Assessment and Improvement of Models of the global Carbon Cycle using Parameter Optimization

Thomas Slawig, Claudia Kratzenstein, Jaroslaw Piwonski, Christina Roschat

Kiel University, Cluster "The Future Ocean"

Cooperation with groups of:
A. Griewank (HU Berlin), N. Gauger (then HU Berlin), A. Oschlies (GEOMAR Kiel).

October 2015 Hamburg
1 Motivation

2 Optimization problem

3 Periodic solutions: why?

4 Mathematical analysis: transient solutions

5 Analysis for periodic solutions

6 Analysis of the optimization problem

7 Computation of a periodic solution

8 Optimization methods
1 Motivation

2 Optimization problem

3 Periodic solutions: why?

4 Mathematical analysis: transient solutions

5 Analysis for periodic solutions

6 Analysis of the optimization problem

7 Computation of a periodic solution

8 Optimization methods
Climate change and the global carbon cycle

Citation from the Intergovernmental Panel on Climate Change (IPCC) Assessment Report 5 (AR5), Summary for Policymakers Sec. E.7, 2013:

"Climate change will affect carbon cycle processes in a way that will exacerbate the increase of CO₂ in the atmosphere (high confidence). Further uptake of carbon by the ocean will increase ocean acidification."

"Ocean uptake of anthropogenic CO₂ will continue under all four RCPs¹ through to 2100, with higher uptake for higher concentration pathways (very high confidence)."

"Based on ... Models, there is high confidence that the feedback between climate and the carbon cycle is positive in the 21st century; that is, climate change will partially offset increases in land and ocean carbon sinks caused by rising atmospheric CO₂. As a result more of the emitted anthropogenic CO₂ will remain in the atmosphere."

¹RCP: Representative Concentration Pathway (greenhouse gas emissions scenario)
RCPs: Representative Concentration Pathways

... or greenhouse gas emissions scenarios:

- Define possible range of additional radiative forcing in the year 2100 relative to pre-industrial values in 1750 ($\approx 340 \text{ W/m}^2$).

<table>
<thead>
<tr>
<th>RCP</th>
<th>2.6</th>
<th>4.5</th>
<th>6</th>
<th>8.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>global annual GHG emissions peak</td>
<td>2010-2020</td>
<td>$\approx 2040$</td>
<td>$\approx 2080$</td>
<td>$&gt;2100$</td>
</tr>
<tr>
<td></td>
<td>and decline substantially thereafter</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>results in</td>
<td>475 ppm GHG $CO_2$ equivalents</td>
<td>630</td>
<td>800</td>
<td>1313</td>
</tr>
<tr>
<td></td>
<td>in 2100</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Prognosted temperature changes

Image: IPCC AR5. Time series of global annual mean surface air temperature anomalies (relative to 1986-2005). Projections are shown for each RCP for the multi-model mean (solid lines) and the 5 to 95% range (±1.64 standard deviation) across the distribution of individual models (shading). Discontinuities at 2100 are due to different numbers of models performing the extension runs beyond the 21st century and have no physical meaning. Numbers in the figure indicate the number of different models contributing to the different time periods.
The global carbon cycle

Atmosphere 589 ± 240 ± 10
(average atmospheric increase: 4 (PgC yr⁻¹))

Net ocean flux
2.3 ± 0.7
0.7

Ocean-atmosphere gas exchange
80 ± 60 ± 20
78.4 ± 60.7 ± 17.7

Net land flux
1.7

Fossil fuels (coal, oil, gas)
0.7 ± 0.6
7.8 ± 0.6

Gross photosynthesis
1.2 ± 10.89 ± 14.1

Net land use change
1.1 ± 0.8

Export from soils to rivers
1.7

Volcanism
0.1

Rock weathering
0.1

Freshwater outgassing
10

Surfex ocean
900

Intermediate & deep sea
37,100

Surface ocean
900

Marine biota
3

Burial
0.2

Rivers
0.9

Dissolved organic carbon
700

Net land use change
1.1 ± 0.8

Export from soils to rivers
1.7

Net land flux
1.7

Volcanism
0.1

Rock weathering
0.1

Fossil fuel reserves
Gas: 383-1135
Oil: 173-264
Coal: 446-541

-365 ± 30

Vegetation
450-650

Soils
1500-3400

Permafrost
~1700

Units
Fluxes: (PgC yr⁻¹)
Stocks: (PgC)

Image: IPCC AR5, Pg = 10^{15}g
Our research: Modeling the marine carbon cycle

Ocean circulation  \[\leftrightarrow\] Marine ecosystem (plankton, photosynthesis)

Velocity, pressure, temperature, salinity, turbulent diffusion

Navier-Stokes equations + System of transport eqn., nonlinear coupling

Thomas Slawig et al. Assessment and Improvement of Models of the global Carbon Cycle using Parameter Optimization
Modeling the *marine* carbon cycle

Ocean circulation "online" → "offline"

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

Velocity, pressure, temperature, salinity, turbulent diffusion

Navier-Stokes equations + System of transport eqn., nonlinear coupling

Marine ecosystem

Tracer concentrations

Longitude [degrees]

Latitude [degrees]
Different models for the marine ecosystem component

different models, different complexity

Marine ecosystem

Tracer concentrations

System of transport eqn., nonlinear coupling

... and new ones to be developed (GEOMAR)
In contrast to the physical system (ocean circulation etc.) ...

... there is no consensus about the model (equations) to be used for the marine ecosystem

... there is not the (best) model (in contrast: Navier-Stokes for flows)

”More complexity $\Rightarrow$ better results” is not true

Thus: new models are under development.

Which model is better? the best?

Quality of the model is assessed by optimizing its parameters

... and looking at the quality of the model-to-data fit

... and the resulting parameters.
In contrast to the physical system (ocean circulation etc.) ...

... there is no consensus about the model (equations) to be used for the marine ecosystem

... there is not the (best) model (in contrast: Navier-Stokes for flows)

"More complexity ⇒ better results" is not true

Thus: new models are under development.

Which model is better? the best?

Quality of the model is assessed by optimizing its parameters

... and looking at the quality of the model-to-data fit

... and the resulting parameters.
In contrast to the physical system (ocean circulation etc.) ...

... there is no consensus about the model (equations) to be used for the marine ecosystem

... there is not the (best) model (in contrast: Navier-Stokes for flows)

”More complexity ⇒ better results” is not true

Thus: new models are under development.

Which model is better? the best?

Quality of the model is assessed by optimizing its parameters

... and looking at the quality of the model-to-data fit

... and the resulting parameters.
Task of the model/parameter optimization

- In contrast to the physical system (ocean circulation etc.) ...
- ... there is no consensus about the model (equations) to be used for the marine ecosystem
- ... there is not *the* (best) model (in contrast: Navier-Stokes for flows)
- ”More complexity $\Rightarrow$ better results” is not true
- Thus: new models are under development.

- Which model is better? the best?
- Quality of the model is assessed by optimizing its parameters
- ... and looking at the quality of the model-to-data fit
- ... and the resulting parameters.
In contrast to the physical system (ocean circulation etc.) ... 
... there is no consensus about the model (equations) to be used for the marine ecosystem 
... there is not *the* (best) model (in contrast: Navier-Stokes for flows) 
"More complexity ⇒ better results" is not true 
Thus: new models are under development. 
Which model is better? the best? 
Quality of the model is assessed by optimizing its parameters 
... and looking at the quality of the model-to-data fit 
... and the resulting parameters.
Task of the model/parameter optimization

- In contrast to the physical system (ocean circulation etc.) ...
- ... there is no consensus about the model (equations) to be used for the marine ecosystem
- ... there is not the (best) model (in contrast: Navier-Stokes for flows)
- ”More complexity $\Rightarrow$ better results” is not true
- Thus: new models are under development.
- Which model is better? the best?
- Quality of the model is assessed by optimizing its parameters
  - ... and looking at the quality of the model-to-data fit
  - ... and the resulting parameters.
Task of the model/parameter optimization

In contrast to the physical system (ocean circulation etc.) ...

... there is no consensus about the model (equations) to be used for the marine ecosystem

... there is not *the* (best) model (in contrast: Navier-Stokes for flows)

”More complexity $\implies$ better results” is not true

Thus: new models are under development.

Which model is better? the best?

Quality of the model is assessed by optimizing its parameters

... and looking at the quality of the model-to-data fit

... and the resulting parameters.
In contrast to the physical system (ocean circulation etc.) ...

... there is no consensus about the model (equations) to be used for the marine ecosystem

... there is not the (best) model (in contrast: Navier-Stokes for flows)

”More complexity ⇒ better results” is not true

Thus: new models are under development.

Which model is better? the best?

Quality of the model is assessed by optimizing its parameters

... and looking at the quality of the model-to-data fit

... and the resulting parameters.
Consequences for the optimization

- Fast optimization (requires fast state equation solvers, see below)
  - Flexible software configuration that allows to compare models ...
  - ... with same forcing ...
  - ... and using the same optimization setting
  - Easy coupling of different models, new models
  - Flexibility w.r.t. optimization methods
    ("global" vs. gradient-based and hybrid methods)
  - Theoretical results about uniqueness of parameters are helpful
  - Estimates of accuracy/uncertainty
Consequences for the optimization

- Fast optimization (requires fast state equation solvers, see below)
- Flexible software configuration that allows to compare models ...
  - ... with same forcing ...
  - ... and using the same optimization setting
- Easy coupling of different models, new models
- Flexibility w.r.t. optimization methods
  - ”global” vs. gradient-based and hybrid methods
- Theoretical results about uniqueness of parameters are helpful
- Estimates of accuracy/uncertainty
Consequences for the optimization

- Fast optimization (requires fast state equation solvers, see below)
- Flexible software configuration that allows to compare models ...
  ... with same forcing ...
  ... and using the same optimization setting
- Easy coupling of different models, new models
- Flexibility w.r.t. optimization methods
  ("global" vs. gradient-based and hybrid methods)
- Theoretical results about uniqueness of parameters are helpful
- Estimates of accuracy/uncertainty
Consequences for the optimization

- Fast optimization (requires fast state equation solvers, see below)
- Flexible software configuration that allows to compare models ...
  - ... with same forcing ...
  - ... and using the same optimization setting
- Easy coupling of different models, new models
  - Flexibility w.r.t. optimization methods
    ("global" vs. gradient-based and hybrid methods)
- Theoretical results about uniqueness of parameters are helpful
- Estimates of accuracy/uncertainty
Consequences for the optimization

- Fast optimization (requires fast state equation solvers, see below)
- Flexible software configuration that allows to compare models ...
- ... with same forcing ...
- ... and using the same optimization setting
- Easy coupling of different models, new models
- Flexibility w.r.t. optimization methods
  ("global" vs. gradient-based and hybrid methods)
- Theoretical results about uniqueness of parameters are helpful
- Estimates of accuracy/uncertainty
Consequences for the optimization

- Fast optimization (requires fast state equation solvers, see below)
- Flexible software configuration that allows to compare models ...
- ... with same forcing ...
- ... and using the same optimization setting
- Easy coupling of different models, new models
- Flexibility w.r.t. optimization methods ("global" vs. gradient-based and hybrid methods)
- Theoretical results about uniqueness of parameters are helpful
- Estimates of accuracy/uncertainty
Consequences for the optimization

- Fast optimization (requires fast state equation solvers, see below)
- Flexible software configuration that allows to compare models ...
- ... with same forcing ...
- ... and using the same optimization setting
- Easy coupling of different models, new models
- Flexibility w.r.t. optimization methods ("global" vs. gradient-based and hybrid methods)
- Theoretical results about uniqueness of parameters are helpful
- Estimates of accuracy/uncertainty
1 Motivation

2 Optimization problem

3 Periodic solutions: why?

4 Mathematical analysis: transient solutions

5 Analysis for periodic solutions

6 Analysis of the optimization problem

7 Computation of a periodic solution

8 Optimization methods
Model calibration and assessment

- Fit model output to data
- Optimize parameters $u$ in the underlying ecosystem model
- $y \in Y$ (either function space or, discretized, $\mathbb{R}^n$):

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

s.t. model equations are fulfilled for biogeochemical tracers $y = (y_j)_j$:

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

+ initial and boundary conditions.

- Given forcing data $\vec{v}, \kappa$.
- $u \in U_{ad} \subset U := \mathbb{R}^m$: constant model parameters to be determined.
- $u_{est}$: parameter estimate (optional).
(Simplest) example: N-DOP model by Parekh et al.

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

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

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

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

1. Motivation
2. Optimization problem
3. Periodic solutions: why?
4. Mathematical analysis: transient solutions
5. Analysis for periodic solutions
6. Analysis of the optimization problem
7. Computation of a periodic solution
8. Optimization methods
Calibration is performed for a *periodic* solution: why?

- Straightforward idea for parameter identification/model calibration:
  - fix some historic time slice (e.g., one decade)
  - compute ocean circulation fields that match available circulation data
  - perform parameter optimization for marine ecosystem model with this forcing
  - and use available ecosystem measurement data

- But: There is no ”typical” decade
  ⇒ Forcing for the next decade might be completely different
  ⇒ results may be useless for prediction
  and: sparse measurement data.

- Instead: take “climatological“ (annually periodic) averaged forcing data

- Compute annually periodic solution:

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

- Optimize with sampled and averaged measurement data.

- Idea: rather understanding the model than determination of concrete parameters
Example: Sampled data
State Equations

- State equations:

\[
\frac{\partial}{\partial t} y_j - \text{div}(\kappa \nabla y_j) + \mathbf{v} \cdot \nabla y_j + q_j(y, u) = 0, \quad Q := \Omega \times (0, T) \\
\frac{\partial \eta}{\partial t} y_j + b_j(y, u) = 0, \quad \Sigma := \Gamma \times (0, T), \\
y(0) = y_0, \quad \Omega, \quad j = 1, \ldots, s.
\]

- Desired tracer balance leads to nonlinear boundary terms \(b_j\).

- Weak form: Find \(y \in W(0, T)^s\) :

\[
y' + By + F(y) = 0 \quad \text{in} \quad L^2(0, T; H^1(\Omega)^*)^s \\
y(0) = y_0.
\]

with

\[
\langle By, w \rangle_{L^2(0, T; H^1(\Omega)^*)^s} = \sum_{j=1}^s \int_Q (\kappa \nabla y_j \cdot \nabla w_j + (\mathbf{v} \cdot \nabla y_j) w_j) \, dxdt,
\]

\[
\langle F(y), w \rangle_{L^2(0, T; H^1(\Omega)^*)^s} = \sum_{j=1}^s \left( \int_Q q_j(y) w_j \, dxdt + \int_{\Sigma} b_j(y) w_j \, dsdt \right).
\]
Transient solutions

Theorem (C. Roschat 2012)

Consider \( F = F_1 + F_2 \) with \( F_1 \) to be Lipschitz continuous and \( F_2 \) to be monotone with \( F_2(0) = 0 \). Then, there is a unique solution \( y \in W(0, T)^s \) of the initial value problem (??). Furthermore, the estimate

\[
\|y\|_{W(0,T)^s} \leq C \left( \|f\|_{L^2(0,T;H^1(\Omega)^*)^s} + \|y_0\|_{L^2(\Omega)^s} \right)
\]

holds with a constant \( C > 0 \) independent of \( y \) and \( y_0 \).

Proof.

Banach’s fixed-point theorem or Galerkin’s method.
1 Motivation
2 Optimization problem
3 Periodic solutions: why?
4 Mathematical analysis: transient solutions
5 Analysis for periodic solutions
6 Analysis of the optimization problem
7 Computation of a periodic solution
8 Optimization methods
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

$$y' + Ay = f, \quad u(0) = u(T),$$

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

Problem: Coercivity usually not fulfilled by ecosystem model equations
Existence of nontrivial periodic solutions

- A result was obtained for a two-tracer model of the general form

\[
\begin{align*}
y_1' + B(y_1) - \lambda y_2 + F_1(y) &= 0 \\
y_2' + B(y_2) + \lambda y_2 + F_2(y) &= 0
\end{align*}
\]

- The N-DOP-model (presented above) fits in this setting:

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

- Two tracers imported because the prove is based on treating one equation first, then the sum, and finally the difference. (Works for the special model structure).
Existence of nontrivial periodic solutions

**Theorem (C. Roschat 2014)**

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

- $q, b$ are continuous
- \[
\sum_{j=1}^{2} \left( \int_{\Omega} q_j(y, x, t) \, dx + \int_{\Gamma} b_j(y, s, t) \, ds \right) = 0 \text{ for almost all } t \text{ and all } y
\]
- \[
\max\{ |q_j(y, x, t)|, |b_j(y, x, t)| \} \leq M \text{ for all } y, j = 1, 2 \text{ and almost all } (x, t)
\]

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.
Existence + optimality system: transient problems

Theorem (C. Roschat 2015)

Let the Banach space $U$ be reflexive, $F: U \times Y \rightarrow L^2(0, T; H^1(\Omega)^*)$ and $f: U \rightarrow L^2(0, T; H^1(\Omega)^*)$ be weakly sequentially continuous. Assume that the set of admissible states $Y_{ad} := \{ y \in W(0, T)^s : \exists u \in U_{ad} : (u, y) \in X_{ad} \}$ is bounded in $W(0, T)^s$. Then there is an optimal parameter $\bar{u} \in U_{ad}$ that satisfies

$$
\begin{align*}
\bar{y}' + B\bar{y} + F(\bar{u}, \bar{y}) &= f(\bar{u}) \\
\bar{y}(0) &= y_0 \\
-p' + B^*(p) + F^*(p) &= \bar{y} - y_d \\
p(T) &= 0 \\
([f'(\bar{u}) - F'_u(\bar{u}, \bar{y})]p + \gamma(\bar{u} - u_d), u - \bar{u})_U &\geq 0 \quad \text{for all } u \in U_{ad}.
\end{align*}
$$

For a periodic solution, existence of an optimal parameter can be shown.
Computation of a periodic solution

- System of transport equations:

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

- Discretized in space and summarized for all tracers at time step \( i \):

\[ y_i = \left( y_j(t_i, x_k) \right)_{j=1, \ldots, s, k=1, \ldots, n_x} \]

- One discrete time-step (e.g. 3 hours) with operator splitting:

\[
\begin{align*}
y_{i+1/2} &= (I + \Delta t L_{exp,i})y_i + \Delta t q_i(y_i, u) \\
y_{i+1} &= (I - \Delta t L_{imp,i})^{-1}y_{i+1/2} =: g_i(y_i, u), \quad i = 0, \ldots, n_{year} - 1.
\end{align*}
\]

- One year: \((y^k)\): all tracers at all grid points at one time step in year \( k \).

\[ y^{k+1} = g_{n_{year}-1} \circ \cdots \circ g_0(y^k, u) =: G \text{ (one model year)} \]
Off-line simulation with transport matrices

- Instead of directly using the pre-computed forcing data ...

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

- ... use discretized transport operators (matrices) in the \( i \)-th time step:

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

\[ y_i: \text{vector of all tracers at all grid points at time step } i. \]

- Sparse implicit and explicit transport matrices:

\[ A_{imp,i} = (I - \Delta t L_{imp,i})^{-1} \]
\[ A_{exp,i} = (I + \Delta t L_{exp,i}) \]

- Example resolution: \( 2.8125^\circ \), i.e., \( 128 \times 64 \) surface grid, 15 vertical layers, 2880 time steps, matrix dimension per tracer \( \approx 50^{000}. \)
Computation of the periodic solution

- One year: \( (y^k): \) all tracers at all grid points at one time step in year \( k \).

\[
y^{k+1} = g_{n_{\text{year}}-1} \circ \cdots \circ g_0(y^k, u)
\]

\[= G \text{ (one model year)}\]

- Standard method in marine science: “Spin-up“ (pseudo time-stepping, fixed-point iteration):

\[
y^{k+1} := G(y^k, u), \quad k = 1, 2, \ldots
\]

About 3’000-10’000 years, about 1 hour cputime (spatially parallelized).

- Newton-Krylov (matrix-free):

\[
solve \quad G_y(y^k, u)\Delta y = -G(y^k, u)
\]

\[
y^{k+1} := y^k + \Delta y, \quad k = 1, 2, \ldots
\]
Convergence towards a periodic solution.
Parallelization via domain decomposition

Theoretical and actual speedup of parallelized simulation runs.
1 Motivation
2 Optimization problem
3 Periodic solutions: why?
4 Mathematical analysis: transient solutions
5 Analysis for periodic solutions
6 Analysis of the optimization problem
7 Computation of a periodic solution
8 Optimization methods
Decay of the weighted cost function. Weights are relative volumes. Model eval.: $\approx 1$ hour on a HPC cluster.
Behavior of the parameter $u_1$. 

$u_1$, DOP remineralization rate
Behavior of the parameter $u_3$. 

Thomas Slawig et al. Assessment and Improvement of Models of the global Carbon Cycle using Parameter Optimization
Lagrange method: Oneshot Optimization

**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: 
\[
L(y, \bar{y}, u) = J(y, u) + \bar{y}^\top (G(y, u) - y)
\]

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

\[
0 = \frac{\partial L}{\partial y} = J_y(y^*, u^*) + \bar{y}^* \bar{G}_y(y^*, u^*) - \bar{y}^* \bar{G}_y
\]

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

\[
0 = \frac{\partial L}{\partial u} = J_u(y^*, u^*) + \bar{y}^* \bar{G}_u(y^*, u^*)
\]

\[\Rightarrow \text{Coupled Oneshot iteration:} \quad \text{(no box constraints on } u)\]

\[
\begin{align*}
y_{k+1} &= G(y_k, u_k) & \text{feasibility} \\
\bar{y}_{k+1} &= \bar{y}_k G_y(y_k, u_k) + J_y(y_k, u_k) & \text{adjoint feasibility} \\
u_{k+1} &= u_k - B_k^{-1} (J_u(y_k, u_k) + \bar{y}_k \bar{G}_u(y_k, u_k)) & \text{optimality}
\end{align*}
\]
Equivalent number of model spin-ups

(Twin experiment, counting 1 model spin-up $\approx 10^{\prime}000$ model years, $u_{est} \neq u_{opt}, \alpha = 0.01$).
Direct optimization (using the reduced functional) works, but takes very long.

One-shot has problems for periodic solutions when multiple data points are used (usually designed for steady solutions).

Modelers’ choice: take evolutionary algorithm and use HLRN ;-)

At the moment: Use different ideas to reduce effort for the state equation solver:
- POD
- parallel-in-time computation

Question of uniqueness of optimal parameters remains and is likely not answerable.