# Towards better biogeochemical ocean models

## Assessment of global biogeochemical ocean models via multi-objective optimisation

*V. Sauerland, I. Kriest, C. von Hallern, A. Srivastav, Kiel University, GEOMAR Helmholtz Centre for Ocean Research Kiel* model assessment is possible by calculating a couple of well-distributed compromise parameterisations at once. For example, if parameter values show a

## In Short

- Calibration of biogeochemical ocean models facilitates model intercomparison, but which parameters are optimal depends on the objective
- It is difficult to pose proper bounds on the modeldata misfit w.r.t. one objective while minimising the misfit related to another target or to properly join different misfits into one objective function
- Multi-objective optimisation eases the task to find good compromise solutions
- We integrate and apply multi-objective optimisation into a framework using an efficient tool for global simulation of oceanic biogeochemical processes

Biogeochemical (BGC) models mirror important processes of the global marine ecosystem. They have a similar influence on model predictions [1] as ocean circulation [2] but depend on roughly constrained parameters. There is a plethora of biogeochemical models of different complexity, each of which appears to be particularly skilled to meet certain (types of) observations. However, due to the parametric uncertainties, no BGC model is superior for all research questions.

Therefore, the parameters of any BGC model must be tuned until model predictions meet corresponding observations. This tuning procedure is also known as model calibration. Since ocean circulation and BGC processes influence each other, reliable model calibration requires to couple BGC components with global ocean circulation models. Up-to several millenia of simulated ocean evolution is necessary in order to obtain steady annual states. Thus, reaching (nearly) optimal model-data misfit values by calibration is a computational demanding task.

**Optimisation Framework** Automated parameter optimisation facilitates model intercomparison. However, good parameters with respect to one type of observations can be bad for another type of observations. Single-objective optimisation converges to a single parameter set, only.

Facing diverging model calibration results with respect to different target observations, enhanced

model assessment is possible by calculating a couple of well-distributed compromise parametrisations at once. For example, if parameter values show a clean monotonic tendency while traversing the final collection of compromise parameter sets from one objective's best solution to the other objective's one (cf. color gradients in Figure 1), we gain more evidence about dependencies between processes and the parameters than from two single-objective calibration results. Multi-objective optimisation methods are designated for this purpose. They can provides a collection of parameter sets such that the user obtains a nearly optimal parameter set with respect to each target but also has the option to choose a best compromise parametrisation in his/her opinion.

The field of multi-objective (MO) optimisation is treated with EAs [3] and EDAs [4], including CMA-ES [5].

Concerning model simulation, we keep using the "Transport Matrix Method" (TMM) for easy and generic coupling between different biogeochemical models and circulation [6]

Achievements and further goals Our singleobjective and bi-objective CMAES configurations are yet applied in a broader research context, regarding model skills with respect to predictions about oxygen minimum zones and nitrous oxide, as experimentally addressed within HLRN project shk00033. In a publication about this issue [7], we applied the biobjective CMAES implementation developed within this project. It is available as a GitHub repository, a permanent version of which is [8].

In the third year of this project we aim to continue the enhancements of our calibration framework with respect to more then two objectives, as compromise parameter sets regarding three or four types of target observations should further help to reveal reasons behind parametric uncertainties.

In rare cases, we observed the "good distribution" criterion for the collection of compromise parameter sets violated. Therefore, we will also integrate and test an avoiding strategy for this behavior, as intended by the concept of  $\epsilon$ -dominance [9].

In addition to addressing different data types [10] (e.g., nutrients, plankton, ...), and different modeldata misfit metrics [7] (e.g., RMSE of tracer distributions or the Cabre metric for oxygen minimum zones), further natural applications will target compromise parameters for different regions of the ocean (see, e.g., section 7.2 in [11]) or different BGC and circulation models.



**Figure 1:** bi-objective (simultaneous) optimisation of RMSE misfit w.r.t nutrients and oxygen distributions and OMZ mismatch. The results (published in [7]) stem from optimizing 6 BGC parameters of model MOPS [12] coupled to a 2.8 degree configuration of the MIT global circulation model with 15 depth layers. Projection of parameter values (normalized, colour scale) onto model misfits, in the vicinity of the optimal solution of optimization. Parameter values have been scaled by  $(\theta - \theta_c)/(\theta_u - \theta_l)$ , where  $\theta$  is the parameter value,  $\theta_u$  and  $\theta_l$  are upper and lower boundary constraints, and  $\theta_c = (\theta_u + \theta_l)/2$  is the center of the allowed interval.

### WWW

http://www.uni-kiel.de

### **More Information**

- [1] I. Kriest, S. Khatiwala, A. Oschlies, *Prog. Oceanogr.* 86, 337–360 (2010). doi: 10.1016/j.pocean.2010.05.002
- [2] R.G. Najjar et al., *Glob. Biogeochem. Cycl.* 21, GB3007 (2007). doi: 10.1029/2006GB002857
- K. Deb, Multi-objective optimization using evolutionary algorithms John Wiley & Sons (2009)
- [4] M. Hauschild, M. Pelikan, http://medal-lab. org/files/2011004\_rev1.pdf (2011)
- [5] C. Igel, N. Hansen, S. Roth, *Evolution-ary Computation*, **15**(1), 1–28 (2007). doi: 10.1162/evco.2007.15.1.1

- [6] S. Khatiwala, Glob. Biogeochem. Cycl. 21, GB3001 (2007). doi:10.1029/2007GB002923
- [7] V. Sauerland, I. Kriest, A. Oschlies, A. Srivastav, AGU JAMES, textbf11, (2019). doi: 10.1029/2018MS001510
- [8] V. Sauerland. GitHub (2018). doi: 10.5281/zenodo.1432629
- [9] O. Schütze, et al. Evolutionary Computation, 18(1), 65–96 (2010). doi: 10.1162/evco.2010.18.1.18103
- [10] I. Kriest, et al. Geosci. Model Dev. 10, 127– 154 (2017). doi:10.5194/gmd-10-127-2017
- [11] M. Schartau, et al. *Biogeosciences* 14, 1647– 1701 (2017). doi:10.5194/bg-14-1647-2017
- [12] I. Kriest, A. Oschlies. Geosci. Model Dev. 8, 2929–2957 (2015). doi:10.5194/gmd-8-2929-2015