A Parareal Algorithm Applied on the Computation of Periodic States in Marine Ecosystems

Thomas Slawig, Jaroslaw Piwonski, Björn Peiler, Christina Roschat

Kiel University, Dep. Computer Science, Cluster "The Future Ocean"

Workshop *Parallel in Time* – May 2015 Dresden
Outline

1 Motivation: Marine Ecosystem Modeling

2 Mathematical Formulation: State Equation

3 Numerical Simulation

4 Optimization

5 Parallelization in Time
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 m mol P/m$^3$. The longitudinal and latitudinal resolution is at $1.0^\circ$. 
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, $\mathbb{R}^n$), $u \in \mathbb{R}^m$, $m \ll n$

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

s.t. model equations are fulfilled: $e(y, u) = 0$. 
Online vs. offline simulation

Ocean circulation fields

Navier-Stokes equations

\[ \text{online} \quad \longleftrightarrow \quad \text{offline} \]

Tracer distribution

System of transport equations with nonlinear coupling

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

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

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

+ diffusion \quad advection \quad biogeochemistry

+ initial and boundary conditions.

- \( u \): constant model parameters to be determined.

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

\[
y(t + T) = y(t), \quad t \in [0, T],
\]

for "climatological“ (annually periodic) forcing data and coupling terms

\[
\kappa(t + T) = \kappa(t), \\
\vec{v}(t + T) = \vec{v}(t), \\
q_j(t + T, \cdot, \cdot) = q_j(t, \cdot, \cdot), \quad t \in [0, T].
\]
(Simplest) example: N-DOP model by Parekh et al.

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

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

- **Typical nonlinearity and non-local term:**
  \[ F(y_1) = \alpha \frac{y_1}{y_1 + K}, \]
  \[ E(y_1) = c \left[ \frac{z}{\text{depth}} \right]^m \int_{\text{depth}}^0 (1 - \nu)F(y_1)dz \]

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

- **Model:** [Dutkiewicz et al., 2005], mitgcm.org.
Ecosystem model equation of $N$-DOP type in classical form:

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

with

- $\Omega \subset \mathbb{R}^n$, $n \leq 3$, 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 $\vec{v}(t) \cdot \eta = 0$ on $\Gamma$
- $\kappa \in L^\infty(\Omega \times [0, T])$ with $\kappa_{\text{min}} := \text{ess inf}_{(x,t)} \kappa(x,t) > 0$
- $d$, $b$ reaction terms depending on space, time and an appropriate function space with $d(0) = 0 = b(0)$
- $\lambda > 0$
Existence of nontrivial periodic solutions

Theorem (Roschat 2014)

Let $C, \lambda > 0$. Under the assumptions

- $d, b$ are continuous

\[
\sum_{j=1}^{2} \left( \int_{\Omega} d_j(y, x, t)\,dx + \int_{\Gamma} b_j(y, s, t)\,ds \right) = 0 \text{ for almost all } t \text{ and all } y
\]

- \( \max\{|d_j(y, x, t)|, |b_j(y, x, t)|\} \leq M \text{ for all } y, j = 1, 2 \text{ and almost all } (x, t). \)

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

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

In particular, the solution is not trivial.

But: No result about uniqueness.
Off-line simulation with transport matrices

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

- ... use discretized transport operators (matrices) in the \( i \)-th time step: 
  \[ y_{i+1} = A_{\text{imp},i} \left( A_{\text{exp},i} y_i + \Delta t q_i(y_i, u) \right), \quad i = 1, \ldots \]

\( y_i \): 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 deg (128 × 64) surface grid, 15 vertical layers, 2880 time steps, matrix dimension per tracer \( \approx 50'000 \).

- Software M\textsc{etos3d} [Piwonski, Slawig 2015]: Marine Ecosystem Toolkit for Optimization and Simulation in 3-D, on github.
Fixed-Point vs. Newton

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

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

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

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

Theoretical and actual speedup of parallelized simulation runs.
Numerical solution of the optimization problem

- Find $y \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $m \ll n$

$$\min J(y, u) = \frac{1}{2} \|y - y_{data}\|^2 + \frac{\alpha}{2} \|u - u_{est}\|^2 \quad \text{s.t. } e(y, u) = 0.$$  

- State $y$ is converged solution of a iteration (fixed-point or Newton):

$$y_{k+1} = G(y_k, u)$$

- Fixed-point: 3-10’000 years model time with 2880 times-steps each
- Newton: $\approx$ 1’000 years model time.
- Spatial parallelization already exploited
- Ideally: even finer discretization
- Gradient based, iterative optimization: typically 100s of function evaluations
- $\implies$ Acceleration techniques needed
- Motivation for additional parallelization in time
First approach: *Parareal* with coarser time-steps

- Same spatial, but coarser temporal resolution:
  
  \[
  \text{time-step} \approx \frac{3}{192} \text{ hrs.} = \frac{1}{64} = \frac{2880}{45} \text{ steps/year}
  \]

- Result for 8 time intervals/processors (one year model time):

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Iteration</th>
<th>Accuracy</th>
<th>Tracer 1</th>
<th>Tracer 2</th>
<th>mean time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>standard</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>221.000</td>
</tr>
<tr>
<td>parareal</td>
<td>1</td>
<td>7.69e-03</td>
<td>3.37e-03</td>
<td></td>
<td>56.126</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.50e-03</td>
<td>7.28e-04</td>
<td></td>
<td>108.441</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>7.73e-04</td>
<td>3.72e-04</td>
<td></td>
<td>160.303</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>2.59e-04</td>
<td>1.17e-04</td>
<td></td>
<td>211.828</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>5.65e-05</td>
<td>3.19e-05</td>
<td></td>
<td>262.946</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>8.69e-06</td>
<td>4.64e-06</td>
<td></td>
<td>313.622</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>5.55e-07</td>
<td>2.78e-07</td>
<td></td>
<td>363.646</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>1.74e-13</td>
<td>8.89e-14</td>
<td></td>
<td>413.596</td>
</tr>
</tbody>
</table>

- Spin-up has $\approx 1e-04$ accuracy, so (especially for the first steps) 1-3 parareal iterations may be enough.
Future work

Parameter ID in marine ecosystems:
- Repeat tests for complete spin-up.
- Micro-macro parareal:
  - Generate matrices corresponding to coarser spatial mesh.
  - Has to be done algebraically, since no coarser model available.
- Later: use current mesh (2.8 deg) as macro and an even finer one (1 deg) as micro.

Different project: Long-time paleo climate simulation
- Micro-macro setting for global climate models.