Parameter Estimation in Marine Ecosystem Models

Thomas Slawig, Claudia Kratzenstein, Jaroslaw Piwonski, Malte Prieß, Christina Roschat

Kiel University, Cluster of Excellence "The Future Ocean"

co-operation with groups of:
A. Griewank (HU Berlin), N. Gauger (RWTH Aachen), A. Oschlies (GEOMAR Kiel), S. Koziel (Reykjavik).

Sept. 12, 2014, Dortmund
1 Motivation
2 Mathematical Aims and Challenges
3 Mathematical Analysis of the PDE
4 Numerical Solution
   ■ Transport Matrix Method
   ■ Computation of a steady annual Cycle
   ■ Parallelization
5 Optimization Methods
   ■ Surrogate-based Optimization
   ■ Oneshot Optimization
6 Numerical Results
7 Summary
Motivation: Modeling the global carbon cycle

- **Ocean is a carbon sink**: $CO_2$ uptake via the sea surface
- **Transport and distribution** via ocean currents
- **Reactions** determined by ocean biogeochemical dynamics

Simulated concentration of nutrients (phosphate, $PO_4^{3-}$) at the surface layer in mmol P/m$^3$. The longitudinal and latitudinal resolution is at 1.0°.
Motivation: Model calibration and assessment

**Fit model output to observed data**

Optimize parameters \( u \) in the underlying model

Formulation of the mathematical problem:

\( y \in Y \) (either function space or, discretized, \( \mathbb{R}^n \)), \( u \in \mathbb{R}^m \), \( m \ll n \)

\[
\min J(y, u) = \frac{1}{2} \| y - y_{data} \|^2 + \frac{\alpha}{2} \| u - u_{est} \|^2
\]

s.t. model equations are fulfilled: \( e(y, u) = 0 \).
1 Motivation

2 Mathematical Aims and Challenges

3 Mathematical Analysis of the PDE

4 Numerical Solution
   ■ Transport Matrix Method
   ■ Computation of a steady annual Cycle
   ■ Parallelization

5 Optimization Methods
   ■ Surrogate-based Optimization
   ■ Oneshot Optimization

6 Numerical Results

7 Summary
Mathematical Challenges

- Mathematical analysis of the PDEs:
  - existence + uniqueness
  - non-zero periodic solutions (steady annual cycle).

- Mathematical analysis of the parameter identification problem:
  - identifiability of parameters
  - injectivity of mapping: $u \mapsto y$.

Model evaluation is expensive + optimization needs many model runs:

- Efficient optimization methods
- Acceleration of the numerical simulation itself:
  - for a single model year
  - for the “spin-up“ (computation of a steady periodic solution).
Coupled system

Ocean circulation + Biogeochemistry for nutrients, plankton etc.

Navier-Stokes equations + System of transport equations with nonlinear coupling
Online vs. offline simulation

Ocean circulation fields

velocity, turbulent mixing
temperature, salinity

Tracer distribution

“tracers“:
nutrients, plankton, ...

Offline simulations:
- circulation fields pre-computed by ocean model
- used as forcing data for tracer transport model
Tracer transport model

- System of transport equations for biogeochemical tracers \( y = (y_i)_j \):

\[
\partial_t y_j = \text{div}(\kappa \nabla y_j) - \text{div}(\vec{v} y_j) + q_j(t, y, u), \quad j = 1, \ldots, s
\]

+ + initial and boundary conditions.

- Pre-computed, “climatological“ (annually periodic) forcing data

\[
\kappa(t + T) = \kappa(t), \quad \vec{v}(t + T) = \vec{v}(t), \quad t \in [0, T]
\]

coupling terms

\[
q_i(t + T, \cdot, \cdot) = q_i(t, \cdot, \cdot)
\]

and boundary conditions.

- Desired solution is a **steady annual cycle**:

\[
y(t + T) = y(t), \quad t \in [0, T].
\]
(Simplest) example: N-DOP model by Parekh et al.

■ Transport equations:
\[
\partial_t y_j = \text{div}(\kappa \nabla y_j) + \text{div}(\nu y_j) + q_j(y, u) \quad \text{in } \Omega \times (0, T), j = 1, 2.
\]

■ Coupling terms:
\[
q_1 = \lambda y_2 + \begin{cases} 
- F(y_1), & \Omega_1 \text{ (upper layer)} \\
(1 - \nu) E(y_1), & \Omega_2 \text{ (lower layer)}
\end{cases}
\]
\[
q_2 = -\lambda y_2 + \begin{cases} 
\nu F(y_1), & \Omega_1 \\
0, & \Omega_2
\end{cases}
\]

■ Typical nonlinearity and non-local term:
\[
F(y_1) = \alpha \frac{y_1}{y_1 + K},
\]
\[
E(y_1) = c \left[ \frac{z}{\text{depth}} \right]^m \int_0^{\text{depth}} (1 - \nu) F(y_1) dz
\]

■ Nonlinear and nonlocal boundary condition (for mass conservation):
\[
\partial_n y_j = b_j(y, u) \quad \text{on } \Gamma \times (0, T).
\]
Obvious problems for parameter identifiability

- Typical nonlinear functions:

\[
F(y_1) = \frac{\alpha y_1}{y_1 + K}, \quad F(y_1) = \frac{\alpha Ky_1^2}{\alpha + Ky_1^2}
\]

- Both have two parameters, but are almost linear in certain ranges of \(y_1\).
Mathematical Formulation of the PDE

Classical form of an ecosystem model equation of \( N\text{-DOP} \) type

\[
\begin{align*}
\partial_t y_1 + \text{div}(\vec{v}(t)y_1) - \text{div}(\kappa \nabla y_1) - \lambda y_2 + d_1(y) &= 0 \quad \text{in } \Omega \times [0, T] \\
\partial_t y_2 + \text{div}(\vec{v}(t)y_2) - \text{div}(\kappa \nabla y_2) + \lambda y_2 + d_2(y) &= 0 \quad \text{in } \Omega \times [0, T] \\
\nabla y_j \cdot (\kappa \vec{n}) + b_j(y) &= 0 \quad \text{on } \Gamma \times [0, T], \ j = 1, 2
\end{align*}
\]

- \( \Omega \subset \mathbb{R}^n \) open, bounded, \( \Gamma := \partial \Omega \)
- \( \vec{v} \in L^\infty(0, T; H^1(\Omega))^n \) with \( \text{div}(\vec{v}(t)) = 0 \) in \( \Omega \) and \( \vec{v}(t) \cdot \vec{n} = 0 \) on \( \Gamma \)
- \( \kappa \in L^\infty(\Omega \times [0, T]) \) with \( \kappa_{\text{min}} := \text{essinf}_{(x,t)} \kappa(x, t) > 0 \)

**Definition**

- \( y = (y_1, y_2) \) is called periodic if \( y_j(0) = y_j(T) \) for \( j = 1, 2 \).
- Let \( C > 0 \). A solution \( y = (y_1, y_2) \) has a constant mass \( C \) if

\[
\text{mass}(y(t)) := \int_{\Omega} (y_1(t) + y_2(t)) \, dx = C \quad \text{for all } t \in [0, T].
\]
Weak formulation

Space of weak solutions:

\[ W(0, T; H^1(\Omega)) := \{ v \in L^2(0, T; H^1(\Omega)); v' \in L^2(0, T; H^1(\Omega)^*) \} \]

Weak formulation of ecosystem model equations

\[
\int_0^T \left\{ \langle y'_1, w_1 \rangle + B(y_1, w_1; t) + (\lambda y_2 + d_1(y), w_1)_{\Omega} + (b_1(y), w_1)_\Gamma \right\} dt = 0
\]

\[
\int_0^T \left\{ \langle y'_2, w_2 \rangle + B(y_2, w_2; t) + (\lambda y_2 + d_2(y), w_2)_{\Omega} + (b_2(y), w_2)_\Gamma \right\} dt = 0
\]

for all test functions \((w_1, w_2) \in L^2(0, T; H^1(\Omega))^2\)

Definition

Time dependent bilinear form

\[ B(a, b; t) := \int_{\Omega} (\kappa(t) \nabla a \cdot \nabla b) dx + \int_{\Omega} \text{div}(\vec{\nu}(t)a)b dx \quad \forall a, b \in L^2(0, T; H^1(\Omega)). \]
Weak formulation in the dual space

More general view on elements of weak formulation:

\[ B : L^2(0, T; H^1(\Omega)) \rightarrow L^2(0, T; H^1(\Omega)^*); \quad \langle B(y), v \rangle := \int_0^T B(y, v; t)dt, \]

\[ F_j : L^2(0, T; H^1(\Omega)) \rightarrow L^2(0, T; H^1(\Omega)^*); \]

\[ \langle F_j(y), v \rangle := \int_0^T \{- (d_j(y), v)_\Omega - (b_j(y), v)_\Gamma \} dt \]

for \( j = 1, 2 \) and \( y, v \in L^2(0, T; H^1(\Omega)) \)

Problem in the dual space

\[ y_1' + B(y_1) - \lambda y_2 = F_1(y) \]
\[ y_2' + B(y_2) + \lambda y_2 = F_2(y) \]
\[ y(0) = y(T) \]
Existence theorem

**Theorem (C. Roschat 2014)**

Let \( C, \lambda > 0 \). Under the assumptions

- \( d, b \) are continuous
- \( \sum_{j=1}^{2} \left( \int_{\Omega} d_{j}(y, x, t) dx + \int_{\Gamma} b_{j}(y, s, t) ds \right) = 0 \) for almost all \( t \) and all \( y \)
- \( \max\{|d_{j}(y, x, t)|, |b_{j}(y, x, t)|\} \leq M \) for all \( y, j = 1, 2 \) and almost all \( (x, t) \)

there is at least one periodic weak solution \( y \in W(0, T; H^{1}(\Omega))^{2} \) of the ecosystem model equation with

\[
\text{mass}(y(t)) = C \quad \text{for all } t \in [0, T].
\]

**Corollary**

The N-DOP model by Parekh et al., extended by suitable boundary conditions, has a periodic solution.
Existence theorem of Gajewski et al., 1974

Let \( V \subset H \subset V^* \) be an evolution triple, \( X := L^2(0, T; V) \). If \( A : X \to X^* \) is a continuous, monotone and coercive operator, the problem

\[
  u' + Au = f, \quad u(0) = u(T),
\]

has a solution \( u \in W(0, T; V) \hookrightarrow C([0, T]; H) \) for every \( f \in X^* \).

Definitions:

- A monotone : \( \iff \langle Au - Av, u - v \rangle \geq 0 \) for all \( u, v \in X \)

- A coercive : \( \iff \frac{\langle Au, u \rangle}{\|u\|_X} \to \infty \) if \( \|u\|_X \to \infty \)

Problem: Coercivity usually not fulfilled by ecosystem model equations.
1 For every $z \in L^2(0, T; L^2(\Omega))^2$, solve

\[
\begin{align*}
    y_1' + B(y_1) - \lambda y_2 &= F_1(z) \\
    y_2' + B(y_2) + \lambda y_2 &= F_2(z) \\
    y(0) &= y(T) \\
    \text{mass}(y(t)) &= C.
\end{align*}
\]

2 Find a fixed point of the map $z \mapsto y$ where $y$ is the solution of the linearized problem.
Step 1: Second equation

- Second equation of the linearized system can be solved independently:
  \[ y_2' + B(y_2) + \lambda y_2 = F_2(z) \]
  \[ y_2(0) = y_2(T) \]

- Gajewski Theorem: Existence of unique periodic \( y_2 \in W(0, T; H^1(\Omega)) \).
Step 2: Equation for the sum

- Equation for the sum $S := y_1 + y_2$:
  
  $$
  S' + B(S) = F_1(z) + F_2(z) \\
  S(0) = S(T) \\
  \text{mass}(S(t)) = C.
  $$

- $B$ is not coercive on the usual solution space $L^2(0, T; H^1(\Omega))$.

- Use $V := \{y \in H^1(\Omega) : \int_\Omega y dx = 0\}$ with $y \mapsto \|\nabla y\|_{L^2(\Omega)^n}$.

- Restricted $B$ is coercive in the space $L^2(0, T; V)$.

- Gajewski: Existence of periodic solution $S \in L^2(0, T; V)$ with distributional derivative $S' \in L^2(0, T; V^*)$.

- $S'$ has to be (and can be) extended from $\in L^2(0, T; V^*)$ to $L^2(0, T; H^1(\Omega)^*)$.

- Problem: $\text{mass}(S(t)) = 0$. Using a special transient solution, a periodic $S_C$ with $\text{mass}(S_C(t)) = C$ for all $t$ can be constructed.
Step 3: Definition of $y_1$

$y_1 := S_C - y_2$ solves the first equation.

Thus:

Result of Steps 1-3

For every $z \in L^2(0, T; L^2(\Omega))^2$ there is a periodic solution $(y_1, y_2) \in W(0, T; H^1(\Omega))^2$ with $\text{mass}(y_1(t), y_2(t)) = C$ for all $t \in [0, T]$. 
Step 4: The fixed point problem

Define

\[ A : L^2(0, T; L^2(\Omega))^2 \rightarrow L^2(0, T; L^2(\Omega))^2, \quad z \mapsto y(z), \]

where \( y(z) \) is the weak solution of the equation linearized by inserting \( z \).

- \( A \) well-defined according to Steps 1-3
- \( y \) fixed point of \( A \) \( \iff \) \( y \) solves the original nonlinear equations.

**Schauder Fixed Point Theorem**

Let \( M \) be a nonempty, closed, bounded, convex subset of a Banach space \( X \), and \( A : M \rightarrow M \) a compact operator. Then \( A \) has a fixed point.

Properties of \( A \) can be shown by boundedness assumption for \( d_j \) and \( b_j \) for

\[ M := \{ y \in L^2(0, T; L^2(\Omega))^2; \| y \|_{L^2(0,T;L^2(\Omega))^2} \leq K \}. \]
For every $C > 0$, the existence of a weak periodic solution of the ecosystem model

$$
\begin{align*}
\partial_t y_1 + \text{div}(\vec{v}(t)y_1) - \text{div}(\kappa \nabla y_1) - \lambda y_2 + d_1(y) &= 0 \quad \text{in } \Omega \times [0, T] \\
\partial_t y_2 + \text{div}(\vec{v}(t)y_2) - \text{div}(\kappa \nabla y_2) + \lambda y_2 + d_2(y) &= 0 \quad \text{in } \Omega \times [0, T] \\
\nabla y_j \cdot (\kappa n) + b_j(y) &= 0 \quad \text{on } \Gamma \times [0, T], j = 1, 2
\end{align*}
$$

with $\int_\Omega (y_1(t) + y_2(t)) \, dx = C$ for all $t \in [0, T]$ was proved.

- **N-DOP** model by Parekh et al. fulfills the assumptions

- Proof adapted to **N-DOP** type models, thus extension to other problems difficult

- Identifiability of parameters: can be shown only for some (work in progress).
1 Motivation

2 Mathematical Aims and Challenges

3 Mathematical Analysis of the PDE

4 Numerical Solution
   - Transport Matrix Method
   - Computation of a steady annual Cycle
   - Parallelization

5 Optimization Methods
   - Surrogate-based Optimization
   - Oneshot Optimization

6 Numerical Results

7 Summary
Off-line simulation with transport matrices

- Instead of directly using the pre-computed forcing data ...
  \[
  \frac{\partial y_j}{\partial t} = \text{div}(\kappa \nabla y_j) - \text{div}(\vec{v}y_j) + q_j(t, y, u), \quad j = 1, \ldots, s
  \]

- ... we use discretized transport operators (Transport Matrix Method [Khatiwala et al., 2005]) in the \(i\)-th time step:
  \[
  y_{i+1} = A_{\text{imp},i} (A_{\text{exp},i} y_i + \Delta t q_i(y_i, u)), \quad i = 1, \ldots
  \]

  \(y_i\): vector of all tracers at all grid points at time step \(i\).

- Sparse implicit and explicit transport matrices:
  \[
  A_{\text{imp},i} = (I - \Delta t L_{\text{imp},i})^{-1} \\
  A_{\text{exp},i} = (I + \Delta t L_{\text{exp},i})
  \]

- Example resolution: 2.8125\(^\circ\), i.e., 128 \(\times\) 64 surface grid, 15 vertical layers, 2880 time steps, matrix dimension per tracer \(\approx 50'000\).
Flexible biogeochemical model interface

- Main aim: flexibility wrt. biogeochemical models for model comparison.
- Designed for **water column** models (1-D): biogeochemical processes happen in the water column (reaction and sinking).
- Designed for **any number** of
  - tracers, $n$
  - vertical discretization layers, $n_y$
  - model parameters to optimize, $m$
  - boundary data (e.g., light intensity, ice cover), $n_b$
  - domain data (e.g., temperature, salinity), $n_d$.

- Example: realization in **Fortran**:

```
subroutine bgc(n, ny, m, nb, nd, dt, q, t, y, u, b, d)
  integer :: n, ny, m, nb, nd
  real*8 :: dt, q(ny, n), t, y(ny, n), u(m), b(nb), d(ny, nd)
end subroutine
```
Main aim: flexibility wrt. HPC hardware: different clusters, GPU and other accelerator cards.

**Metos3D** [Piwonski and Slawig, 2014]: Marine Ecosystem Toolkit for Optimization and Simulation in 3-D

Alternatives for computation of steady annual state:
1. classical “spin-up“: Fixed point iteration, takes about 3’000-10’000 years model time with time steps of hours (ocean circulation is slow)
2. Matrix-free Newton-Krylov method

Platform-independent by usage of *PETSc: Portable Extendable Toolkit for Scientific Computing* (Argonne) [Balay et al., 1997].

Fixed-Point vs. Newton

- One time-step (e.g. 3 hours):
  \[ y_{i+1} = A_{imp,i} \left[ A_{exp,i} y_i + \Delta t q_i(y_i, u) \right] := g_i(y_i, u), \quad i = 0, \ldots, n_{\text{year}} - 1. \]

- One year: \((y^k):\) all tracers at all grid points at one time step in year \(k.\)
  \[ y^{k+1} = g_{n_{\text{year}} - 1} \circ \cdots \circ g_0(y^k, u) \]
  \[ =: G \text{ (one model year)} \]

- "Spin-up" (pseudo time-stepping, fixed-point iteration):
  \[ y^{k+1} := G(y^k, u), \quad k = 1, 2, \ldots \]

- Newton-Krylov (matrix-free):
  \[
  \text{solve } \left\{ G_y(y^k, u) \Delta y = -G(y^k, u) \right\}
  \]
  \[ y^{k+1} := y^k + \Delta y, \quad k = 1, 2, \ldots \]
Convergence towards a periodic solution.
Parallelization via domain decomposition

Theoretical and actual speedup of parallelized simulation runs.
Porting on GPUs

- Biogeochemical model (Fortran): PGI CUDA Fortran compiler
- Transport Matrix Framework:
  - PETSc-dev: preliminary version (2012), GPU enabled PETSc version
  - MatCopy(), MatScale(), MatAXPY() had to be added.

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>Max</th>
<th>Average</th>
<th>StdDev</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU</td>
<td>621.43 s</td>
<td>626.79 s</td>
<td>622.14 s</td>
<td>0.540</td>
</tr>
<tr>
<td>GPU</td>
<td>28.17 s</td>
<td>28.20 s</td>
<td>28.18 s</td>
<td>0.003</td>
</tr>
<tr>
<td>CPU : GPU</td>
<td>22.06</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Routine</th>
<th>CPU</th>
<th>GPU</th>
<th>CPU : GPU</th>
</tr>
</thead>
<tbody>
<tr>
<td>BGCStep</td>
<td>469.76 s</td>
<td>13.05 s</td>
<td>36.00</td>
</tr>
<tr>
<td>MatCopy</td>
<td>34.04 s</td>
<td>3.91 s</td>
<td>8.70</td>
</tr>
<tr>
<td>MatScale</td>
<td>23.33 s</td>
<td>1.99 s</td>
<td>11.70</td>
</tr>
<tr>
<td>MatAXPY</td>
<td>37.49 s</td>
<td>2.89 s</td>
<td>12.96</td>
</tr>
<tr>
<td>MatMult</td>
<td>58.19 s</td>
<td>5.87 s</td>
<td>9.92</td>
</tr>
</tbody>
</table>

GPU: GeForce GTX 480, CPU: Intel Xeon E5520 2.27 GHz single core
Outline

1. Motivation
2. Mathematical Aims and Challenges
3. Mathematical Analysis of the PDE
4. Numerical Solution
   - Transport Matrix Method
   - Computation of a steady annual Cycle
   - Parallelization
5. Optimization Methods
   - Surrogate-based Optimization
   - Oneshot Optimization
6. Numerical Results
7. Summary
Optimization problems and strategies:

- **Cost function:**
  \[
  \min_{y,u} J(y, u) = \frac{1}{2} \| y - y_d \|^2_Y + \frac{\alpha}{2} \| u - u_{est} \|^2
  \]

- **Parameter bounds:** \( b_l \leq u \leq b_u \).

- **Up to now:** state constraints \( y \geq 0 \) not taken into account.

- **Two types of experiments:**
  - Twin experiment with \( y_d = y(u_d) \)
  - Real data experiment: \( y_d \): World Ocean Atlas (WOA) data, preprocessed.

- **Different optimization strategies:**
  - direct optimization of reduced cost function (model as black-box)
  - surrogate-based approach for reduced cost function
  - Lagrangian approach: oneshot

- **Different optimization algorithms:**
  - derivative based
  - evolutionary/genetic algorithms.
Surrogate-based Optimization (SBO): Idea

- For optimization: expensive fine model replaced by faster surrogate.
- Construction of coarse/simplified model.
- Surrogate: Alignment of coarse model by fine model information at reference points (response correction).
- Aim: fine model information still valid in vicinity of the reference point.

When $u$ changes: Compute only coarse model, but evaluate surrogate.
Surrogate-based Optimization Algorithm

Trust region algorithm with general, not quadratic approximative model. Iterate:

1. Evaluate fine model: compute $y_f(u)$ at reference point $u$.
2. Evaluate coarse model: compute $y_c(u)$ at reference point $u$.
3. Define response correction operator:

$$A : Y \rightarrow Y \quad \text{satisfying} \quad Ay_c(u) = y_f(u)$$

and optionally

$$\frac{dAy_c}{du}(u) = \frac{dy_f}{du}(u)$$

4. Define surrogate:

$$s : U \rightarrow Y, \quad s(u) := Ay_c(u)$$

5. Minimize

$$J(s(u), u).$$
SBO: Construction of coarse model

Options:

- Output interpolation
- Different model with simplified dynamics
- Reduced order model
- Coarser mesh in space and/or time
- 0-D, 1-D, 2-D instead of 3-D
- Here: reduced iterations in spin-up/fixed-point iteration
SBO: Response correction operator

- Has to be cheap.
- Here: Multiplicative, point-wise in space and time:

\[ Ay(x, t) := \frac{y_f(x, t)}{y_c(x, t)} y(x, t) \quad \text{for all grid-points } (x, t). \]
SBO: Benefit in optimization runs

- Options in the choice of steps $j^c$ for the coarse model
- Hybrid strategy is best in this case ...
- ... but requires experiments to obtain best parameters.

$$J(y^J(u))$$

$\begin{align*}
    j^c &= 25 \\
    j^c &= 50 \\
    j^c &= 100 \\
    j^c &= 200 \\
    j^f &= 3000 \text{ direct} \\
    (j^c)_k \text{ hybrid} \\
    (j^c)_k \text{ hybrid}
\end{align*}$
Oneshot Optimization Strategy

**Idea:** Lagrange multiplier approach.

**Problem:**

$$\min_{y,u} J(y, u) \quad s.t. \quad e(y, u) = 0 \iff y = G(y, u)$$

Lagrangian: $L(y, \bar{y}, u) = J(y, u) + \bar{y}^T (G(y, u) - y)$

1st order necessary optimality conditions at $(y^*, \bar{y}^*, u^*)$:

$$0 = \frac{\partial L}{\partial y} = J_y(y^*, u^*) + \bar{y}^* T G_y(y^*, u^*) - \bar{y}^* T$$

$$0 = \frac{\partial L}{\partial \bar{y}} = G(y^*, u^*) - y^*$$

$$0 = \frac{\partial L}{\partial u} = J_u(y^*, u^*) + \bar{y}^* T G_u(y^*, u^*)$$

$\Rightarrow$ **Coupled Oneshot iteration:** (no box constraints on $u$)

\[
\begin{align*}
y_{k+1} &= G(y_k, u_k) \quad \text{feasibility} \\
\bar{y}_{k+1}^T &= \bar{y}_k^T G_y(y_k, u_k) + J_y(y_k, u_k) \quad \text{adjoint feasibility} \\
u_{k+1} &= u_k - B_k^{-1} (J_u(y_k, u_k) + \bar{y}_k^T G_u(y_k, u_k)) \quad \text{optimality}
\end{align*}
\]
Oneshot: Choice of Preconditioner B

- Suggestion by [Griewank, Hamdi 2008]:

\[
B = \frac{1}{\sigma} (\alpha G_u^\top G_u + \beta L_y^\top L_y u + L_{uu})
\]

where \( \sigma, \alpha \) and \( \beta \) are carefully chosen weights.

- Derivative information using Automatic/Algorithmic Differentiation (AD)
- ... computed from **only one** model year.
- Applied only to nonlinear biogeochemistry terms \( q \):

\[
y_{i+1} = A_{imp,i} (A_{exp,i} y_i + \Delta t q_i(y_i, u)).
\]

<table>
<thead>
<tr>
<th>Derivative</th>
<th>Mode of Computation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( J_y )</td>
<td>analytically</td>
</tr>
<tr>
<td>( J_u )</td>
<td>analytically</td>
</tr>
<tr>
<td>( \bar{y}^\top G_y, \bar{y}^\top G_u )</td>
<td>one reverse sweep of AD + analytically for linear parts</td>
</tr>
<tr>
<td>( G_u )</td>
<td>finite differences</td>
</tr>
<tr>
<td>( J_{yu} = 0 )</td>
<td>analytically</td>
</tr>
<tr>
<td>( \bar{y}^\top G_{yu}, \bar{y}^\top G_{uu} )</td>
<td>finite differences after AD computation of ( \bar{y}^\top G_y, \bar{y}^\top G_u )</td>
</tr>
</tbody>
</table>
Oneshot: Optimized Strategy

Computational time:
\[
\frac{\text{one step for } (y, \bar{y})}{\text{one step for } y} = \frac{4}{1}, \quad \frac{\text{one step for } (y, \bar{y}, u)}{\text{one step for } y} = \frac{27}{1}.
\]

Same for complete converged iteration.

\[
\begin{align*}
y_{k+1} &= G(y_k, u_k) \quad \text{feasibility} \\
\bar{y}_{k+1}^T &= \bar{y}_k^T G_y(y_k, u_k) + J_y(y_k, u_k) \quad \text{adjoint feasibility} \\
u_{k+1} &= u_k - B_k^{-1} \left( J_u(y_k, u_k) + \bar{y}_k^T G_u(y_k, u_k) \right) \quad \text{optimality}
\end{align*}
\]

Optimized/adapted Oneshot:
- Perform \(n_p\) pre-iterations only on \(y, \bar{y}\)
- Perform \(n_e\) exact steps for \(y, \bar{y}, u\) (with \(B\) computed in every step).
- Perform inexact steps for \(y, \bar{y}, u\) (with \(B\) computed every \(n_i\) steps) until convergence.

Computational time for \(n_p = n_e = 500, n_i = 5\):
\[
\frac{\text{converged iteration for } (y, \bar{y}, u)}{\text{converged iteration for } y} = \frac{14}{1}.
\]
Equivalent number of model spin-ups

(Twin experiment, counting 1 model spin-up $\triangleq 10^{\prime}000$ model years, $u_{est} \neq u_{opt}, \alpha = 0.01$).
Outline

1 Motivation
2 Mathematical Aims and Challenges
3 Mathematical Analysis of the PDE
4 Numerical Solution
   ■ Transport Matrix Method
   ■ Computation of a steady annual Cycle
   ■ Parallelization
5 Optimization Methods
   ■ Surrogate-based Optimization
   ■ Oneshot Optimization
6 Numerical Results
7 Summary
Parameter Convergence (Twin experiment)

$u_1$, DOP remineralization rate

Convergence towards the reference parameter $u_{d,1}$. 
Real Data: World Ocean Atlas

Decay of the weighted cost function. Weights are relative volumes.
Non-unique parameters (World Ocean Atlas)

Behavior of different parameters for two different starting values.
Outline

1 Motivation
2 Mathematical Aims and Challenges
3 Mathematical Analysis of the PDE
4 Numerical Solution
   ■ Transport Matrix Method
   ■ Computation of a steady annual Cycle
   ■ Parallelization
5 Optimization Methods
   ■ Surrogate-based Optimization
   ■ Oneshot Optimization
6 Numerical Results
7 Summary
Optimization Summary:

- Optimization methods work for synthetic (attainable) data.
- Different optimization algorithm give similar results.
- Cost reduction reduction worse for real data.
- Real data: Identified parameters are not unique → requires further analysis.
- Both SBO and Oneshot reduce the computational effort.
- Detailed performance comparison has to be done.