

Fast spin-up of geochemical tracers in ocean circulation and climate models

Samar Khatiwala

¹Department of Earth Sciences, University of Oxford

Key Points:

- Geochemical tracers have provided great insight into oceanic processes but are prohibitively expensive to simulate in climate models
- A new “sequence acceleration” method is introduced offering speed-ups of 10–25 times for a range of typical geochemical tracer problems
- The new method is completely “black box” and can be applied to any model

Corresponding author: Samar Khatiwala, samar.khatiwala@earth.ox.ac.uk

Abstract

Ocean geochemical tracers such as radiocarbon, protactinium and thorium isotopes, and noble gases are widely used to constrain a range of physical and biogeochemical processes in the ocean. However their routine simulation in global ocean circulation and climate models is hindered by the computational expense of integrating them to a steady state. Here, a new approach to this long-standing “spin-up” problem is introduced to efficiently compute equilibrium distributions of such tracers in seasonally-forced models. Based on “Anderson Acceleration”, a sequence acceleration technique developed in the 1960s to solve nonlinear integral equations, the new method is entirely “black box” and offers significant speed-up over conventional direct time integration. Moreover, it requires no preconditioning, ensures tracer conservation and is fully consistent with the numerical time-stepping scheme of the underlying model. It thus circumvents some of the drawbacks of other schemes such as matrix-free Newton Krylov that have been proposed to address this problem. An implementation specifically tailored for the batch HPC systems on which ocean and climate models are typically run is described, and the method illustrated by applying it to a variety of geochemical tracer problems. The new method, which provides speed-ups by over an order of magnitude, should make simulations of such tracers more feasible and enable their inclusion in climate change assessments such as IPCC.

Plain Language Summary

Radiocarbon and other geochemical tracers have provided great insight into the workings of the ocean but are prohibitively expensive to simulate in climate models. This study introduces a new computational method that can be applied to any model to greatly speed-up simulations of such tracers, enabling their routine inclusion in climate models and thus more effective use of those tracers.

1 Introduction

Numerical models of ocean geochemical tracers such as radiocarbon and $^{231}\text{Pa}/^{230}\text{Th}$ have long been used to gain insight into physical and biogeochemical processes. The distribution of natural radiocarbon has been vital to quantifying ocean ventilation and the uptake of anthropogenic carbon and heat by the ocean, as well as constraining ocean inverse and general circulation models (Toggweiler et al., 1989; Matsumoto et al., 2004; Schlitzer, 2007; Khatiwala et al., 2009; DeVries & Primeau, 2011; Khatiwala et al., 2012; Gebbie & Huybers, 2012; Orr et al., 2017). It is one of the tracers recommended by the Ocean Model Intercomparison Project (OMIP; Orr et al. (2017)) for simulation by models participating in the IPCC’s Coupled Model Intercomparison Project (CMIP). On longer time scales, radiocarbon (Sarnthein et al., 2013; Skinner et al., 2017), the isotope pair ^{231}Pa and ^{230}Th (Yu et al., 1996; McManus et al., 2004) and neodymium isotopes (Rutberg et al., 2000; Goldstein & Hemming, 2003; Bohm et al., 2015), are widely used as proxies of past ocean circulation. Other tracers of physical and biogeochemical relevance include, ventilation tracers such as ideal age (Thiele & Sarmiento, 1990; England, 1995), preformed tracers (Ito et al., 2004; Ito & Follows, 2005; Williams & Follows, 2011; Lauderdale et al., 2013; Khatiwala et al., 2019) and noble gases (Nicholson et al., 2016; Hamme et al., 2019; Cassar et al., 2021).

A principal challenge to simulating such tracers in ocean general circulation models (GCMs) and their more effective and routine use is the computational cost. In many cases, the tracers need to be integrated to a seasonally-repeating equilibrium, an expensive undertaking given that this may take several thousand years due to the slow adjustment time scales of the deep ocean (Holzer & Primeau, 2006; Wunsch & Heimbach, 2008; Khatiwala et al., 2012). For example, the OMIP criteria for equilibrium are that the net air-sea flux of CO_2 be $<0.01 \text{ PgC/y}$ for dissolved inorganic carbon and 98% of ocean volume has a ^{14}C drift of $<0.001\text{‰}$ per year (or a radiocarbon age drift of <10

y per 1000 y). Achieving the first target can take $O(4000-5000)$ years, while radiocarbon takes even longer (as much 8000 years; (Graven et al., 2012; Orr et al., 2017)). As a consequence, only two models participating in CMIP6, the latest IPCC exercise, report having undertaken this simulation. It should be noted that while ocean GCMs and, more generally climate models, exhibit variability on many time scales, the dominant one is the seasonal cycle, which, either via the underlying circulation or forcing (e.g., wind speed for carbon), has a significant impact on the equilibrium solution. The focus of the present study is therefore on finding equilibrium solutions of such seasonally-forced models.

A number of approaches have been proposed to speed-up ocean tracer simulations. One is “offline” models which advect and diffuse passive tracers using a precomputed circulation field. These can be time-stepped much more quickly to an equilibrium than running the full GCM. Some GCMs, such as MITgcm (Marshall et al., 1997) and NEMO (Gurvan Madec and NEMO System Team, n.d.), have offline versions, while the transport matrix method (TMM), which represents tracer advection-diffusion as a sequence of sparse matrix-vector products, is another offline scheme (Khatiwala et al., 2005; Primeau, 2005; Khatiwala, 2007; DeVries & Primeau, 2011). An alternative approach is to directly compute the equilibrium tracer field without performing a transient integration (Merlis & Khatiwala, 2008). The basic idea is to pose the problem as a large, nonlinear system of equations implicitly defined via the model time-stepper. Matrix-free Newton-Krylov (MFNK) has been proposed as a way to solve this system (Merlis & Khatiwala, 2008; Khatiwala, 2008; Li & Primeau, 2008; Bardin et al., 2014; Lindsay, 2017). While for many biogeochemical problems MFNK has been shown to be up to two orders of magnitude faster than conventional time integration, it has a number of drawbacks (discussed in more detail below) which have prevented it from being adopted more widely.

Here, a new approach is presented that, like MFNK, attempts to directly compute a steady state solution but overcomes some of the latter’s drawbacks and limitations. The new scheme is based on a numerical technique known as Anderson Acceleration (AA) or Mixing. Developed in the 1960s by D. G. Anderson (Anderson, 1965, 2019) in the context of nonlinear integral equations, it primarily found application to electronic structure problems in quantum chemistry and material science (Walker & Ni, 2011; Zhang et al., 2020). More recently, there has been a resurgence of interest in AA as a solver for partial differential equations (Walker et al., 2010; Brune et al., 2015) and optimization problems (Zhang et al., 2020; Fu et al., 2020; Tang & Daoutidis, 2022).

The basic idea behind AA, and other so-called sequence acceleration methods (Brezinski, 2000; Brezinski et al., 2018), is to exploit the previous history of the model to extrapolate to a solution that is closer to equilibrium. One major advantage of AA in the context of spin-up is that it only requires the ability to integrate the model with a given initial condition and return the solution. Unlikely MFNK, there is no need for a preconditioner (see below). For the spin-up problem this generally requires a transport matrix representation of the circulation (Khatiwala, 2008; Li & Primeau, 2008), which is rarely available, and a custom implementation tailored to each GCM and tracer problem, a non-trivial undertaking (Lindsay, 2017). Furthermore, and again in contrast with MFNK, the new proposed approach is demonstrated to work with multi-step time-stepping algorithms, such as leapfrog, that are common in ocean models. In that sense, this method is completely “black box”, making it significantly easier to apply to a variety of GCMs and well suited to the batch HPC systems on which they are typically run.

In the next section, the theoretical underpinnings of AA are outlined, followed by details on the practical implementation. Next, the performance of AA is demonstrated by applying it to several tracer problems, including radiocarbon, ventilation tracers and $^{231}\text{Pa}/^{230}\text{Th}$. The paper concludes with a summary and future directions for research.

2 Anderson Acceleration

2.1 Mathematical formulation

We start with a mathematical statement of the problem. A numerical model can be expressed as a function \mathbf{g} that takes in an initial tracer field $\mathbf{x}(0)$ at time $t = 0$, time-steps forward, and returns the tracer field $\mathbf{x}(T)$ at time $t = T$, where T is the forcing period:

$$\mathbf{x}(T) = \mathbf{g}(\mathbf{x}(0)). \quad (1)$$

Here, \mathbf{x} is a vector representation of the tracer field at the “wet” grid points of the GCM mapped into a vector. Note that \mathbf{x} can also represent multiple fields if the model contains more than one tracer or requires the tracer field at more than one time step to step forward the model. The spin-up problem is to seek a solution that repeats after one period, i.e.,

$$\mathbf{x} = \mathbf{g}(\mathbf{x}). \quad (2)$$

Mathematically, the conventional (slow) approach of simply integrating the model until an equilibrium is reached is just a fixed point (FP) iteration scheme:

Given \mathbf{x}_0 ,
for $k = 0, 1, \dots$ **until convergence**
 $\mathbf{x}_{k+1} = \mathbf{g}(\mathbf{x}_k)$

2.2 Matrix-free Newton Krylov

One way to solve eq. 2 is to pose the problem as a nonlinear algebraic system of equations (Merlis & Khatiwala, 2008; Khatiwala, 2008; Li & Primeau, 2008):

$$\mathbf{f}(\mathbf{x}) = \mathbf{g}(\mathbf{x}) - \mathbf{x} = \mathbf{0}. \quad (3)$$

To solve this system, which is not explicitly constructed but is implicitly defined via the model time-stepper, Merlis and Khatiwala (2008) proposed to apply matrix-free Newton Krylov. Recall that Newton’s method is based on constructing a local linear model of the function \mathbf{f} and then iterating (Kelley, 1995; Dennis & Schnabel, 1996):

Given an initial iterate \mathbf{x}_0 ,
for $k = 0, 1, \dots$ **until convergence**
 Solve $\mathbf{J}(\mathbf{x}_k)\Delta\mathbf{x}_k = -\mathbf{f}(\mathbf{x}_k)$, where $\mathbf{J} = \partial\mathbf{f}/\partial\mathbf{x} = \partial\mathbf{g}/\partial\mathbf{x} - \mathbf{I}$
 $\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta\mathbf{x}_k$

Newton’s method is attractive because, with a “good” initial guess, it converges quadratically and can be paired with a “globalization” method such as line search to find a global minimum. Much of the computational effort of this method is in solving the linear Newton equation: $\mathbf{J}\Delta\mathbf{x} = -\mathbf{f}(\mathbf{x})$. This is especially difficult for the current problem because the explicit computation and storage of the Jacobian matrix \mathbf{J} is impractical since here this matrix is large and dense. For example: \mathbf{J} is O(30 TB) for a 1° resolution model with two tracers. For systems such as this it is natural to apply a Krylov subspace method like GMRES (Saad, 2003). The advantage of Krylov methods is that it is not necessary to explicitly form the coefficient matrix; only its action on a given vector is needed (Saad, 2003). Such a method is known as “matrix-free” and the combination of Newton with Krylov is called matrix-free or Jacobian-free Newton Krylov (JFNK), an approach originally motivated by the solution of nonlinear ordinary and partial differential equations (Gear & Saad, 1983; Brown & Hindmarsh, 1986; Brown & Saad, 1990; Chan & Jackson,

1984; Knoll & Keyes, 2004). In the present case, the Jacobian-vector product can be accurately computed via finite differences, which only requires the ability to evaluate $\mathbf{f}(\mathbf{x})$, i.e., integrate the model for one period:

$$\mathbf{J}(\mathbf{x})\Delta\mathbf{x} \approx \frac{\mathbf{f}(\mathbf{x} + \sigma\Delta\mathbf{x}) - \mathbf{f}(\mathbf{x})}{\sigma}$$

Krylov methods are iterative and work by expressing the solution in a small subspace that is built up through repeated evaluation of the product between the coefficient matrix and a given vector (Saad, 2003). They are thus only practical when the number of such evaluations, which for the present problem involves integrating the model for one period, is small. That is rarely the case and it is almost always necessary to precondition the matrix (Saad, 2003). In essence, the preconditioned coefficient matrix is closer to the identity matrix leading to faster convergence. Historically, MFNK has been applied to the solution of nonlinear systems derived from the discretization of partial differential equations (Knoll & Keyes, 2004). This results in sparse Jacobians for which preconditioners can usually be readily constructed. The fact that the Jacobian in the spin-up problem is dense, severely limits the choice to methods in which the preconditioner is applied implicitly, i.e., by solving another linear system. Khatiwala (2008) and Li and Primeau (2008) proposed such an implicit preconditioner that could be applied to any generic biogeochemical model. With this preconditioner, Khatiwala (2008) showed that MFNK can accelerate spin-up by up to two orders of magnitude for a variety of biogeochemical models.

While MFNK can be very effective it has not been widely adopted by the modeling community (CESM is perhaps the only model in which it has been implemented for spinning-up radiocarbon (Lindsay, 2017)). There are several reasons for this. First, preconditioning requires a transport matrix for the ocean model (Khatiwala, 2008; Li & Primeau, 2008) and this is only available for a small handful of GCMs. Preconditioning also requires the ability to compute a Jacobian for the underlying (bio)geochemical model. Second, tracer conservation is not guaranteed and is tricky to ensure. Third, MFNK has not been shown to work with multi-step time stepping schemes, which are commonly used in ocean GCMs. For example, the Adams-Bashforth and leapfrog schemes require information from $t-\Delta t$ and t to compute the solution at $t+\Delta t$, where Δt is the model time step. MFNK can be used with such models by providing a single initial condition, in which case the first time step will typically be an Euler step, with subsequent time steps using information from two or more previous steps. However, the equilibrium solution found by MFNK will be different from that obtained by direct time-stepping to steady state, i.e., if the solution were inserted into the ocean model as an initial condition it will quite likely start drifting. While it is unclear how large or significant this drift would be in practice, the inconsistency may matter if the MFNK solution were, for example, used as an initial condition for a subsequent transient simulation. Lastly, the overall algorithm is cumbersome and difficult to apply, particularly when the model is run on batch HPC systems. This has required writing custom implementations tailored to the tracer problem and GCM (Lindsay, 2017).

2.3 Anderson Acceleration

An alternative approach to MFNK is to try to accelerate the original fixed point iteration, which generates a (slowly converging) sequence \mathbf{x}_k . Sequence acceleration or extrapolation methods seek to transform such sequences into ones that converge faster, an idea that has a rich history in numerical methods (for example Richardson extrapolation) (Brezinski, 2000; Smith et al., 1987; Brezinski et al., 2018). A simple modification of the FP algorithm illustrates the idea:

Given \mathbf{x}_0 ,
for $k = 0, 1, \dots$ **until convergence**
 $\mathbf{x}_{k+1} = \beta \mathbf{g}(\mathbf{x}_k) + (1 - \beta) \mathbf{x}_k$.

Here, $0 < \beta \leq 1$ is known as the damping parameter. This scheme is known variously as “Krasnosel’skil Mann iteration”, “averaged iteration” and “simple mixing” (Zhang et al., 2020). Note that the iteration can also be written as $\mathbf{x}_{k+1} = \mathbf{x}_k + \beta(\mathbf{g}(\mathbf{x}_k) - \mathbf{x}_k) = \mathbf{x}_k + \beta \mathbf{f}(\mathbf{x}_k)$, which shows that the new iterate is the current one plus a fraction β of the current residual. This simple modification can sometimes improve convergence at no extra cost and suggests that it may be possible to exploit the information contained in previous iterates to speed-up FP iteration.

Such a scheme was devised in the 1960s by D. G. Anderson (Anderson, 1965, 2019), who came up with an elegant approach based on taking a wighted average of several previous iterates such that, were \mathbf{g} linear, the residual is minimized (Fang & Saad, 2009; Zhang et al., 2020). Known alternatively as Anderson Acceleration, Mixing or Extrapolation, this method is also called “direct inversion in the iterative subspace” in computational chemistry and “Pulay mixing” in material science, the fields in which it first found wide application. Mathematically, this gives the iteration

$$\mathbf{x}_{k+1} = \sum_{j=0}^{m_k} \alpha_j^{(k)} \mathbf{g}(\mathbf{x}_{k-m_k+j}), \quad (4)$$

where the number of previous iterates is m_k+1 and the α_j ’s minimize the norm of the weighted residual \mathbf{f} of those iterates:

$$\text{minimize } \left\| \sum_{j=0}^{m_k} \alpha_j^{(k)} \mathbf{f}(\mathbf{x}_{k-m_k+j}) \right\|_2^2$$

subject to the normalization $\sum_{j=0}^{m_k} \alpha_j^{(k)} = 1$. Usually, this iteration is combined with “damping” to give:

$$\mathbf{x}_{k+1} = \beta_k \sum_{i=0}^{m_k} \alpha_i^{(k)} \mathbf{g}(\mathbf{x}_{k-m_k+i}) + (1 - \beta_k) \sum_{i=0}^{m_k} \alpha_i^{(k)} \mathbf{x}_{k-m_k+i}. \quad (5)$$

Also in practice, the constrained least-squares problem for the m_k+1 α_j ’s is replaced by an unconstrained one for m_k γ_i ’s (Fang & Saad, 2009; Walker & Ni, 2011) so that the next iterate can be written as:

$$\mathbf{x}_{k+1} = \mathbf{g}(\mathbf{x}_k) - \sum_{j=0}^{m_k-1} \gamma_j^{(k)} [\mathbf{g}(\mathbf{x}_{k-m_k+j+1}) - \mathbf{g}(\mathbf{x}_{k-m_k+j})]. \quad (6)$$

The coefficients γ_i are found by minimizing

$$\|\mathbf{f}_k - \tilde{\mathbf{F}}_k \boldsymbol{\gamma}^{(k)}\|_2^2, \quad (7)$$

where $\tilde{\mathbf{F}}_k$ is a matrix whose m_k columns are $\Delta \mathbf{f}_{k-m}, \dots, \Delta \mathbf{f}_{k-1}$ and $\Delta \mathbf{f}_{k-m}, \dots, \Delta \mathbf{f}_{k-1}$. The AA solution (eq. 6) can be seen as an extrapolation of the original iterate $\mathbf{g}(\mathbf{x}_k)$ plus m_k previous iterates (Zhang et al., 2020). Note that substituting $\beta_k = 1$ in eq. 5 gives the undamped step (eq. 4), and $\beta_k = 1, m_k = 0$ recovers the original FP iteration $\mathbf{x}_{k+1} = \mathbf{g}(\mathbf{x}_k)$.

Two aspects of AA are worth emphasizing. First, it has negligible overhead, with the computationally expensive part of the calculation being the cost of running the ocean model. Indeed, it can be run on the frontend of a HPC system or a single core of a compute node. This is in contrast to MFNK which has significant overhead and resource requirements for computing, factoring and applying the preconditioner. Second, tracer conservation is always satisfied as AA takes a linear combination of previous iterates (which are simply the outputs of the model).

3 Implementation

While the basic AA algorithm is quite straightforward, it is convenient to start with an existing implementation, of which there are several. For example, PETSc, a widely used numerical software library for solution of linear and nonlinear equations (Balay et al., 2022), has a full-featured AA implementation (and other variations that fall under the category of nonlinear GMRES (Brune et al., 2015)). However, its disadvantage is that PETSc’s architecture does not allow for checkpointing and restarts. This is essential for any practical implementation of AA that will be applied to the ocean spin-up problem. Instead, the “reference” implementation of Walker (2010) written in MATLAB is used as a starting point as it is well documented, compact, transparent, easy to modify and port to other languages such as python.

In most AA implementations, the algorithm has two main parameters, the damping parameter, β_k , and the maximum memory parameter, m_{\max} , the maximum number of previous iterates stored from which the new iterate is computed. Typically, these are constant, but in so-called non-stationary AA they can vary over time (Chen & Vuik, 2022b, 2022a, see below). Note that AA starts to be applied immediately after the first iteration, and at any iteration k there will be $m_k = \min(k, m_{\max})$ iterates in memory from which the new iterate is constructed. When m_k reaches m_{\max} , the oldest iterate in memory (on the LHS of $\tilde{\mathbf{F}}_k$, the coefficient matrix for the least-squares problem (eq. 7)) is discarded and the latest one added (to the RHS of $\tilde{\mathbf{F}}_k$). In Walker’s implementation, at each iteration, $\tilde{\mathbf{F}}_k$ is monitored and if found to be poorly conditioned, the oldest vector is discarded. In practice, the QR factorization of $\tilde{\mathbf{F}}_k$ is used to solve the least squares problem, and each time the matrix is updated the factorization is also updated without recomputing it using MATLAB’s QRDELETE function (Walker, 2010). Termination of the algorithm is based on $\|\mathbf{f}(\mathbf{x}_k) - \mathbf{g}(\mathbf{x}_k) - \mathbf{x}_k\|_2$ going below a specified tolerance or the maximum number of specified iterations being reached.

For the spin-up problem, Walker’s original implementation has been modified to incorporate a number of additional features. These include:

1. Checkpointing to facilitate running on HPC machines with batch submission systems;
2. Hooks to signal convergence based on external criteria (e.g., the air-sea flux of CO_2 meeting the OMIP criterion);
3. The ability to restart AA by “zeroing” the memory periodically, if the algorithm stagnates, or because of an external signal based on a user-specified condition;
4. The ability to run multiple instances of AA to spin-up multiple independent tracers (or independent sets of tracers) simulated simultaneously in the model (the tracers/tracer sets can have different termination tolerances which can take different number of iterations to reach); and
5. The facility to scale different tracers in a multi-tracer problem so that they have the same order of magnitude.

A python version (which can also be called from MATLAB) is under development and is currently being tested.

To use the code, a user must supply:

- A driver routine that read in the initial iterate and calls AA.
- A “wrapper” (the \mathbf{g} function seen by AA) around their model to exchange data between AA and the model (the initial condition on input, the state after running the model for one period on output).
- An optional function that checks for convergence and passes back that information to AA via the wrapper.

The details of the driver and wrapper function will depend on the model and computing system on which the calculations are performed. Fig. 1 illustrates a typical workflow on a batch HPC system but the implementation of AA presented here is general enough to allow a variety of different workflows.

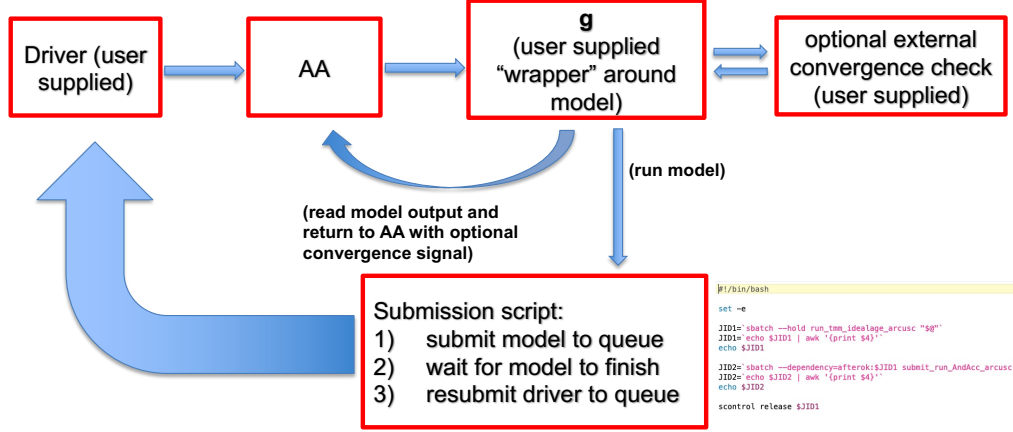


Figure 1. Schematic of workflow for using the implementation of Anderson Acceleration presented here on a HPC machine with a batch queuing system. The sequence starts by submitting the driver to the queue. This calls the AA algorithm which in turn, after saving the full current state of the algorithm to a checkpoint file, calls the wrapper function in “run mode.” The wrapper function maps the input vector to model fields, calls a script that submits both the model and driver to the queue, and then returns control to AA, which then terminates. As seen in the example script in the figure, execution of the driver is conditional on successful completion of the model run. Once that happens, the driver relaunches and calls AA, which restores the state of the algorithm by reading the checkpoint file and then calls the wrapper in “read mode”. The wrapper reads the model output, remaps it to a vector and returns the solution back to AA along with (optional) convergence information. AA computes the next iterate and calls the wrapper again in “run mode”. And so on. This sequence is repeated until convergence. The example script shown is for the SLURM scheduling software, but PBS and other systems can be similarly used.

4 Examples

In the following, the application of AA to spinning-up a variety of geochemical and ventilation tracers is illustrated. For each case the number of simulated years required to reach equilibrium using AA is compared with that for direct integration (DI). Also compared are the final steady state solutions. Unless stated otherwise, AA was used with $\beta = 1$ (no damping), $m_{\max} = 50$ and no restarts. And with one exception, all examples were run on HPC systems with either the SLURM or PBS schedulers. The exception is the first example which was run on a local server as it involves a model that does not require parallel computing resources.

4.1 Example 1: Online spin-up of abiotic carbon cycle model

The first example concerns the spin-up of an OCMIP-type abiotic carbon cycle model (Orr et al., 1999) running online within the University of Victoria Earth System Climate Model (UVic ESCM) (version 2.9) (Weaver et al., 2001). This particular configuration of UVic ESCM was tuned to preindustrial conditions (Muglia & Schmittner, 2015; Muglia et al., 2018; Khatiwala et al., 2019). While UVic ESCM features a sophisticated biogeochemical model, MOBI (Model of Ocean Biogeochemistry and Isotopes; Schmittner and Somes (2016)), it can be trivially reduced to a simple abiotic model for dissolved inorganic carbon (DIC) and ^{14}C in DIC (DI^{14}C) by switching off the biological source/sink terms with C preprocessor directives. DI^{14}C is coupled to DIC via the air-sea gas exchange term, which is parameterized in a manner similar to the OCMIP-2 and OMIP protocols (Orr et al., 1999, 2017), although with different gas transfer coefficients. The ocean GCM component of UVic ESCM has a resolution of $1.8^\circ \times 3.6^\circ$ and 19 vertical layers. Tracers are time-stepped with a leapfrog scheme requiring two initial conditions.

As is typical of ocean GCMs and climate models, at the start of a run UVic ESCM reads initial conditions from a (netcdf format) restart file, if one is available. The file contains all the physical (dynamical) and tracer fields required to continue a previous integration. Here, a restart file from a previous dynamical model spin-up (Muglia & Schmittner, 2015) is used. At each iteration of AA, the DIC and DI^{14}C fields (for both leapfrog time steps) in this file are overwritten by those provided by the AA algorithm (via the wrapper function). At the end of the (year long) integration, the final DIC and DI^{14}C fields (written to another restart file) are read by the wrapper, remapped to a vector, and passed back to the AA algorithm.

Fig. 2 (top) shows the air-sea flux of CO_2 for DIC and the fraction of ocean volume with a radiocarbon drift of $<0.001\%$ per year versus the number of years of integration. The horizontal lines are the respective OMIP criteria for equilibrium. Evidently, with DI it takes ~ 5200 y to reach the OMIP criteria for air-sea flux, and ~ 6500 y for radiocarbon drift. In contrast, AA requires ~ 450 and ~ 470 years, respectively, to reach those criteria, implying a speed-up of over a factor of 11 for the coupled system. The bottom plots compare the equilibrium DI and AA solutions, showing that AA reproduces the DI solution. Note that both leapfrog time steps are plotted.

An additional experiment (not shown) was performed in which AA was applied to the same carbon cycle model run offline via the TMM (Khatiwala et al., 2005; Khatiwala, 2007, 2018) with transport matrices (TMs) from the same UVic ESCM configuration used for the above online simulations. The TMM requires only a single initial condition to be specified, leading to a problem that is half the size of that with AA applied online with leapfrog. No noticeable difference in performance was found between the online and offline cases. To the extent that these results can be generalized, it suggests that AA can be successfully applied without loss of performance to find equilibrium solutions of GCMs with multi-step time-stepping schemes. Moreover, the solutions would be fully consistent with the numerics of the model.

4.2 Example 2: Abiotic carbon cycle with ECCO state estimates

To demonstrate that AA performs well at even significantly higher spatial resolution, a similar abiotic model as above was spun-up, but simulated offline using the TMM. For this, TMs extracted from the “Estimating the Circulation and Climate of the Ocean” (ECCO) version 4 (release 4) ocean state estimate (Forget et al., 2015) were used. ECCO is based on fitting the MITgcm model (Marshall et al., 1997) to a variety of observations using an adjoint approach to derive a dynamically-consistent ocean state estimate (Stammer et al., 2004; Wunsch & Heimbach, 2007). Version 4 uses observations between 1992–2017, and a MITgcm configuration with a “latitude-longitude-cap” grid (LLC90) with horizontal resolution ranging from 22–110 km and 50 vertical levels. To further assess the

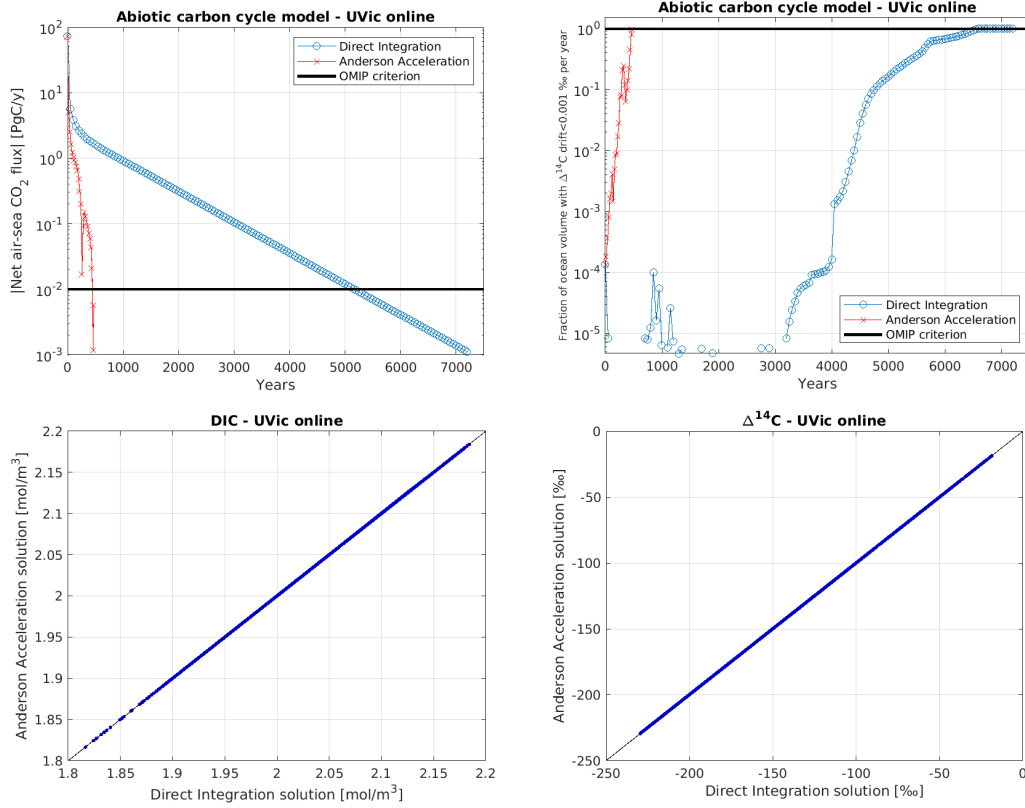


Figure 2. Top: Plots of the net air-sea flux of CO_2 (left) and fraction of ocean volume with a radiocarbon drift of $<0.001\text{‰}$ per year (right; equivalent to a ^{14}C age drift of <10 y per 1000 y) versus number of simulated years for direct time integration and AA. Simulations were performed online with UVic ESCM. Horizontal lines are the respective OMIP criteria for equilibrium, namely, a net CO_2 flux <0.01 PgC/y for DIC and 98% for radiocarbon drift (Orr et al., 2017). Bottom: Comparison of the AA equilibrium solution (vertical axis) with that computed by direct time integration (horizontal axis). The left plot is for DIC and the right for $\Delta^{14}\text{C}$. The black diagonal line is the 1:1 relationship. The solution at both the initial time steps required by UVic ESCM’s leapfrog time-stepping scheme are shown.

impact of resolution on the performance of AA, a second spin-up experiment was performed using TMs extracted from the MITgcm ECCO-GODAE ocean state estimate (Stammer et al., 2004; Wunsch & Heimbach, 2007). This version, which was constrained to observations between 1992–2004, has a lower resolution of $1^\circ \times 1^\circ$ and 23 vertical levels. In both cases, TMs representing a monthly mean climatology over the estimation period were used. The OCMIP-2/OMIP abiotic carbon cycle model was forced with 6-hourly winds from the CORE-II atmospheric reanalysis (Large & Yeager, 2004), and temperature, salinity and sea ice concentration from the corresponding state estimate. TMs and forcing fields were interpolated to the current time step before being applied.

With ECCO-v4 (top row of Fig. 3), direct integration takes ~ 4300 y and ~ 8000 y to reach the OMIP CO_2 flux and radiocarbon drift criteria, respectively, while AA requires ~ 350 y and ~ 470 y, respectively, to meet them. AA is thus faster by an overall factor of ~ 12 . With the lower resolution ECCO-GODAE configuration, the corresponding times are ~ 5400 y and ~ 7200 y for DI, and ~ 350 y for meeting both criteria with AA, a speed-up of ~ 15 . Resolution thus does not seem to significantly impact the per-

formance of the method. For completeness, Fig. 4 shows that for ECCO-v4 the equilibrium AA solution agrees very well with the DI solution.

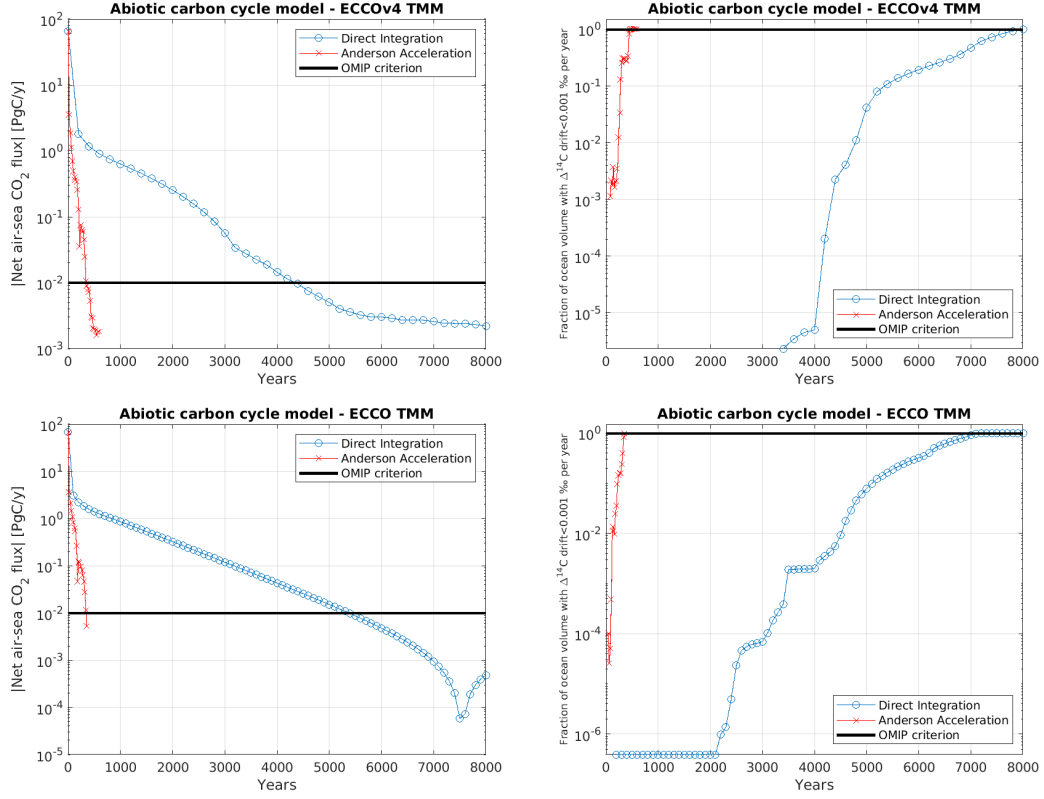


Figure 3. Plots of the net air-sea flux of CO_2 (left) and fraction of ocean volume with a radiocarbon drift of $<0.001\%$ per year (right; equivalent to a ^{14}C age drift of <10 y per 1000 y) versus number of simulated years for direct time integration and AA. Horizontal lines are the respective OMIP criteria for equilibrium, namely, a net CO_2 flux <0.01 PgC/y for DIC and 98% for radiocarbon drift (Orr et al., 2017). The top row shows simulations with the TMM with transport matrices from the ECCO-v4 state estimate, while the bottom row are simulations performed with the lower resolution ECCO-GODAE TMs.

4.3 Example 3: Ideal age

Next, AA is applied to the ideal age tracer (Thiele & Sarmiento, 1990; England, 1995; Holzer & Hall, 2000), a passive tracer with a zero surface boundary condition and an interior source of “1” representing aging. The steady state solution is the “mean age”, the average time since a water parcel was last in contact with the surface (Holzer & Hall, 2000), a widely used metric of ocean ventilation time scales. The ideal age was simulated with the TMM using TMs extracted from the MITgcm ECCO-GODAE ocean state estimate (see above). Fig. 5 (left) comparing direct integration and AA shows that the former takes ~ 4500 y to reach the equivalent OMIP criterion for radiocarbon age drift, while the latter takes ~ 200 y. This is a speed-up by a factor of ~ 22 . The right panel compares the two equilibrium solutions. AA large reproduces the DI solution with a RMS difference of ~ 1.85 y. The solutions differ at a few isolated grid points where the model evolves slowly due to weak exchange with surrounding waters. At those points the AA mean age

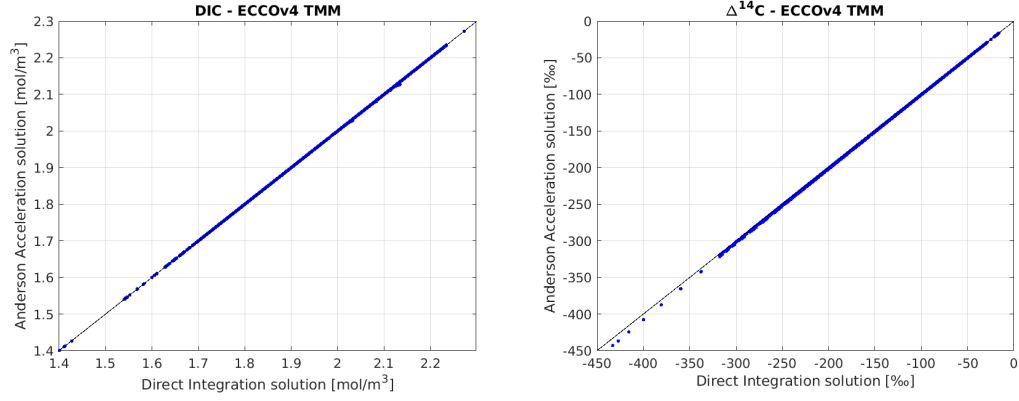


Figure 4. Comparison of the AA equilibrium solution (vertical axis) with that computed by direct time integration (horizontal axis) for an abiotic carbon cycle model simulated with the TMM with ECCO-v4 TMs. The left plot is for DIC and the right for $\Delta^{14}\text{C}$. The black diagonal line is the 1:1 relationship.

is systematically older than the corresponding DI values, suggesting that the DI solution hasn't fully equilibrated.

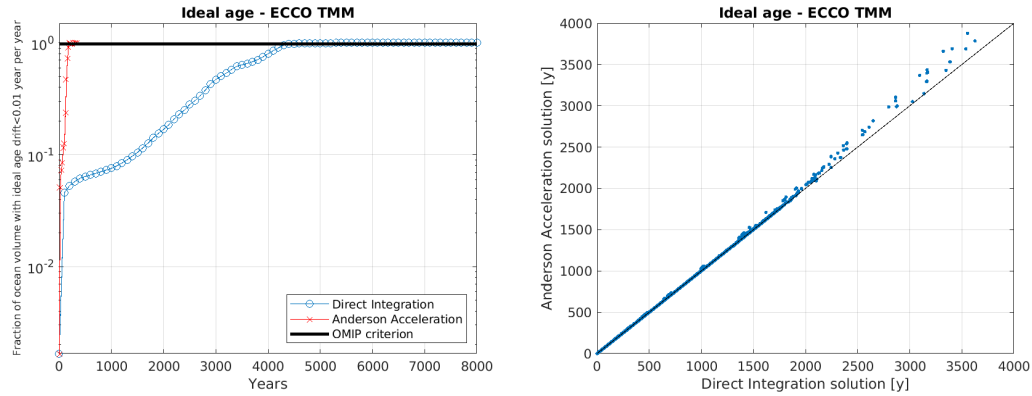


Figure 5. Left: Fraction of ocean volume with an ideal age drift of <10 y per 1000 y versus number of simulated years for direct time integration and AA. Horizontal line is the OMIP radiocarbon criterion for equilibrium (Orr et al., 2017). Right: Comparison of the AA equilibrium solution (vertical axis) with that computed by direct time integration (horizontal axis).

4.4 Example 4: Preformed tracers

A “preformed” tracer is a conservative tracer whose concentration is set at the sea surface and is passively transported into the interior by ocean circulation. Such tracers are often used to study and quantify the strength of ocean carbon pumps by propagating surface distributions of nutrients, dissolved oxygen (O_2) and DIC (Ito et al., 2004; Ito & Follows, 2005; Williams & Follows, 2011; Lauderdale et al., 2013; Khatiwala et al., 2019). For this example, preformed PO_4 and O_2 were spun-up using monthly mean fields from World Ocean Atlas 2018 (WOA18; Garcia et al. (2018)) as surface boundary conditions. The boundary conditions were propagated into the ocean interior using the TMM

with TMs extracted from the same configuration of UVic ESCM as above. AA was applied separately to each tracer (until a specified tolerance for the norm of the residual was reached), but both tracers were simulated simultaneously. That is, two instances of AA were run within the same overall driver.

Fig. 6 shows the results. Unlike the previous examples, here and in the following set of examples, there is no physical criterion for “convergence”. Such tracers are typically integrated for several thousand years (Lauderdale et al., 2013; Khatiwala et al., 2019), at which point the solution is considered to be in equilibrium. Therefore, the number of iterations required to achieve the same residual norm as that reached by direct integration after a specified number of years is used to assess AA’s performance. (Recall that the residual norm is the norm of $\mathbf{f}(\mathbf{x})$, the difference between the initial condition \mathbf{x} and the solution after one year of integration.) If this is 4000 y, AA requires 260 y for O_2 and 160 y for PO_4 , a speed-up of ~ 15 and ~ 25 , respectively.

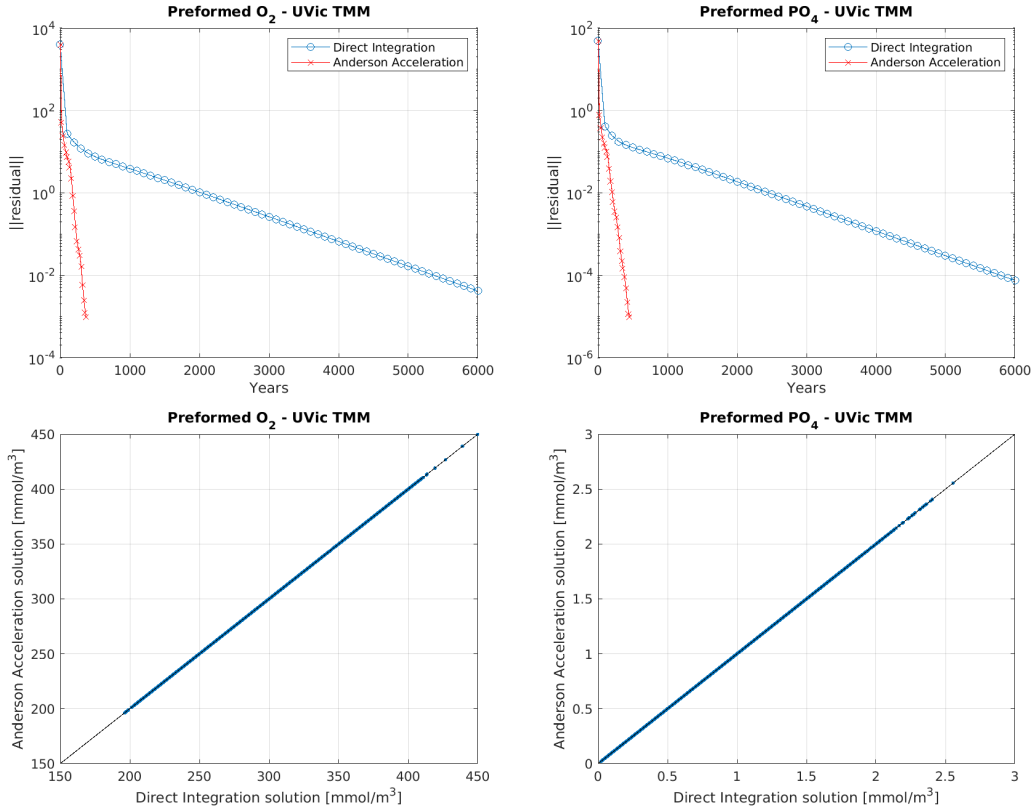


Figure 6. Top: Residual norm versus number of simulated years for direct time integration and AA for preformed O_2 (left) and preformed PO_4 (right). Bottom: Comparison of DI and AA equilibrium solutions.

This example illustrates how the AA implementation described here can be used to spin-up multiple, independent tracers simultaneously. A problem where this facility can be particularly advantageous is that of computing the distribution of source water fractions, i.e., the fraction of water at any point in the ocean that was last in contact with a given surface patch (e.g., Khatiwala et al., 2001; Haine & Hall, 2002; Primeau, 2005). The problem is in fact quite similar to that of preformed tracers, with the difference being that the boundary condition is fixed at “1” on the patch in question and “0” on the rest of the ocean surface. The steady state solution is the source water fraction. Like the

mean age, water mass fractions are an important and widely used metric of ocean ventilation but aren't routinely simulated because of the need for extended integrations to capture the ocean's long diffusive time scales (Primeau, 2005; Holzer & Primeau, 2006; Khatiwala et al., 2012). It should be noted that there is value in the directly integrated solution to this problem. The time derivative of the transient solution (Haine & Hall, 2002) is the so-called "boundary propagator" (BP), a type of Green's function for the advection-diffusion equation that can be interpreted as a probability density function of the time and location of last surface contact for any water parcel (Holzer & Hall, 2000). In addition to its intrinsic value in rigorously characterizing ocean circulation, the BP has been used to estimate uptake of anthropogenic carbon (Khatiwala et al., 2009) and heat (Zanna et al., 2019) by the ocean. The 0th moment of the BP is the water mass fraction and the 1st moment is the mean age (e.g., Waugh et al., 2003). There is work underway to investigate whether AA can be used to compute the moments of the BP efficiently, and if those moments can be used to approximate the full boundary propagator.

4.5 Example 5: Protactinium and thorium isotopes

In this and the next example two geochemical problems are considered. The first involves the tracers ²³¹Pa and ²³⁰Th, whose ratio is widely used as a paleoproxy for the strength of the Atlantic Meridional Overturning Circulation (Yu et al., 1996; McManus et al., 2004). These particle reactive tracers are produced by uranium decay at (different) constant, spatially-uniform rates in the ocean, and in turn undergo radioactive decay. They are absorbed onto and desorbed from sinking particles in a process termed reversible scavenging (Bacon & Anderson, 1982). Weaker scavenging of ²³¹Pa relative to ²³⁰Th causes it to be advected further and have a longer residence time than the latter (Yu et al., 1996; Henderson & Anderson, 2003), which is the basis for the use of the ratio of these tracers as a circulation proxy. These tracers are now implemented in many ocean and climate models used for paleoclimate studies (e.g., Rempfer et al., 2017; Gu & Liu, 2017; van Hulst et al., 2018; Missiaen et al., 2020; Sasaki et al., 2022).

The two tracers have also been incorporated into MOBI, the biogeochemical model embedded within UVic ESCM (see above), as a separate module. A full description, which is in the process of being written for publication elsewhere, is beyond the scope of this paper but, briefly, their implementation closely follows Siddall et al. (2005). Scavenging from 4 different particle types (particulate organic carbon, opal, calcium carbonate and lithogenic particles) are treated, with particle concentration fields taken from Siddall et al. (2005). (The implementation also allows particle concentration fields to be taken directly from those simulated simultaneously by MOBI, but that feature is not used here.) Scavenging coefficients are from Hayes et al. (2015). MOBI and the Pa/Th module can be run either online within UVic ESCM or offline via the TMM (in which case TMs from any ocean model can be used). Here, for computational efficiency it was run via the TMM using TMs from UVic ESCM as described above. Both tracers were spun-up simultaneously.

As in the above example, in the absence of physical criteria for convergence the residual norm during AA iterations is monitored. As seen in Fig. 7 (top), both tracers reach equilibrium relatively quickly, with direct integration taking ~3000 y for ²³¹Pa and even less for ²³⁰Th. Since both tracers are almost always simulated simultaneously the residual norm at 3000 y is taken as convergence criteria. Using AA this takes ~325 y for ²³¹Pa and ~130 y for ²³⁰Th, i.e., AA offers a speed-up by a factor of ~10. The bottom plots show there is very good agreement between the DI and AA solutions.

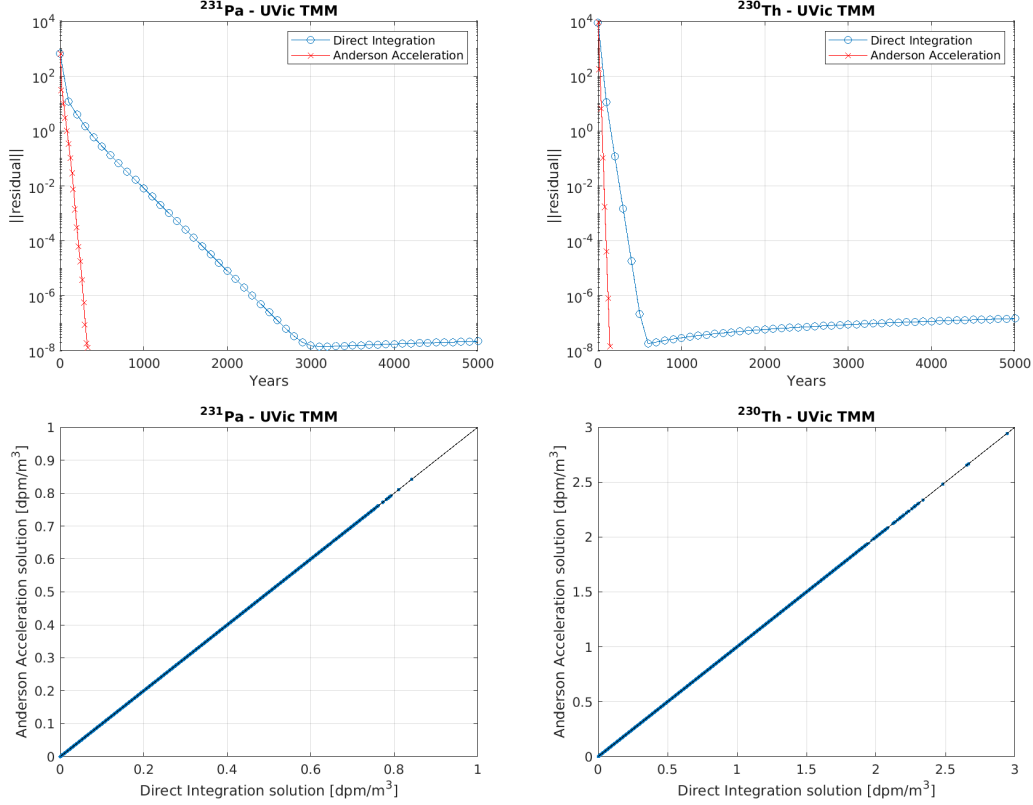


Figure 7. Top: Residual norm versus number of simulated years for direct time integration and AA for ^{231}Pa (left) and ^{230}Th (right). Bottom: Comparison of DI and AA equilibrium solutions.

4.6 Example 6: Zinc cycling model

Lastly, a model of the oceanic cycle of zinc (Zn), a micronutrient important for phytoplankton growth (Morel et al., 2014; Vance et al., 2017), is spun-up. The model is that of de Souza et al. (2018) (see also Vance et al. (2017)) in which the biological uptake of Zn is linked to that of phosphate (PO_4) via a stoichiometric parameter that is a nonlinear function of the Zn concentration. The uptake of PO_4 is in turn diagnosed by restoring the surface concentration of PO_4 (Najjar et al., 2007) toward a seasonally-evolving climatology from WOA18 (Garcia et al., 2018), with a fraction of the uptake instantaneously converted to dissolved organic phosphorus (DOP). The remaining uptake (along with that of Zn) is exported out of the euphotic layer as a particulate flux which is remineralized with depth according to a power law (Martin et al., 1987). The model thus consists of three coupled tracers. The model is coupled to the TMM and, as in de Souza et al. (2018), run with TMs from a MITgcm configuration with a resolution of $2.8^\circ \times 2.8^\circ \times 15$ levels.

In this example, to investigate the effect of the maximum memory parameter on AA's performance, three different values of m_{max} were tried: $m_{\text{max}} = 50$ with no restart as in the previous examples; and $m_{\text{max}} = 30$ and $m_{\text{max}} = 40$, both with restarts when the maximum memory was reached. As is evident in Fig. 8 (top), the latter two performed better than $m_{\text{max}} = 50$ without restart. While one might naively think that the more information retained for AA to exploit the better its performance, this example shows that is not necessarily the case. Retaining more iterates may lead to out-of-date infor-

mation being used, degrading performance (Walker, 2010). Or it may lead to poor conditioning of the least-squares problem as redundant information is added to $\tilde{\mathbf{F}}$.

Using the year 3000 residual norm with direct integration as a convergence criteria, AA with $m_{\max} = 30$ takes ~ 440 y for both PO_4 and Zn, a factor of ~ 7 speed-up. Regardless of the value of m_{\max} , AA converges to the same equilibrium solution, which agrees well with DI (bottom plots).

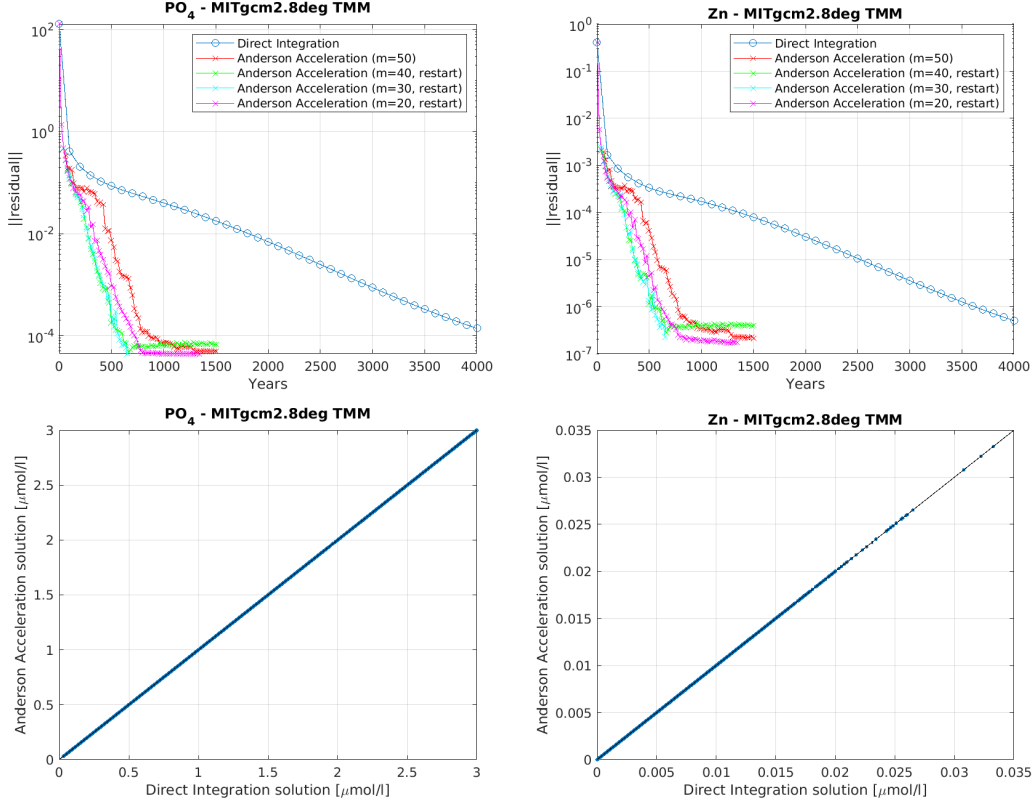


Figure 8. Top: Residual norm versus number of simulated years for direct time integration and AA for PO_4 (left) and zinc (right). Bottom: Comparison of DI and AA equilibrium solutions. The results for DOP are similar and not shown.

In this problem, the inventories of Zn and total phosphorus (the sum of PO_4 and DOP) should be conserved (within numerical accuracy). Reassuringly, the global mean concentration of Zn (phosphorus) for the AA solution is found to deviate from its initial value by $\sim 0.058\%$ ($\sim 0.04\%$), while that for the DI solution is $\sim 0.047\%$ ($\sim 0.033\%$). This demonstrates another important aspect of AA, namely the ability to conserve tracer.

5 Summary and Future Directions

In this study a new method for spin-up of passive tracers in periodically-forced ocean models is described and applied to several widely simulated geochemical and ventilation tracers. The new approach, based on a sequence acceleration technique called Anderson Acceleration or Mixing, offers speed-ups of between 10–25 times over conventional direct time integration. Also described is an implementation that is tailored for the spin-up problem and designed to work on the batch HPC systems on which ocean GCMs and

climate models are typically run. The algorithm has two main tunable parameters, the damping parameter β and the maximum memory parameter m_{\max} . For the problems treated here, $\beta = 1$ and $m_{\max} = 50$ were found to work well. But their optimal values are likely to depend on the specific tracer problem and spatial resolution of the underlying ocean GCM and some experimentation on the problem at hand will likely be needed to find them.

While for some problems matrix-free Newton Krylov, another proposed approach to the spin-up problem (Merlis & Khatiwala, 2008; Khatiwala, 2008; Li & Primeau, 2008), may perform better, AA offers a number of advantages. Unlike MFNK, AA is completely black box, requiring no preconditioners and thus no transport matrices for the underlying GCM; it has been demonstrated to work on models with multi-step time stepping schemes; by construction, tracer conservation is ensured; and it is well suited to batch HPC systems, a particularly complicated aspect of using MFNK (Lindsay, 2017).

The promising results shown by AA in this study suggest a number of avenues for future research. An obvious one is whether AA can be applied to more complex biogeochemical models. In preliminary experiments with a typical NPZD-type model with several interacting tracers, it was found that, after a large initial drop in the residual norm, AA tends to stagnate or displays oscillatory behavior. Further experiments exploring whether this behavior can be ameliorated with appropriate choices of β and m_{\max} are underway. Another possibility is to apply the non-stationary variants of AA that have been recently proposed (Chen & Vuik, 2022b, 2022a). In such schemes, β_k and m_k are adjusted over time, possibly dynamically (by solving an optimization subproblem) (Chen & Vuik, 2022a). Such variants come at the cost of additional function evaluations but may help stabilize AA for more complex problems, especially stiff problems with multiple time scales.

A second avenue is to explore whether AA can be combined with MFNK to overcome some of the latter’s drawbacks. In particular, Eyert (1996) and Fang and Saad (2009) have shown that there is a deep connection between quasi-Newton (QN) methods and AA (and other acceleration methods). In quasi-Newton, the Jacobian is not recalculated at each iteration but is “updated” by using information from previous iterations. A well known method for this is due to Broyden who came up with a remarkable scheme to modify the Jacobian (or its inverse) via a low rank update (which requires storing just a few vectors) (Nocedal & Wright, 2006; Fang & Saad, 2009; Brown & Brune, 2013). Such “limited memory” QN methods are widespread in scientific computing. The relation between AA and QN can be seen by writing eq. 6 as:

$$\mathbf{x}_{k+1} = \mathbf{g}(\mathbf{x}_k) - \tilde{\mathbf{G}}_k \boldsymbol{\gamma}^{(k)}$$

where $\tilde{\mathbf{G}}_k$ is a matrix whose m_k columns are $\Delta \mathbf{g}_{k-m}, \dots, \Delta \mathbf{g}_{k-1}$, with $\Delta \mathbf{g}_i = \mathbf{g}(\mathbf{x}_{i+1}) - \mathbf{g}(\mathbf{x}_i)$ (Walker, 2010). Substituting the normal equations solution to the least squares problem $\|\mathbf{f}_k - \tilde{\mathbf{F}}_k \boldsymbol{\gamma}^{(k)}\|_2^2$, and rearranging, AA can be written as:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \left[(-\mathbf{I}) + \tilde{\mathbf{G}}_k (\tilde{\mathbf{F}}_k' \tilde{\mathbf{F}}_k)^{-1} \tilde{\mathbf{F}}_k' \right] \mathbf{f}_k.$$

Comparing this with Newton

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{J}_k^{-1} \mathbf{f}_k$$

we see that AA implicitly constructs an approximate inverse Jacobian that is a rank m update to $-\mathbf{I}$. An interesting question is whether this can be exploited in some way. One possibility is to insert a few iterations of AA between each iteration of MFNK, using the former to precondition the inner GMRES iterations within the latter. Updating Jacobians via low-rank updates is a common strategy in quasi-Newton (Brown & Brune, 2013) and, as envisioned here, the preconditioner can be applied efficiently and recursively by storing just a few vectors (Nocedal, 1980; Nocedal & Wright, 2006).

A third direction is to investigate whether AA be used to accelerate the *dynamical* equilibration of seasonally-forced ocean models. Indeed, this problem was one of the

primary motivations for the development of the TMM (Khatiwala et al., 2005). And it was also the context in which MFNK was first applied to the ocean spin-up problem (Merlis & Khatiwala, 2008). One could either apply AA to the full model state or, simpler and as proposed by Khatiwala et al. (2005), interleave equilibration by AA of active tracers (temperature and salinity) with conventional direct integration to adjust the dynamical (velocity and pressure) fields.

Lastly, within ocean models there are a number of components that require efficient, scalable and robust solvers. One such is the sea ice component where the complex, nonlinear physics remains a challenge for the iterative schemes (e.g., MFNK) currently being used (Lemieux et al., 2012; Losch et al., 2014; Kimmritz et al., 2017). Anderson Acceleration may be just the tip of the iceberg in terms of sequence acceleration methods (Brezinski et al., 2018) that may be worth pursuing.

6 Open Research

The TMM software and associated data required to perform the simulations presented here are available from Khatiwala (2018) (<http://doi.org/10.5281/zenodo.1246300>). UVic ESCM and MOBI codes are available from Schmittner et al. (2022) (<http://doi.org/10.5281/zenodo.7159021>). Scripts used to perform the calculations and reproduce the figures presented here, as well as output data are available from Khatiwala (2022) (<http://doi.org/10.5281/zenodo.7164028>).

For the purpose of review, the Anderson Acceleration code described in this study is attached as a supplement. If and when this manuscript is accepted it will be placed in a publicly accessible and citable location such as GitHub/zenodo.

Acknowledgments

This research was supported by UK NERC grants NE/T009357/1 and NE/P019218/1. Computing resources were provided by the University of Oxford Advanced Research Computing (ARC) facility (<http://dx.doi.org/10.5281/zenodo.22558>), and the Climate Simulation Laboratory at NCAR’s Computational and Information Systems Laboratory (ark:/85065/d7wd3xhc), sponsored by the National Science Foundation and other agencies.

References

- Anderson, D. G. (1965). Iterative procedures for nonlinear integral equations. *J. Assoc. Comput. Mach.*, 12, 547-560. doi: 10.1145/321296.321305
- Anderson, D. G. (2019). Comments on “Anderson Acceleration, Mixing and Extrapolation”. *Numer. Algor.*, 80, 135-234. doi: doi:10.1007/s11075-018-0549-4
- Bacon, M. P., & Anderson, R. F. (1982). Distribution of thorium isotopes between dissolved and particulate forms in the deep sea. *J. Geophys. Res.*, 87, 2045, 216-237. doi: 10.1029/JC087iC03p02045
- Balay, S., Abhyankar, S., Adams, M. F., Benson, S., Brown, J., Brune, P., ... Zhang, J. (2022). *PETSc/TAO users manual* (Tech. Rep. No. ANL-21/39 - Revision 3.18). Argonne National Laboratory.
- Bardin, A., Primeau, F., & Lindsay, K. (2014). An offline implicit solver for simulating prebomb radiocarbon. *Ocean Modell.*, 73, 45-58. doi: 10.1016/j.ocemod.2013.09.008
- Bohm, E., Lippold, J., Gutjahr, M., Frank, M., Blaser, P., Antz, B., ... Deininger, M. (2015). Strong and deep Atlantic meridional overturning circulation during the last glacial cycle. *Nature*, 517(7532). doi: 10.1038/nature14059
- Brezinski, C. (2000). Convergence acceleration during the 20th century. *J. Comput. Appl. Math.*, 122, 1-21.

- 589 Brezinski, C., R.-Z., M., & Saad, Y. (2018). Shanks sequence transformations and
590 Anderson acceleration. *SIAM Rev.*, *60*, 646-669. doi: 10.1137/17M1120725
- 591 Brown, J., & Brune, P. R. (2013). Low-rank quasi-newton updates for robust jaco-
592 bian lagging in newton methods. In *International Conference on Mathematics
593 and Computational Methods Applied to Nuclear Science & Engineering (M&C
594 2013)* (p. 2554-2565). American Nuclear Society.
- 595 Brown, P. N., & Hindmarsh, A. C. (1986). Matrix-free methods for stiff systems of
596 ODE's. *SIAM J. Numer. Anal.*, *23*, 610-638.
- 597 Brown, P. N., & Saad, Y. (1990). Hybrid Krylov methods for nonlinear systems of
598 equations. *SIAM J. Sci. Stat. Comput.*, *11*, 450-481.
- 599 Brune, P. R., Knepley, M. G., Smith, B. F., & Tu, X. (2015). Composing scalable
600 nonlinear algebraic solvers. *SIAM Rev.*, *57*, 535-565. doi: 10.1137/130936725
- 601 Cassar, N., Nicholson, D., Khatiwala, S., & Cliff, E. (2021). Decomposing the oxy-
602 gen signal in the ocean interior: Beyond decomposing organic matter. *Geophys.
603 Res. Lett.*, *48*. doi: 10.1029/2021GL092621
- 604 Chan, T. F., & Jackson, K. R. (1984). Nonlinearly preconditioned Krylov subspace
605 methods for discrete Newton algorithms. *SIAM J. Sci. Stat. Comput.*, *5*, 533-
606 542.
- 607 Chen, K., & Vuik, C. (2022a). Composite anderson acceleration method with dy-
608 namic window-sizes and optimized damping.
609 doi: 10.48550/arXiv.2203.14627
- 610 Chen, K., & Vuik, C. (2022b). Non-stationary anderson acceleration with optimized
611 damping.
612 doi: 10.48550/arXiv.2202.05295
- 613 de Souza, G. F., Khatiwala, S., P.Hain, M., H.Little, S., & Vance, D. (2018). On the
614 origin of the marine zinc-silicon correlation. *Earth and Planet. Sci. Lett.*, *492*,
615 22-34. doi: 10.1016/j.epsl.2018.03.050
- 616 Dennis, J. E., & Schnabel, R. B. (1996). *Numerical Methods for Unconstrained Op-
617 timization and Nonlinear Equations*. Society for Industrial and Applied Mathe-
618 matics.
- 619 DeVries, T., & Primeau, F. (2011). Dynamically and observationally constrained
620 estimates of water-mass distributions and ages in the global ocean. *J. Phys.
621 Oceanogr.*, *41*, 2381-2401. doi: 10.1175/JPO-D-10-05011.1
- 622 England, M. H. (1995). The age of water and ventilation timescales in a global
623 ocean model. *J. Phys. Oceanogr.*, *25*, 2756-2777.
- 624 Eyert, V. (1996). A comparative study on methods for convergence acceleration of
625 iterative vector sequences. *J. Comput. Phys.*, *124*, 271-285.
- 626 Fang, H., & Saad, Y. (2009). Two classes of multiseant methods for nonlinear ac-
627 celeration. *Numer. Linear Algebra Appl.*, *16*, 197-221. doi: doi:10.1002/nla
628 .617
- 629 Forget, G., Campin, J. M., Heimbach, P., Hill, C. N., Ponte, R. M., & Wunsch, C.
630 (2015). Ecco version 4: An integrated framework for non-linear inverse model-
631 ing and global ocean state estimation. *Geosci. Model Dev.*, *8*, 3071-3104. doi:
632 10.5194/gmd-8-3071-2015
- 633 Fu, A., Zhang, J., & Boyd, S. (2020). Anderson accelerated Douglas-Rachford split-
634 ting. *SIAM J. Sci. Comput.*, *42*, 3560-3583. doi: 10.1137/19M1290097
- 635 Garcia, H. E., Weathers, K., Paver, C. R., Smolyar, I., Boyer, T. P., Locarnini,
636 R. A., ... Reagan, J. R. (2018). *World Ocean Atlas 2018, Volume 3: Dissolved
637 Oxygen, Apparent Oxygen Utilization, and Oxygen Saturation* (Tech. Rep.).
- 638 Gear, C. W., & Saad, Y. (1983). Iterative solution of linear equations in ODE codes.
639 *SIAM J. Sci. Stat. Comput.*, *4*, 583-601.
- 640 Gebbie, G., & Huybers, P. (2012). The mean age of ocean waters inferred from
641 radiocarbon observations: sensitivity to surface sources and accounting for
642 mixing histories. *J. Atmos. Sci.*, *42*, 291-305.

- Goldstein, S. L., & Hemming, S. R. (2003). Long-lived isotopic tracers in oceanography, paleoceanography, and ice-sheet dynamics. In *Treatise on geochemistry* (Vol. 6, p. 453-489). Elsevier Ltd.
- Graven, H. D., Gruber, N., Key, R., Khatiwala, S., & Giraud, X. (2012). Changing controls on oceanic radiocarbon: New insights on shallow-to-deep ocean exchange and anthropogenic CO₂ uptake. *J. Geophys. Res.*, *117*. doi: 10.1029/2012JC008074
- Gu, S., & Liu, Z. (2017). ²³¹pa and ²³⁰th in the ocean model of the community earth system model (cesm1.3). *Geosci. Model Dev.*, *10*, 4723-4742. doi: 10.5194/gmd-10-4723-2017
- Gurvan Madec and NEMO System Team. (n.d.). Nemo ocean engine (Computer software manual No. 27). Zenodo. doi: 10.5281/zenodo.1464816
- Haine, T. W. M., & Hall, T. M. (2002). A generalized transport theory: water-mass composition and age. *J. Phys. Oceanogr.*, *32*, 1932-1946.
- Hamme, R. C., Nicholson, D. P., Jenkins, W. J., & Emerson, S. R. (2019). Using noble gases to assess the ocean's carbon pumps. *Ann. Rev. Mar. Sci.*, *11*, 75-103. doi: 10.1146/annurev-marine-121916-063604
- Hayes, C. T., Anderson, R. F., Fleisher, M. Q., Huang, K., Robinson, L. F., Lu, Y., ... Moran, S. (2015). ²³⁰th and ²³¹pa on geotraces ga03, the u.s. geotraces north atlantic transect, and implications for modern and paleoceanographic chemical fluxes. *Deep-Sea Res. II*, *116*, 29-41. doi: 10.1016/j.dsr2.2014.07.007
- Henderson, G. M., & Anderson, R. F. (2003). The u-series toolbox for paleoceanography. *Rev. Mineral. Geochem.*, *52*, 493-531. doi: 10.2113/0520493
- Holzer, M., & Hall, T. M. (2000). Transit-time and tracer-age distributions in geophysical flows. *J. Atmos. Sci.*, *57*, 3539-3558.
- Holzer, M., & Primeau, F. (2006). The diffusive ocean conveyor. *Geophys. Res. Lett.*, *33* (doi:10.1029/2006GL026232).
- Ito, T., & Follows, M. J. (2005). Preformed phosphate, soft tissue pump and atmospheric CO₂. *J. Mar. Res.*, *63*(4), 813-839.
- Ito, T., Follows, M. J., & Boyle, E. A. (2004). Is AOU a good measure of respiration in the oceans? *Geophys. Res. Lett.*, *31*(17), L17305. doi: 10.1029/2004GL020900
- Kelley, C. T. (1995). *Iterative Methods for Linear and Nonlinear Equations*. Society for Industrial and Applied Mathematics.
- Khatiwala, S. (2007). A computational framework for simulation of biogeochemical tracers in the ocean. *Glob. Biogeochem. Cy.*, *21*, GB3001. doi: 10.1029/2007GB002923
- Khatiwala, S. (2008). Fast spin up of ocean biogeochemical models using matrix-free Newton-Krylov. *Ocean Modell.*, *23*, 121-129. doi: 10.1016/j.ocemod.2008.05.002
- Khatiwala, S. (2018). *Transport Matrix Method software for ocean biogeochemical simulations* [Software]. doi: 10.5281/zenodo.1246300
- Khatiwala, S. (2022). *Fast spin-up of geochemical tracers in ocean and climate models* [Data and Software]. doi: 10.5281/zenodo.7164028
- Khatiwala, S., Primeau, F., & Hall, T. (2009). Reconstruction of the history of anthropogenic CO₂ concentrations in the ocean. *Nature*, *462*, 346-349. doi: 10.1038/nature08526
- Khatiwala, S., Primeau, F., & Holzer, M. (2012). Ventilation of the deep ocean constrained with tracer observations and implications for radiocarbon estimates of ideal mean age. *Earth and Planet. Sci. Lett.*, *325-326*, 116-125. doi: 10.1016/j.epsl.2012.01.038
- Khatiwala, S., Schmittner, A., & Muglia, J. (2019). Air-sea disequilibrium enhances ocean carbon storage during glacial periods. *Sci. Adv.*, *5*. doi: 10.1126/sciadv.aaw4981

- 697 Khatiwala, S., Visbeck, M., & Cane, M. (2005). Accelerated simulation of passive
698 tracers in ocean circulation models. *Ocean Modell.*, *9*, 51-69. doi: 10.1016/j.
699 .ocemod.2004.04.002
- 700 Khatiwala, S., Visbeck, M., & Schlosser, P. (2001). Age tracers in an ocean GCM.
701 *Deep-Sea Res. I*, *48*, 1423-1441.
- 702 Kimmritz, M., Losch, M., & Danilov, S. (2017). A comparison of viscous-plastic sea
703 ice solvers with and without replacement pressure. *Ocean Modell.*, *115*, 59-69.
704 doi: 10.1016/j.ocemod.2017.05.006
- 705 Knoll, D. A., & Keyes, D. E. (2004). Jacobian-free Newton-Krylov methods: a sur-
706 vey of approaches and applications. *J. Comput. Phys.*, *193*, 357-397.
- 707 Large, W., & Yeager, S. (2004). *Diurnal to decadal global forcing for ocean and*
708 *sea-ice models: the datasets and flux climatologies* (Tech. Rep. Nos. NCAR
709 Technical Note: NCAR/TN-460+STR). CGD Division of the National Centre
710 for Atmospheric Research, Boulder, CO, USA.
- 711 Lauderdale, J. M., Garabato, A. C. N., Oliver, K. I. C., Follows, M. J., & Williams,
712 R. G. (2013). Wind-driven changes in Southern Ocean residual circulation,
713 ocean carbon reservoirs and atmospheric CO₂. *Clim. Dyn.*, *41*(7-8). doi:
714 10.1007/S00382-012-1650-3
- 715 Lemieux, J., Knoll, D. A., Tremblay, B., Holland, D. M., & Losch, M. (2012). A
716 comparison of the jacobian-free newton-krylov method and the evp model
717 for solving the with a viscous-plastic formulation: A serial algorithm study
718 sea ice momentum equation. *J. Comput. Phys.*, *231*, 5926-5944. doi:
719 10.1016/j.jcp.2012.05.024
- 720 Li, X., & Primeau, F. (2008). A fast Newton-Krylov solver for seasonally varying
721 global ocean biogeochemistry models. *Ocean Modell.* (*in press*).
- 722 Lindsay, K. (2017). A newton-krylov solver for fast spin-up of online ocean tracers.
723 *Ocean Modell.*, *109*, 33-43. doi: 10.1016/j.ocemod.2016.12.001
- 724 Losch, M., Fuchs, A., Lemieux, J., & Vanselow, A. (2014). A parallel jacobian-
725 free newton-krylov solver for a coupled sea ice-ocean model. *J. Comput. Phys.*,
726 *257*, 901-911. doi: 10.1016/j.jcp.2013.09.026
- 727 Marshall, J., Adcroft, A., Hill, C., Perelman, L., & Heisey, C. (1997). A finite-
728 volume, incompressible navier-stokes model for studies of the ocean on parallel
729 computers. *J. Geophys. Res.*, *102*, 5733-5752.
- 730 Martin, J. H., Knauer, G. A., Karl, D. M., & Broenkow, W. W. (1987). VERTEX:
731 carbon cycling in the northeast Pacific. *Deep-Sea Res.*, *34*, 267-285. doi: 10
732 .1016/0198-0149(87)90086-0
- 733 Matsumoto, K., Sarmiento, J. L., Key, R. M., Aumont, O., Bullister, J. L., Caldeira,
734 K., ... Orr, J. C. (2004). Evaluation of ocean carbon cycle models with
735 data-based metrics. *Geophys. Res. Lett.*, *31* (doi:10.1029/2003GL018970).
- 736 McManus, J. F., François, R., Gherardl, J. M., Kelgwin, L., & Drown-Leger, S.
737 (2004). Collapse and rapid resumption of Atlantic meridional circulation linked
738 to deglacial climate changes. *Nature*, *428*, 834-837. doi: 10.1038/nature02494
- 739 Merlis, T., & Khatiwala, S. (2008). Fast dynamical spin-up of ocean general cir-
740 culation models using Newton-Krylov methods. *Ocean Modell.*, *21*, 97-105. doi:
741 10.1016/j.ocemod.2007.12.001
- 742 Missiaen, L., Bouttes, N., Roche, D. M., Dutay, J., Quiquet, A., Waelbroeck, C.,
743 ... Peterschmitt, J. (2020). Carbon isotopes and pa/th response to forced
744 circulation changes: a model perspective. *Clim. Past*, *16*, 867-883. doi:
745 10.5194/cp-16-867-2020
- 746 Morel, F. M. M., Milligan, A. J., & Saito, M. A. (2014). Marine Bioinorganic
747 Chemistry: The Role of Trace Metals in the Oceanic Cycles of Major Nutri-
748 ents. In H. D. Holland & K. K. Turekian (Eds.), *Treatise on geochemistry*
749 (*second edition*) (Second Edition ed., p. 123-150). Oxford: Elsevier. doi:
750 10.1016/B978-0-08-095975-7.00605-7

- Muglia, J., & Schmittner, A. (2015). Glacial Atlantic overturning increased by wind stress in climate models. *Geophys. Res. Lett.*, *42*, 9862-9868. doi: 10.1002/2015GL064583
- Muglia, J., Skinner, L. C., & Schmittner, A. (2018). Weak overturning circulation and high Southern Ocean nutrient utilization maximized glacial ocean carbon. *Earth and Planet. Sci. Lett.*, *496*, 47-56. doi: 10.1016/j.epsl.2018.05.038
- Najjar, R. G., Jin, X., Louanchi, F., Aumont, O., Caldeira, K., Doney, S. C., ... Yool, A. (2007). Impact of circulation on export production, dissolved organic matter, and dissolved oxygen in the ocean: Results from Phase II of the Ocean Carbon-cycle Model Intercomparison Project (OCMIP-2). *Glob. Biogeochem. Cy.*, *21*(GB3007). doi: 10.1029/2006GB002857
- Nicholson, D. P., Khatiwala, S., & Heimbach, P. (2016). Noble gas tracers of ventilation during deep water formation in the Weddell Sea. In *The 7th international symposium on gas transfer at water surfaces*. IOP Conference Series: Earth and Environmental Science. doi: 10.1088/1755-1315/35/1/012019
- Nocedal, J. (1980). Updating quasi-newton matrices with limited storage. *Mathematics of Computation*, *35*, 773-782. doi: 10.1090/S0025-5718-1980-0572855-7
- Nocedal, J., & Wright, S. (2006). *Numerical Optimization*. Springer.
- Orr, J. C., Najjar, R., Sabine, C. L., & Joos, F. (1999). Abiotic-HOWTO. In *Internal OCMIP Report (www.ipsl.jussieu.fr/OCMIP)* (p. 15 pp.). LSCE/CEA Saclay, Gif-sur-Yvette, France.
- Orr, J. C., Najjar, R. G., Aumont, O., Bopp, L., Bullister, J. L., Danabasoglu, G., ... Yool, A. (2017). Biogeochemical protocols and diagnostics for the CMIP6 Ocean Model Intercomparison Project (OMIP). *Geosci. Model Dev.*, *10*. doi: 10.5194/gmd-10-2169-2017
- Primeau, F. (2005). Characterizing transport between the surface mixed layer and the ocean interior with a forward and adjoint global ocean transport model. *J. Phys. Oceanogr.*, *35*, 545-564.
- Rempfer, J., Stocker, T. F., Joos, F., Lippold, J., & Jaccard, S. L. (2017). New insights into cycling of ^{231}Pa and ^{230}Th in the atlantic ocean. *Earth and Planet. Sci. Lett.*, *468*, 27-37. doi: 10.1016/j.epsl.2017.03.027
- Rutberg, R. L., Hemming, S. R., & Goldstein, S. L. (2000). Reduced North Atlantic Deep Water flux to the glacial Southern Ocean inferred from neodymium isotope ratios. *Nature*, *405*, 935-938.
- Saad, Y. (2003). *Iterative Methods for Sparse Linear Systems*. Society for Industrial and Applied Mathematics.
- Sarnthein, M., Schneider, B., & Grootes, P. M. (2013). Peak glacial ^{14}C ventilation ages suggest major draw-down of carbon into the abyssal ocean. *Clim. Past*, *9*(6), 2595-2614. doi: 10.5194/cp-9-2595-2013
- Sasaki, Y., Kobayashi, H., & Oka, A. (2022). Global simulation of dissolved ^{231}Pa and ^{230}Th in the ocean and the sedimentary $^{231}\text{Pa}/^{230}\text{Th}$ ratios with the ocean general circulation model coco ver4.0. *Geosci. Model Dev.*, *15*, 2013-2033. doi: 10.5194/gmd-15-2013-2022
- Schlitzer, R. (2007). Assimilation of radiocarbon and chlorofluorocarbon data to constrain deep and bottom water transports in the world ocean. *J. Phys. Oceanogr.*, *37*(2), 259-276.
- Schmittner, A., & Somes, C. J. (2016). Complementary constraints from carbon (^{13}C) and nitrogen (^{15}N) isotopes on the glacial ocean's soft-tissue biological pump. *Paleoceanography*, *31*, 669-693. doi: 10.1002/2015PA002905
- Schmittner, A., Somes, C. J., & Muglia, J. (2022). *OSU-CEOAS-Schmittner/UVic2.9* [Software]. doi: 10.5281/zenodo.7159021
- Siddall, M., Henderson, G. M., Edwards, N. R., Frank, M., Mu, S. A., Stocker, T. F., & Joos, F. (2005). $^{231}\text{Pa}/^{230}\text{Th}$ fractionation by ocean transport, biogenic particle flux and particle type. *Earth and Planet. Sci. Lett.*, *237*, 135-155. doi: 10.1016/j.epsl.2005.05.031

- 806 Skinner, L. C., Primeau, F., Freeman, E., de la Fuente, M., Goodwin, P. A.,
807 Gottschalk, J., ... Scrivner, A. E. (2017). Radiocarbon constraints on the
808 glacial ocean circulation and its impact on atmospheric CO₂. *Nat. Comms.*, 8,
809 16010. doi: 10.1038/ncomms16010
- 810 Smith, D. A., Ford, W. F., & Sidi, A. (1987). Extrapolation methods for vector se-
811 quences. *SIAM Rev.*, 29, 199-233. doi: 10.1137/1029042
- 812 Stammer, D., Ueyoshi, K., Köhl, A., Large, W. G., Josey, S. A., & Wunsch, C.
813 (2004). Estimating air-sea fluxes of heat, freshwater, and momentum through
814 global ocean data assimilation. *J. Geophys. Res.*, 109 (C05023). doi:
815 10.1029/2003JC002082
- 816 Tang, W., & Daoutidis, P. (2022). Fast and stable nonconvex constrained dis-
817 tributed optimization: the ellada algorithm. *Optim. Eng.*, 23, 259-301. doi:
818 10.1007/s11081-020-09585-w
- 819 Thiele, G., & Sarmiento, J. L. (1990). Tracer dating and ocean ventilation. *J. Geo-*
820 *phys. Res.*, 95, 9377-9391.
- 821 Toggweiler, J. R., Dixon, K., & Bryan, K. (1989). Simulations of radiocarbon in a
822 coarse resolution world ocean model, 1. Steady state prebomb distributions. *J.*
823 *Geophys. Res.*, 94, 8217-8242.
- 824 van Hulten, M., Dutay, J., & Roy-Barman, M. (2018). A global scavenging and
825 circulation ocean model of thorium-230 and protactinium-231 with improved
826 particle dynamics (nemo-prothorp 0.1). *Geosci. Model Dev.*, 11, 3537-3556.
827 doi: 10.5194/gmd-11-3537-2018
- 828 Vance, D., Little, S. H., de Souza, G. F., Khatiwala, S., Lohan, M. C., & Middag, R.
829 (2017). Silicon and zinc biogeochemical cycles coupled through the Southern
830 Ocean. *Nature Geosci.*, 10, 202-206. doi: 10.1038/ngeo2890
- 831 Walker, H. (2010). *Anderson Acceleration: Algorithms and Implementations* (Tech.
832 Rep. No. MS-1-13-46). Worcester Polytechnic Institute. doi: 10.1145/321296
833 .321305
- 834 Walker, H., & Ni, P. (2011). Anderson acceleration for fixed-point iterations. *SIAM*
835 *J. Numer. Anal.*, 49, 1715-1735. doi: 10.1137/10078356X
- 836 Walker, H. F., Woodward, C. S., & Yang, U. M. (2010). An accelerated fixed-point
837 iteration for solution of variably saturated flow. In J. Carrera (Ed.), *XVIII In-*
838 *ternational Conference on Water Resources CMWR 2010*.
- 839 Waugh, D. W., Hall, T. M., & Haine, T. W. N. (2003). Relationships among tracer
840 ages. *J. Geophys. Res.*, 108. doi: 10.1029/2002JC001325
- 841 Weaver, A., Eby, M., Wiebe, E., Bitz, C., Duffy, P., Ewen, T., ... Yoshimori, M.
842 (2001, DEC). The UVic Earth System Climate Model: Model description, cli-
843 matology, and applications to past, present and future climates. *Atmos.-Ocean*,
844 39(4), 361-428.
- 845 Williams, R. G., & Follows, M. J. (2011). *Ocean Dynamics and the Carbon Cycle:*
846 *Principles and Mechanisms*. Cambridge University Press.
- 847 Wunsch, C., & Heimbach, P. (2007). Practical global oceanic state estimation. *Phys-*
848 *ica D*, 230, 197-208.
- 849 Wunsch, C., & Heimbach, P. (2008). How long to oceanic tracer and proxy equilib-
850 rium? *Quat. Sci. Rev.*, 27, 637-651.
- 851 Yu, E., Francois, R., & Bacon, M. P. (1996). Similar rates of modern and last-glacial
852 ocean thermohaline circulation inferred from radiochemical data. *Nature*, 379,
853 689-694. doi: 10.1038/379689a0
- 854 Zanna, L., Khatiwala, S., Gregory, J. M., Ison, J., & Heimbach, P. (2019). Global
855 reconstruction of historical ocean heat storage and transport. *Proc. Natl. Acad.*
856 *Sci.*, 116 (4), 1126-1131. doi: 10.1073/pnas.1808838115
- 857 Zhang, J., O'Donoghue, B., & Boyd, S. (2020). Globally convergent Type-I Ander-
858 son acceleration for nonsmooth fixed-point iterations. *SIAM J. Optimization*,
859 30, 3170-3197. doi: 10.1137/18M1232772