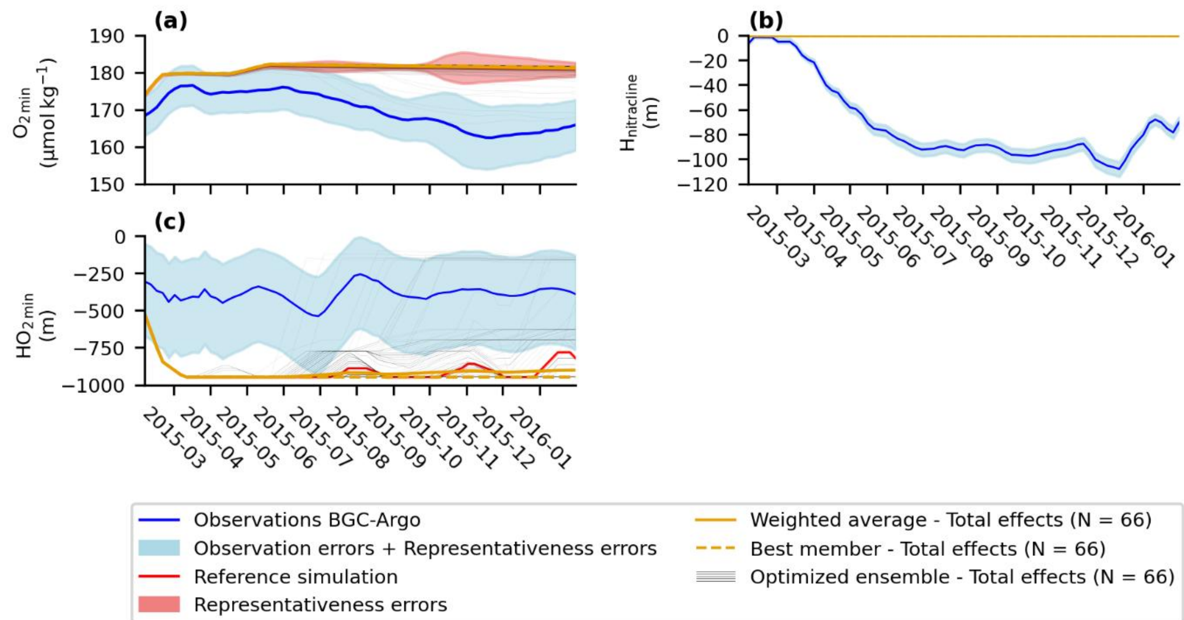

Figure S21. Seasonal cycle of assimilated emerging metrics. Panels show : (**a**) $O_{2_{min}}$, (**b**) $H_{Nitracline}$, (**c**) $HO_{2_{min}}$. Chl_{DCM} and H_{DCM} are not shown, as there were not enough DCM observations to reconstruct these metrics. The blue curve represents observations from BGC-Argo float #6901648, with the blue shading indicating the combined observations and representativity errors. The red curve corresponds to the reference simulation from PISCES-1D, with the red shading representing representativity errors. Orange line indicate the weighted means of the ensemble optimized using the Total effects parameters. The black curves represent the simulations of this ensemble. A six-point moving average was applied to all time series to smooth short-term fluctuations.

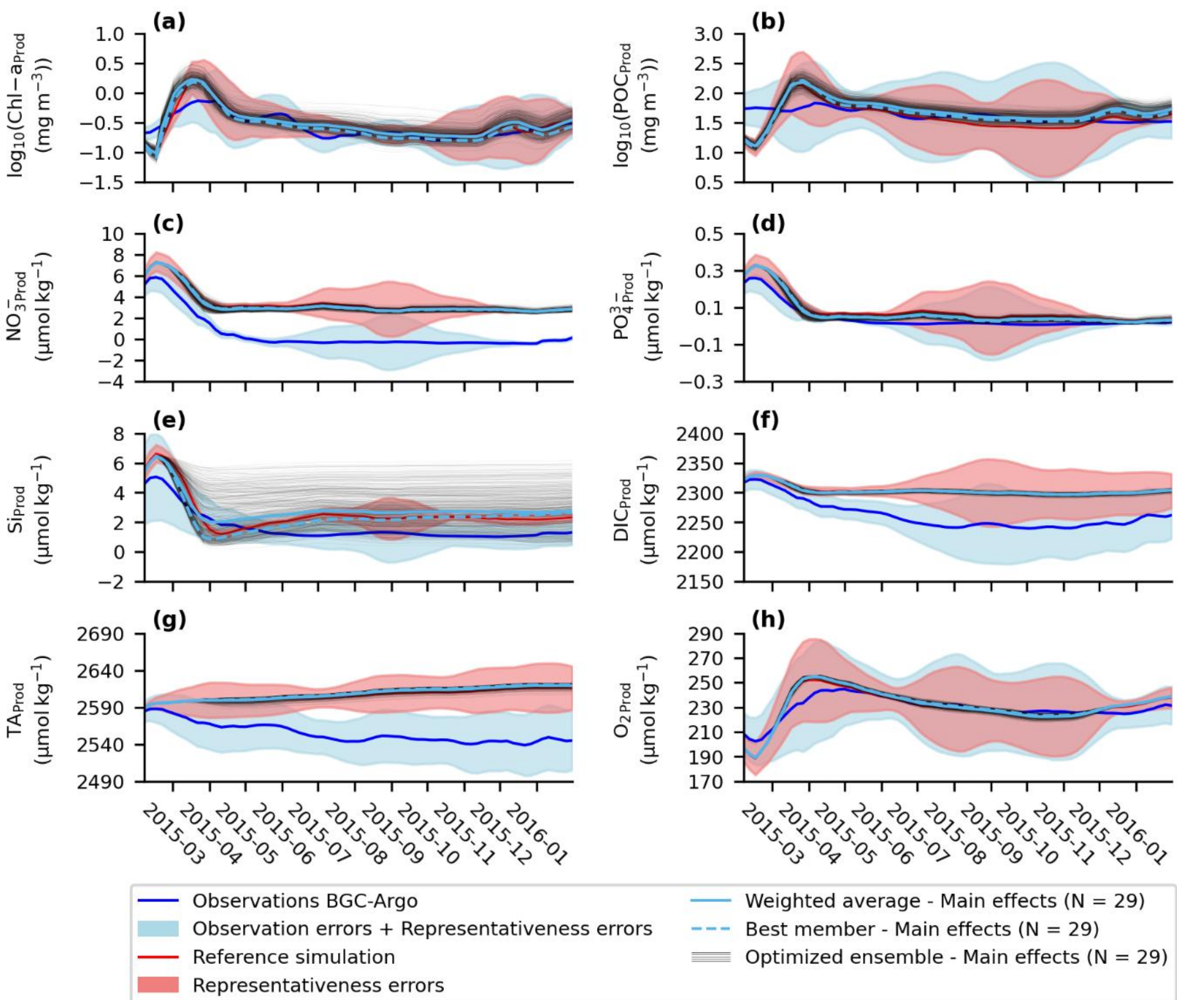

Figure S22. Seasonal cycle of assimilated metrics in the productive layer. Panels show : (**a**) $\log_{10}(\mathrm{Chl-a_{Prod}})$, (**b**) $\log_{10}(\mathrm{POC_{Prod}})$, (**c**) $\mathrm{NO_3^-}_{\mathrm{Prod}}$, (**d**) $PO_4^{3-}{}_{\mathrm{Prod}}$, (**e**) $\mathrm{Si_{Prod}}$, (**f**) $\mathrm{DIC_{Prod}}$, (**g**) $\mathrm{TA_{Prod}}$, and (**h**) $\mathrm{O_2}_{\mathrm{Prod}}$. The blue curve represents observations from BGC-Argo float #6901648, with the blue shading indicating the combined observations and representativity errors. The red curve corresponds to the reference simulation from PISCES-1D, with the red shading representing representativity errors. Blue line indicate the weighted means of the ensemble optimized using the Main effects parameters. The black curves represent the simulations of this ensemble. A six-point moving average was applied to all time series to smooth short-term fluctuations.

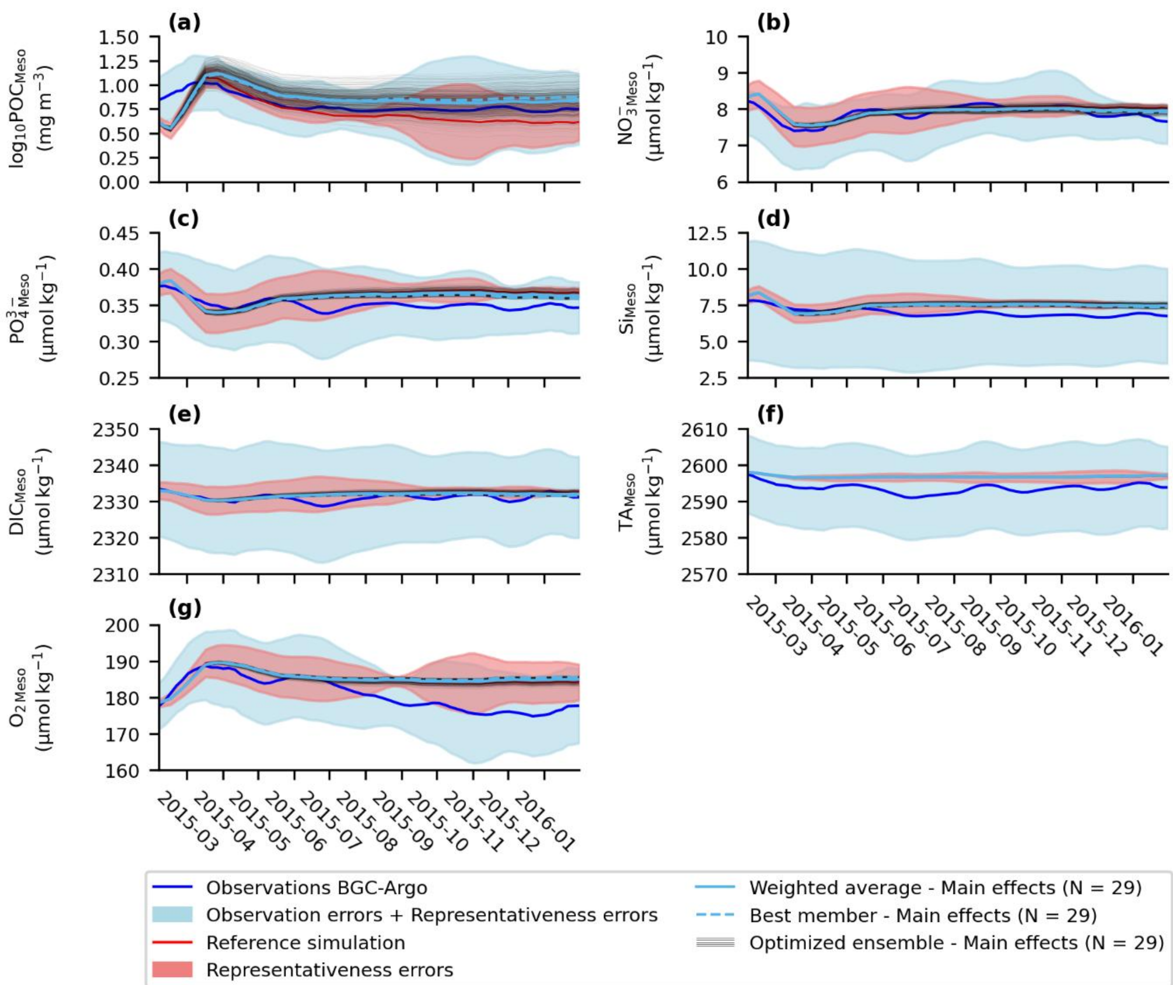

Figure S23. Seasonal cycle of assimilated metrics in the mesopelagic layer. Panels show : (**a**) $\log_{10}(\mathrm{POC_{Meso}})$ (**b**) $\mathrm{NO_3^-}_{\mathrm{Meso}}$, (**c**) $PO_4^{3-}{}_{\mathrm{Meso}}$, (**d**) $\mathrm{Si_{Meso}}$, (**e**) $\mathrm{DIC_{Meso}}$, (**f**) $\mathrm{TA_{Meso}}$, and (**g**) $\mathrm{O_2}_{\mathrm{Meso}}$. The blue curve represents observations from BGC-Argo float #6901648, with the blue shading indicating the combined observations and representativity errors. The red curve corresponds to the reference simulation from PISCES-1D, with the red shading representing representativity errors. Blue line indicate the weighted means of the ensemble optimized using the Main effects parameters. The black curves represent the simulations of this ensemble. A six-point moving average was applied to all time series to smooth short-term fluctuations.

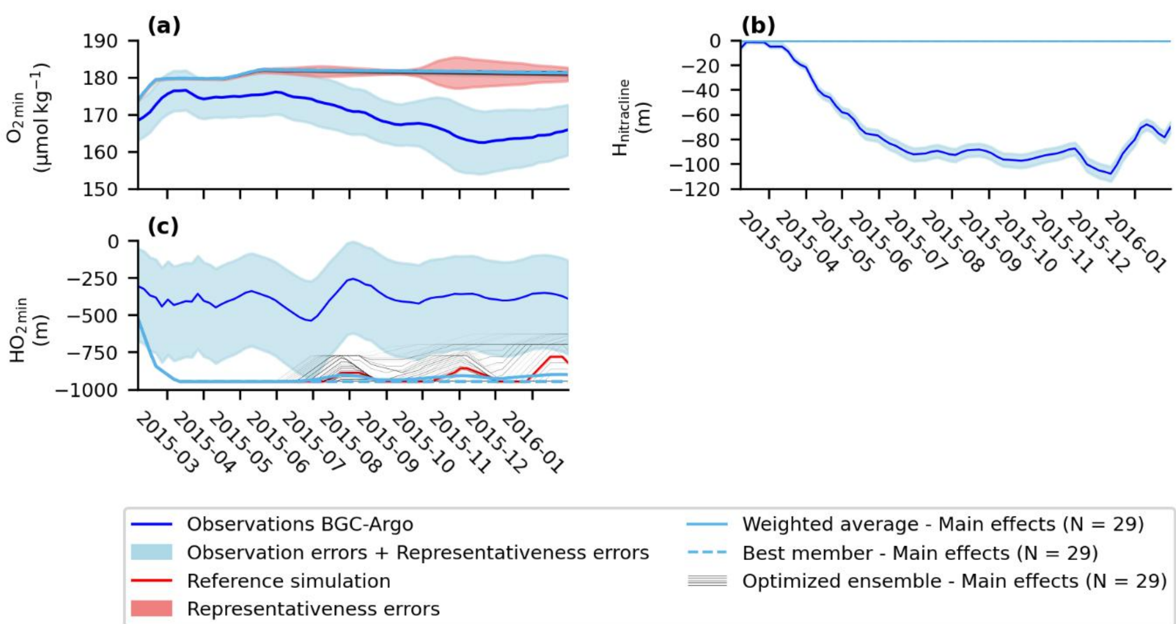

Figure S24. Seasonal cycle of assimilated emerging metrics. Panels show : (**a**) $O_{2_{\min}}$, (**b**) $H_{\mathrm{Nitracline}}$, (**c**) $HO_{2_{\min}}$. $\mathrm{Chl}_{\mathrm{DCM}}$ and H_{DCM} are not shown, as there were not enough DCM observations to reconstruct these metrics. The blue curve represents observations from BGC-Argo float #6901648, with the blue shading indicating the combined observations and representativity errors. The red curve corresponds to the reference simulation from PISCES-1D, with the red shading representing representativity errors. Blue line indicate the weighted means of the ensemble optimized using the Main effects parameters. The black curves represent the simulations of this ensemble. A six-point moving average was applied to all time series to smooth short-term fluctuations.