Algorithmic Optimal Control - CO_2 Uptake of the Ocean

Junior Research Group A3

Surrogate-Based Optimization for Validation of Marine Ecosystem Models

Dipl. Phys. Malte Prieß* - mpr@informatik.uni-kiel.de

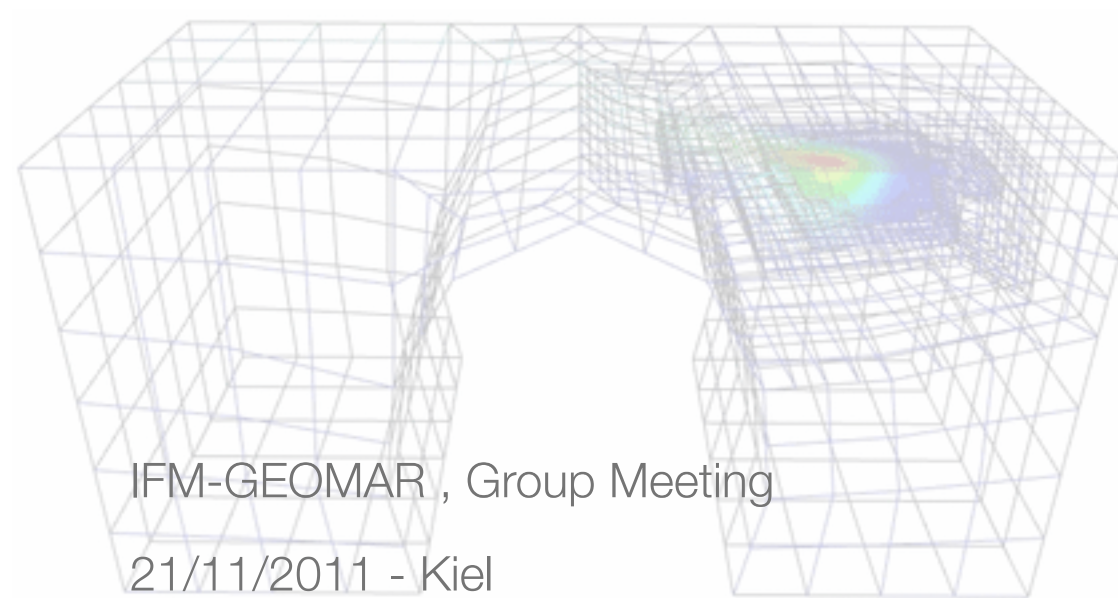
in collaboration with Prof. Thomas Slawig*, Prof. Slawomir Koziel**, Prof. Andreas Oschlies

*

Institute for Computer Science
Cluster The Future Ocean
Christian-Albrechts Universität zu Kiel, Germany

**

Engineering Optimization & Modeling Center
School of Science and Engineering
Reykjavik University, Iceland

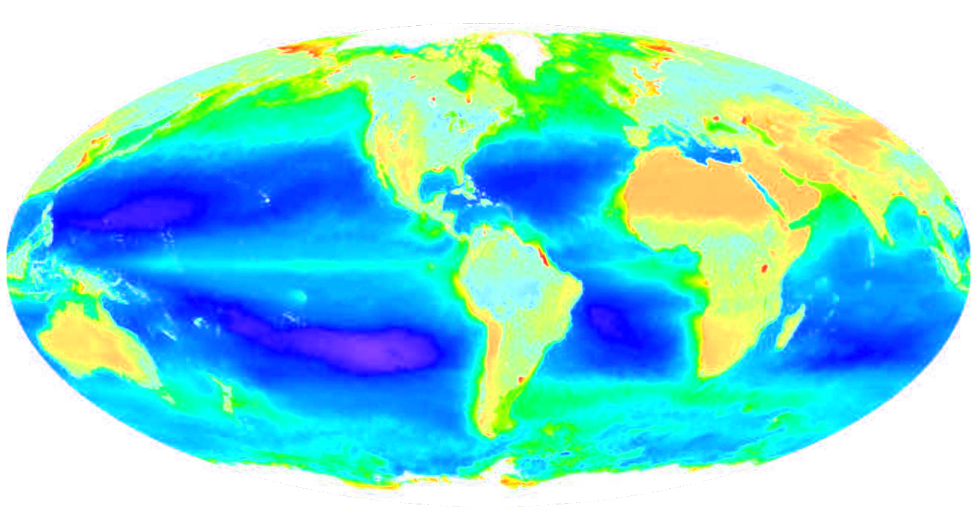
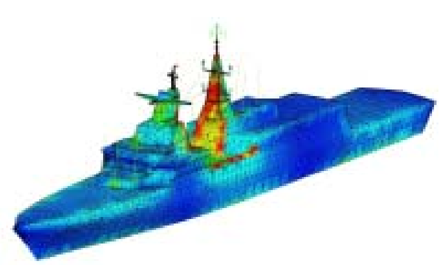
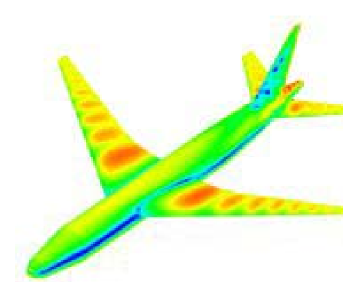
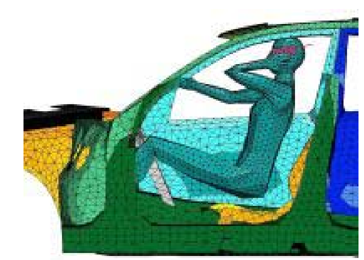
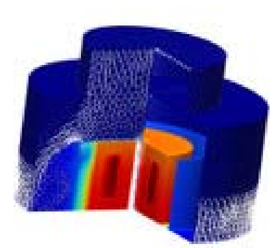
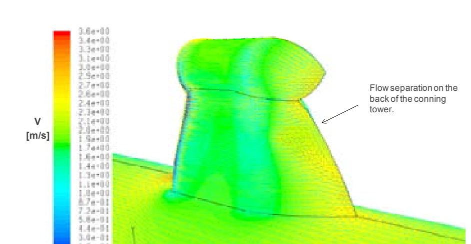
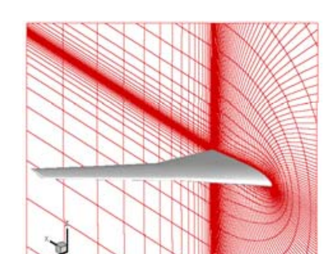


IFM-GEOMAR , Group Meeting

21/11/2011 - Kiel

Some facts ...

- *Solving nonlinear optimization problems* where computation of the objective function involves *time consuming computer simulations* may be quite *challenging*
- *Fundamendal bottleneck*:
  most of conventional optimization algorithms, whether deterministic (e.g., gradient-based) or stochastic (e.g., meta-heuristics), typically require *large number of objective function evaluations*
- This typically translates into *prohibitively high computational cost*
- Development of methods that would reduce the number of expensive simulations necessary to yield a satisfactory solution becomes critical
- Optimization of *complex coupled hydrodynamical marine ecosystem models* is a representative example
- Evaluation times of *hours up to several days for a single model evaluation* are not uncommon

- *Nonlinear optimization problems of the form*

$$\min_{\mathbf{u}\in U_{ad}} \; J(\,\mathbf{y}(\mathbf{u})\,) \qquad (1)$$

(subject to some constraints)

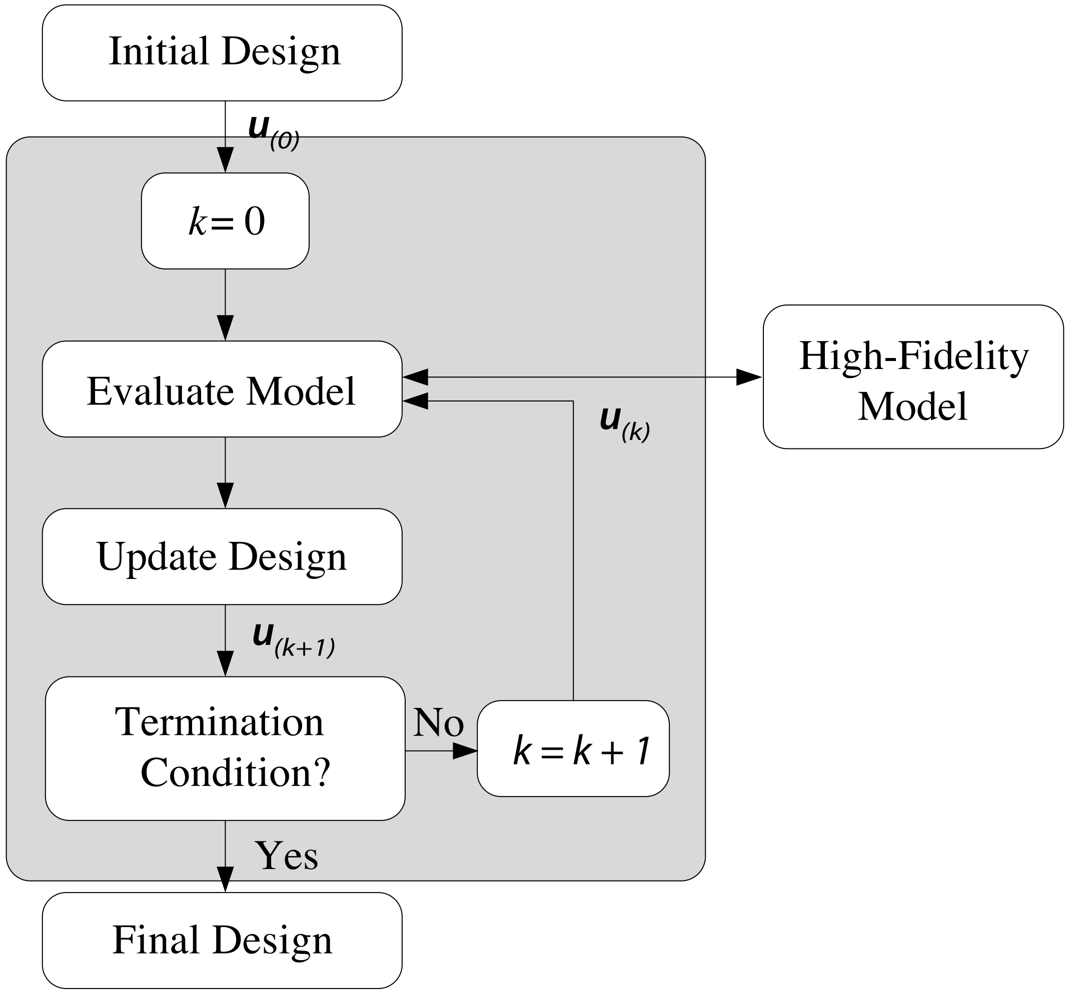
- Complex *high-fidelity (fine) model* y is *computationally expensive*

- *Straightforward attempt:* „Direct" Optimization

- (1) is a *tedious process* or even beyond the capabilities of modern computer power

- Assuming 30 minutes for a single model evaluation a direct optimization will most-likely require *several days up to weeks*

„Direct" Optimization



Source:
L. Leifsson, S. Koziel, Reykjavik University, Iceland

Surrogate-Based Optimization (SBO)

- Idea:
  Exploit a „surrogate“, a *computationally cheap but yet reasonably accurate representation of the fine model*

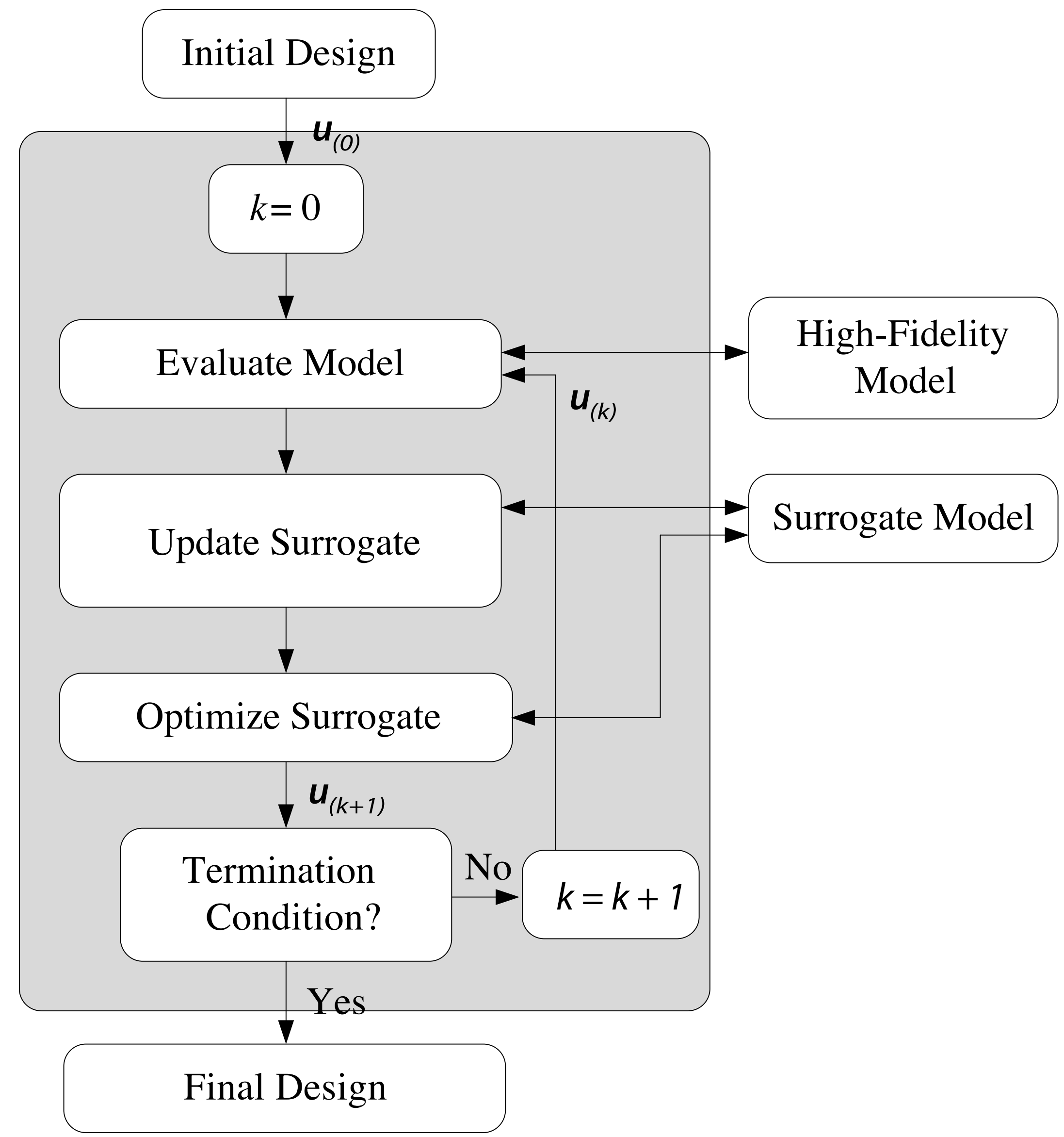
$$\mathbf{u}_{k+1} = \underset{\mathbf{u}\in U_{ad},||\mathbf{u}-\mathbf{u}_k||\leq\delta_k}{\operatorname{argmin}} J\left(\mathbf{s}_k(\mathbf{u})\right) \qquad (2)$$

- It is typically *updated using the fine model data* accumulated during the process

- The *scheme (2) is normally iterated* in order to refine the search and to locate a (local) fine model optimum of (1) as precisely as possible

- ... until some *stopping criteria* are satisfied (e.g. $||\mathbf{u}_{k+1} - \mathbf{u}_k|| < \varepsilon$)

Surrogate-Based Optimization



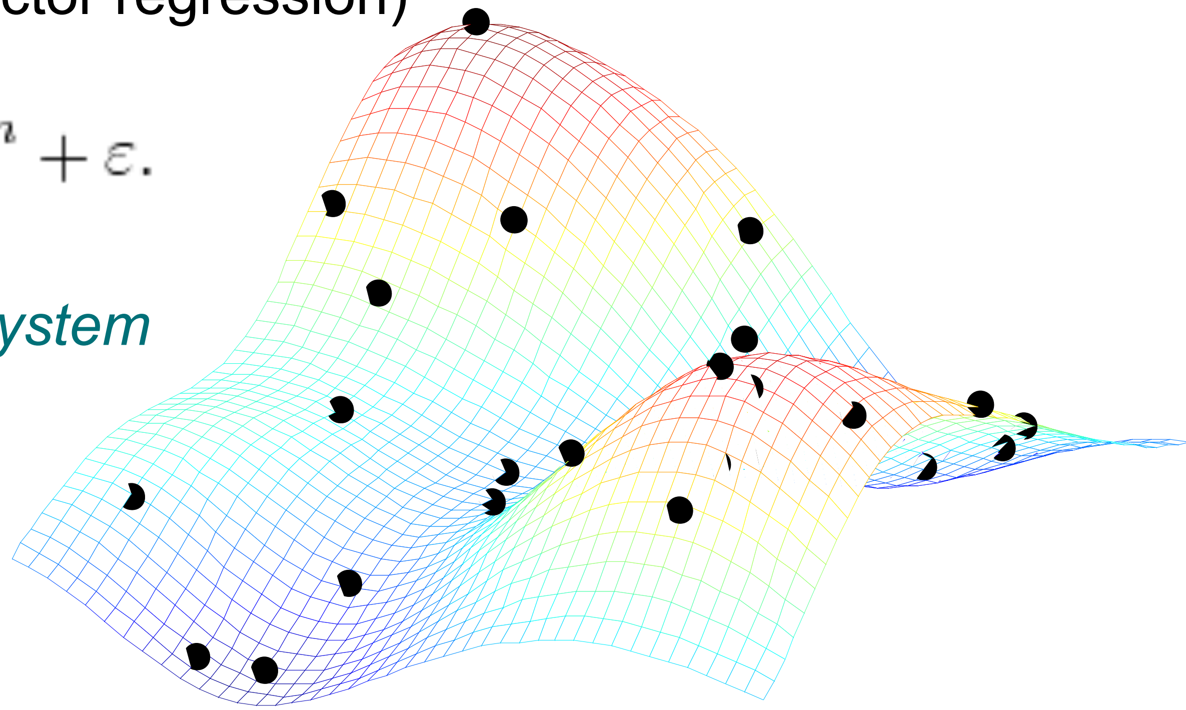
Source:
L. Leifsson, S. Koziel, Reykjavik University, Iceland

Functional vs. Physics-Based Surrogates

Function-Approximation Surrogates

- Suitable *approximations of sampled fine model data* (e.g., polynomial regression, kriging or support-vector regression)

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_m x^m + \varepsilon.$$



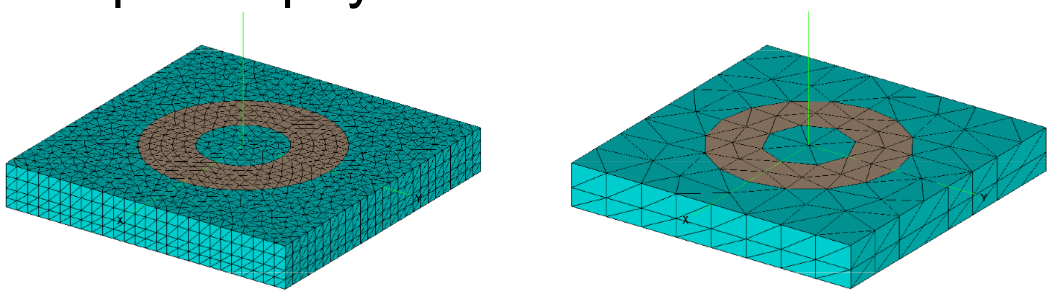
- Constructed *without particular knowledge of the system*
- Do not inherit any physical characteristics
- *Cheap* model evaluation
- But, typically requires *substantial amount of fine model data samples to set up a model* which ensures a good accuracy level
- Their use to ad-hoc optimization may be questionable
- Methodology is rather *generic* ⟶ *applicable to a wide class of problems*

(*) Picture Source: S. Koziel, Reykjavik University, Iceland

future ocean
KIEL MARINE SCIENCES

Physics-Based Surrogates

- Constructed from a *physics-based low-fidelity (or coarse)* model
  - Coarse discretization (clearly, numerical stability issues have to be taken into account)
  - Relaxed convergence criterion (e.g., in a fix-point iteration exploited for a steady-state simulation)
  - Simplified physics



E.g., some pseudo-timestepping scheme

$$\mathbf{y}' = f(\mathbf{y}, t)$$

$$\mathbf{y}_{j+1}(t, \mathbf{x}) = \mathbf{y}_j(t, \mathbf{x}) + \tau \cdot \Phi(t_j, \mathbf{y}_j, h)$$

- Coarse model *enjoys the same underlying physics*, it is typically able to predict the general behavior of the fine model

- However, their *accuracy is typically not sufficient* to directly exploit them in the optimization loop in lieu of the fine model

- *Suitable correction techniques* to yield a reliable surrogate are required (Space Mapping, Response Correction, SPRP)

(*) Picture Source: S. Koziel, Reykjavik University, Iceland

Functional vs. Physics-Based Surrogates

future ocean
KIEL MARINE SCIENCES

Physics-Based Surrogates

- *Fundamental advantage:*
  SBO schemes working with physics-based surrogates normally *require small number of fine model evaluations* to yield a sufficient accuracy (*often, only one* per iteration)

- Thus, the *computational burden is shifted towards the cheap coarse model*

- *Key prerequisites:*
  - Quality of the coarse model is critical ⟶ inaccurate model may result in poor algorithm performance
  - Cheap and yet reasonably accurate coarse model as well as a properly selected and low-cost alignment procedure
  - Agreement of function and derivative information (not necessarily exact)
  - Globalization: Some *standard trust-region/ line-search* approaches

- Underlying coarse model, correction approach is problem specific
  ⟶ Their reuse across different problems is rare

- *One of the most recognized SBO techniques* exploiting physics-based coarse models

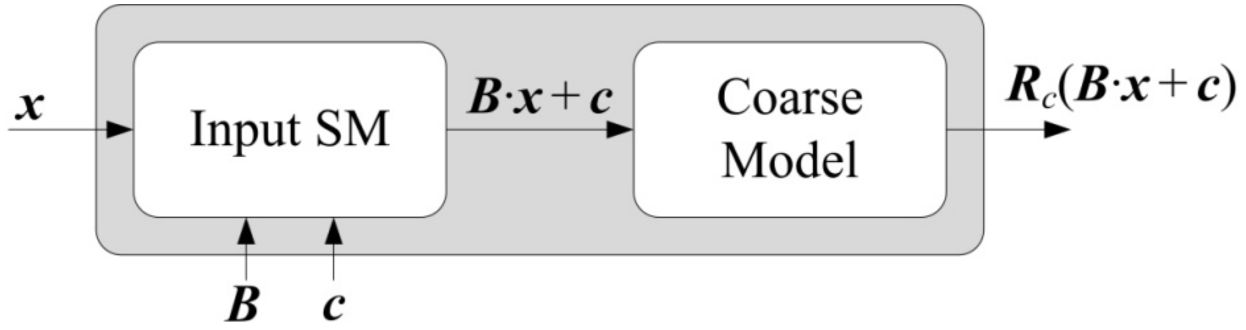
- A *mapping* relating the fine and coarse model parameters is *proposed to calibrate a physics-based coarse model*

- This mapping using so-called *parameter extraction (PE)* is a nonlinear opt. problem itself

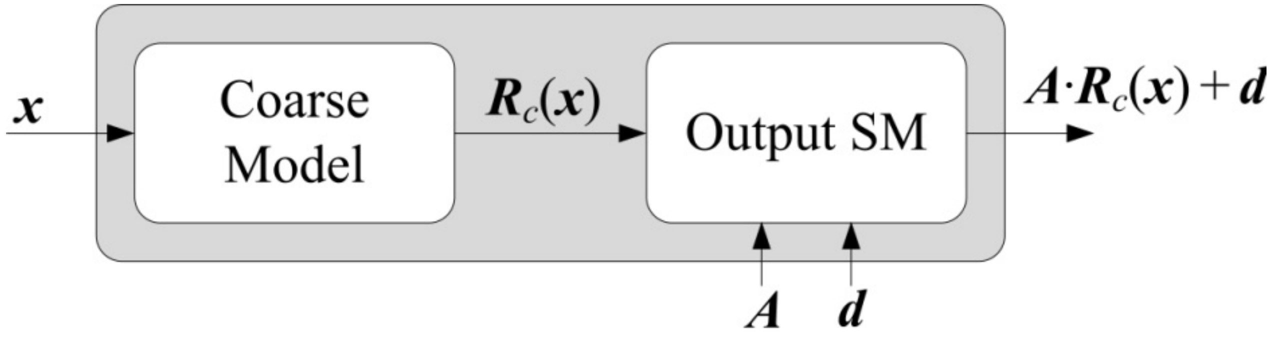
$$\mathbf{s}_k(\mathbf{u}) = \bar{\mathbf{y}}_c(\mathbf{u}, \mathbf{p}_k), \quad \mathbf{p}_k = \underset{\mathbf{p}}{\operatorname{argmin}} \left(\sum_{i=0}^{k} \|\mathbf{y}_f(\mathbf{u}_k) - \bar{\mathbf{y}}_c(\mathbf{u}_k, \mathbf{p})\| \right)$$

(Generic SM surrogate model, i.e., coarse model $\mathbf{y_c}$ with auxiliary mapping $\mathbf{p_k}$)

Domain distortion (*input SM*)



Response distortion (*output SM*)



(*) Picture Source: S. Koziel, Reykjavik University, Iceland

Response Correction (RC) Techniques

- One simple example of RC is *output SM* discussed before
- Yet another simple approach is a *Multiplicative Response Correction (MRC)* approach

$$\mathbf{a}_k := \frac{\mathbf{y}_f(\mathbf{u}_k)}{\mathbf{y}_c(\mathbf{u}_k)}, \quad k = 1, 2, \dots \qquad \mathbf{s}_k(\mathbf{u}) := \mathbf{a}_k\, \mathbf{y}_c(\mathbf{u}),$$

- *By definition*, the surrogate satisfies *agreement in function values*
- Since physics-based, its derivatives are expected to be at least similar to those of the fine model

$$\mathbf{s}_k(\mathbf{u}_k) = \mathbf{y}_f(\mathbf{u}_k), \quad \mathbf{s}'_k(\mathbf{u}_k) \approx \mathbf{y}'_f(\mathbf{u}_k).$$

- If required, exact *agreement in first-order information is „forced“* by an (optional) term *E as*

$$\mathbf{s}_k(\mathbf{u}) = \mathbf{a}_k\, \mathbf{y}_c(\mathbf{u}) + E_k\,(\mathbf{u} - \mathbf{u}_k)$$

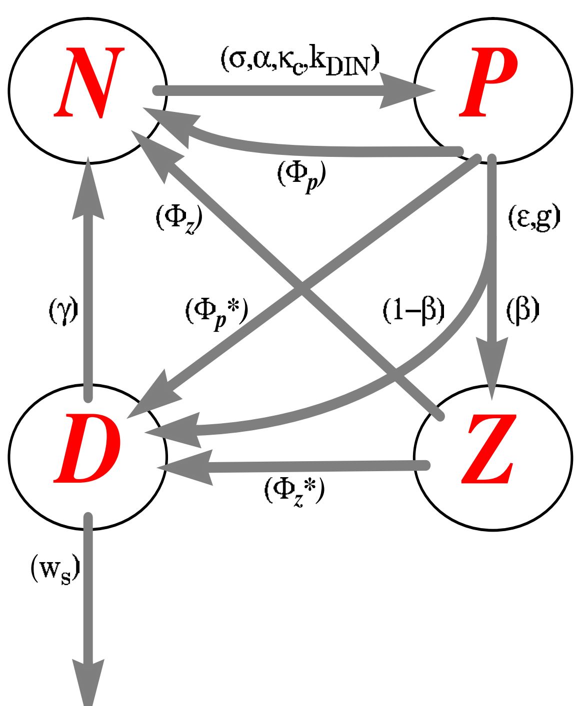
- *Clearly, trade-offs between the solution accuracy and the extra computational overhead related to sensitivity calculation have to be investigated*

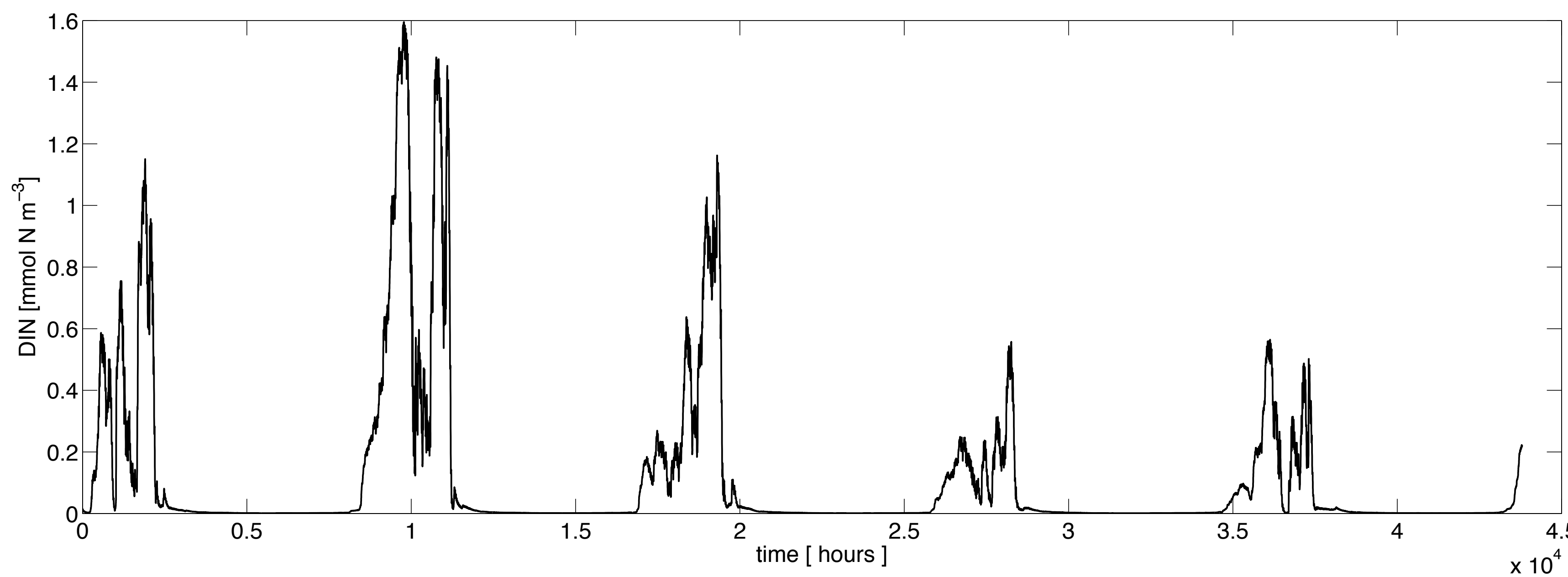
- *Nitrogen-budget* ecosystem model simulating the dynamical evolutions of four tracers, *dissolved inorganic nitrogen, phytoplankton, zooplankton, detritus* (Oschlies and Garcon, 1999; Schartau and Oschlies, 2003)

- One-dimensional, widely used model $$\frac{\partial y_i}{\partial t} = \partial_z (\kappa \, \partial_z y_i) + q_i(y, \mathbf{u}), \quad i = 1, \ldots, 4$$

- Ocean circulation data:
  Used for assembling the system matrices for the differential operators within the simulation

- Numerical solution of the underlying advective-diffusive reaction equations:
  Transient run with the time-dependent forcing data (Pseudo-timestepping scheme)





1D NPZD Model - Coarser Mesh Discretization

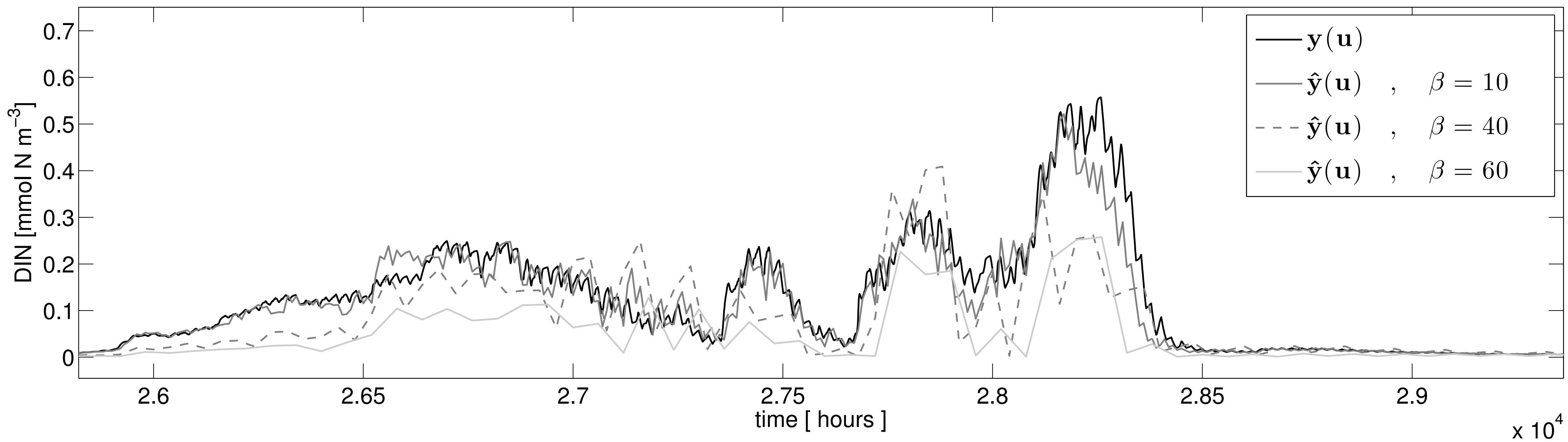


Figure 1: *High- and low-fidelity model output y, ŷ, respectively, for the state dissolved inorganic nitrogen at depth z ≈ −2.68 m for different values of the coarsening factor β and the same randomly chosen parameter vector u.*

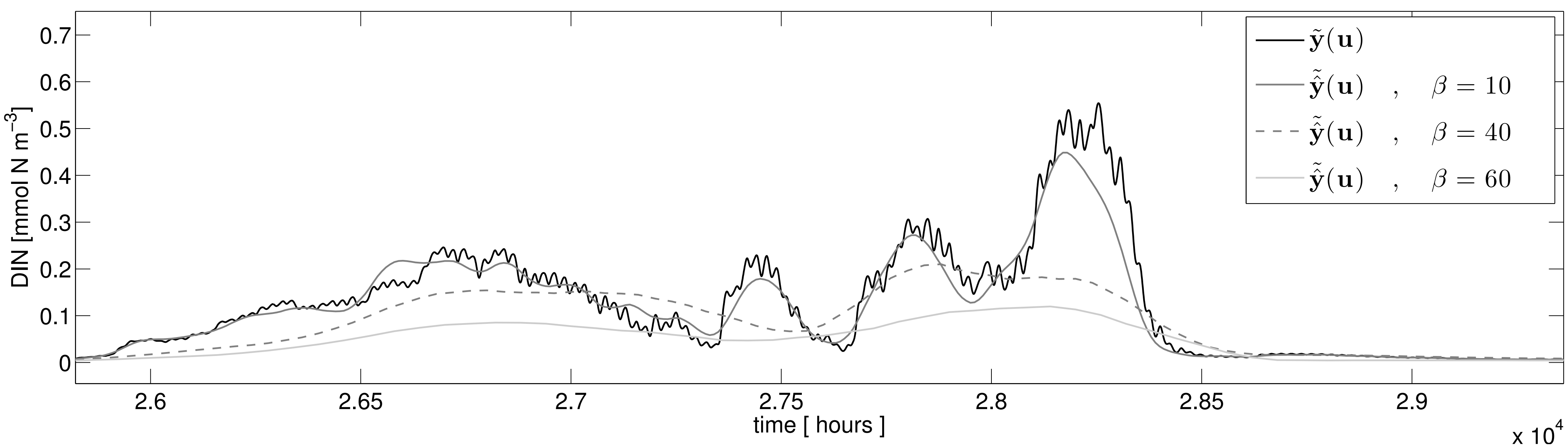


Figure 2: *Same as in Figure 2 but now using smoothing for both the coarse and the fine model. Resulting smoothed response contains the main characteristics of the fine one.*

1D NPZD Model - Generalization Capability

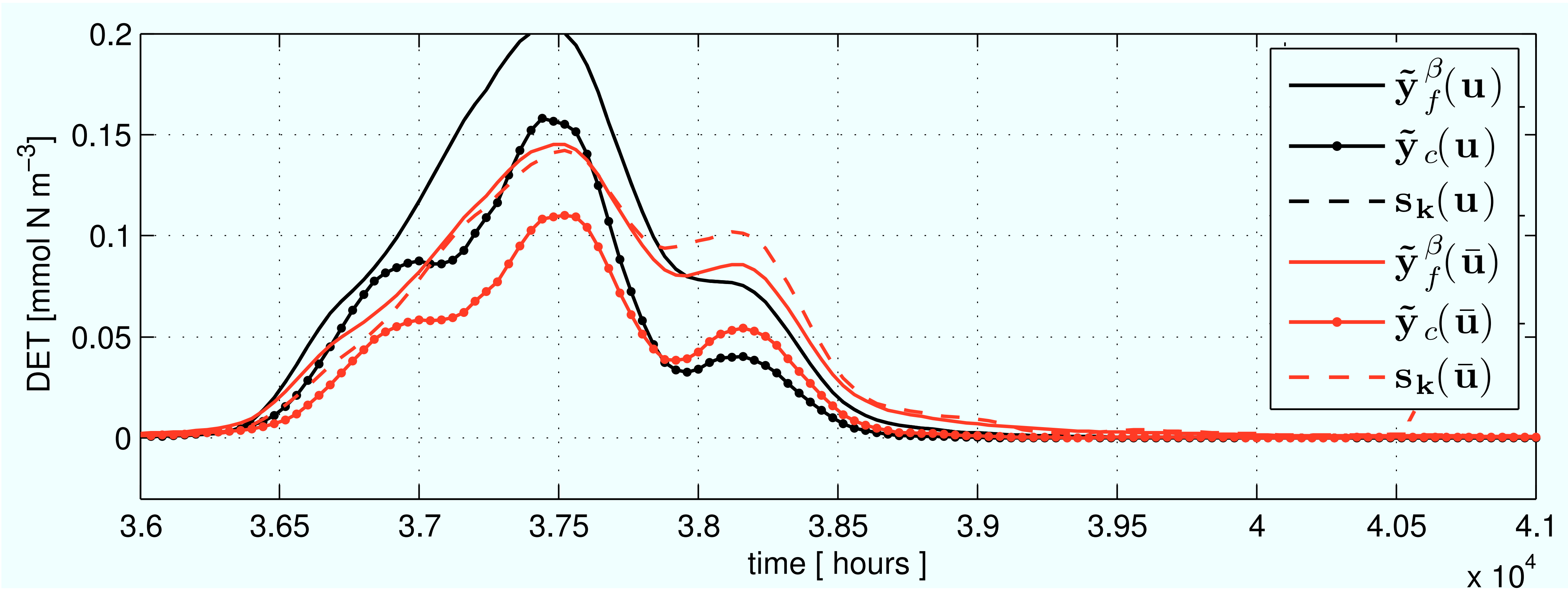


Figure 3: *Surrogate's, fine and coarse model output (some time intervall) for the state detritus at depth $z \approx -2.68$ m and at two iterates* $\mathbf{u}$ *and in a neighbourhood* $\bar{\mathbf{u}}$.

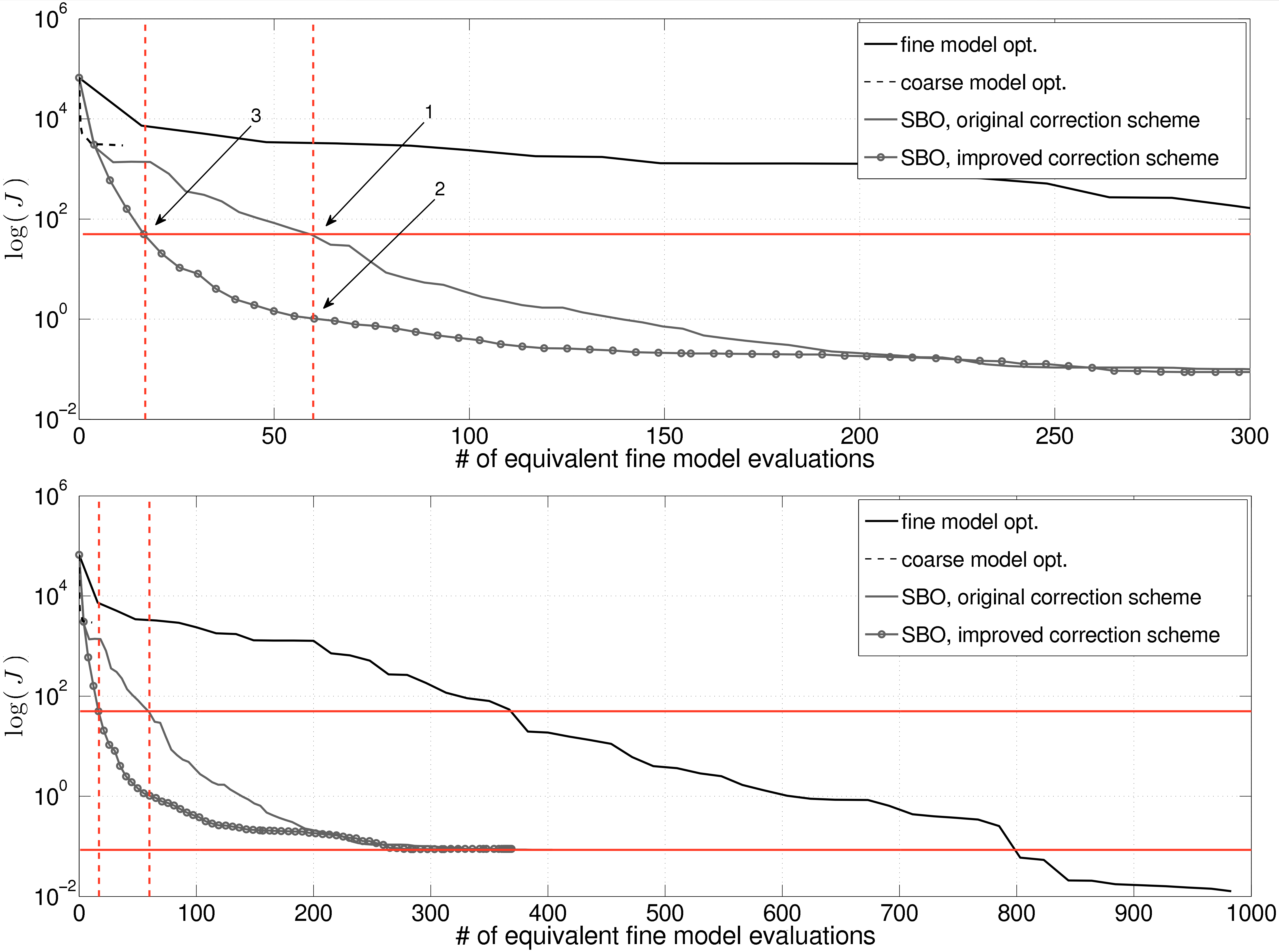


Figure 4: Convergence history - fine, coarse and surrogate optimization - for illustrative optimization runs.

3D N-DOP Model

- Dynamical evolutions among *nitrogen* and *dissolved organic phosphorus* (Kriest et al. 2010; Dutkiewicz et al. (2005); Parekh et al. (2005); Yamanaka and Tajika, 1997)

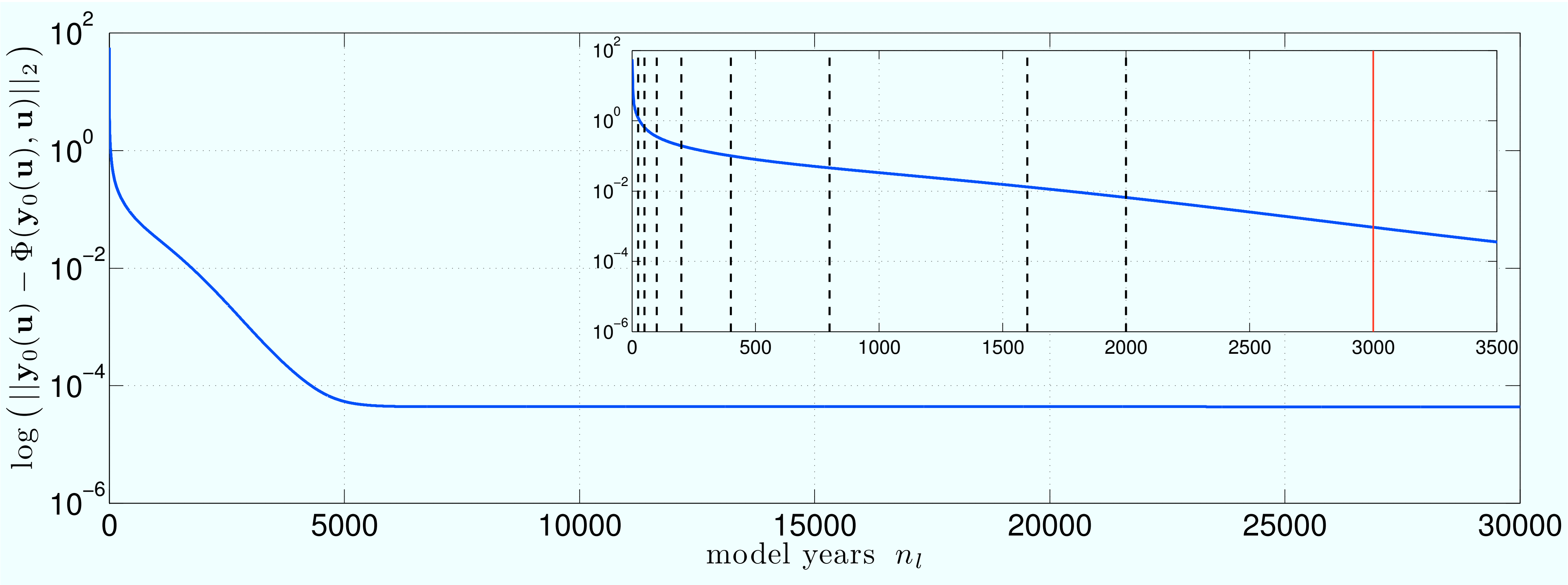
- Coupled to a general ocean circulation model in an *off-line mode*, exploting the *Transport Matrix Method (TMM)*

$$\begin{aligned} \mathbf{y}_{j+1} &= \mathbf{A}_{imp,j}\,(\mathbf{A}_{exp,j}\,\mathbf{y}_j + \tau\,\mathbf{q}_j(\mathbf{y}_j,\mathbf{u})) \\ &=: \varphi_j(\mathbf{y}_j,\mathbf{u}), \qquad j = 0,\ldots,n_\tau - 1 \end{aligned}$$

- The TTM is applied to simulate a *steady annual cycle* (with an initial spin-up)

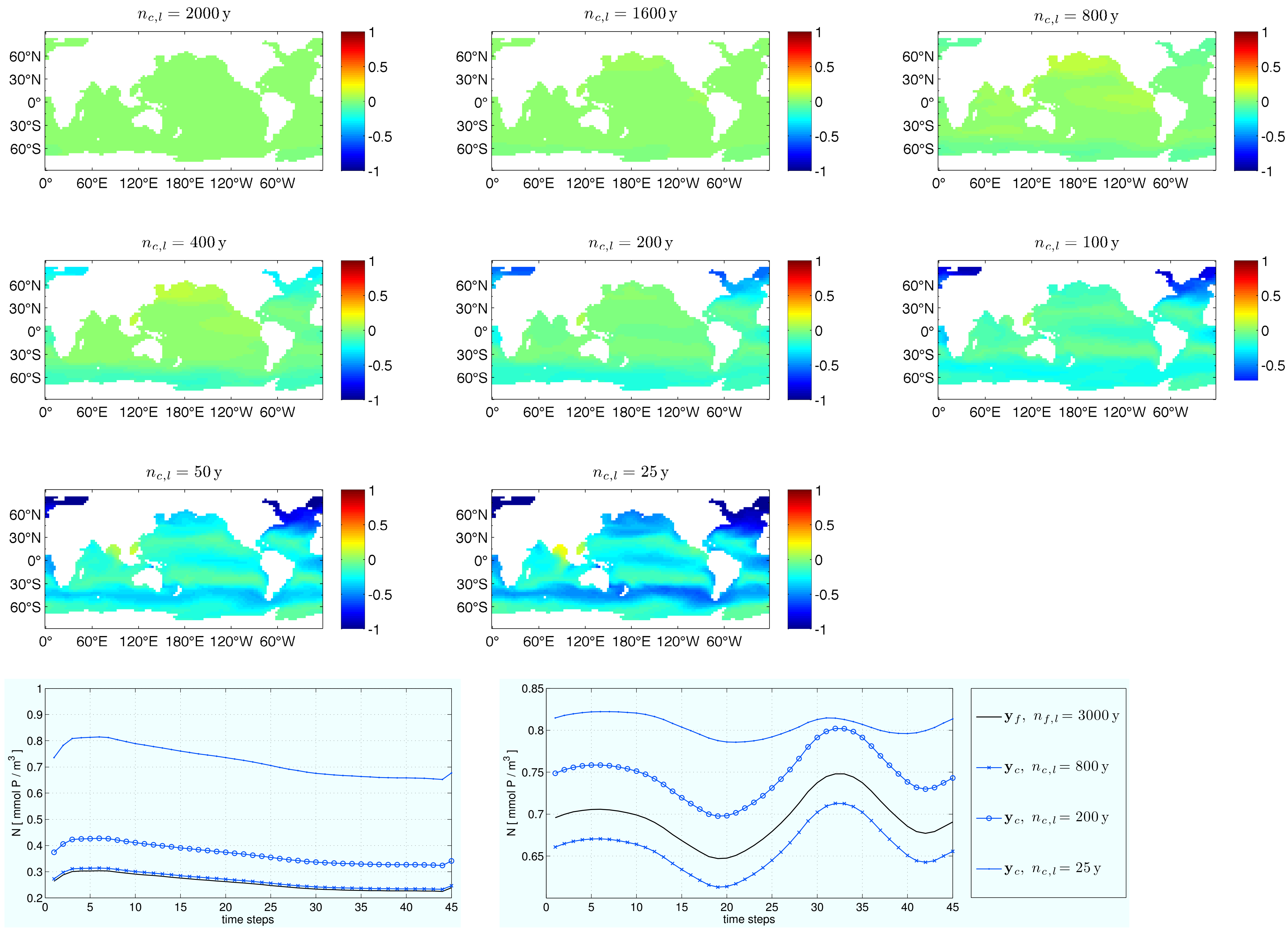
$$\mathbf{y}_{n_\tau} = \Phi(\mathbf{y}_0,\mathbf{u}) = \mathbf{y}_0 \qquad \Phi := \varphi_{n_\tau-1} \circ \cdots \circ \varphi_0$$

- Here, we use a *classical fixed point iteration* $\mathbf{y}^{l+1} = \Phi(\mathbf{y}^l,\mathbf{u}), \quad l = 0,\ldots,n_l - 1$



- *„Sufficient“ accuracy after 3000 iterations* (converged „reference“ fine model)

3D N-DOP Model - Relaxed Convergence Criterion

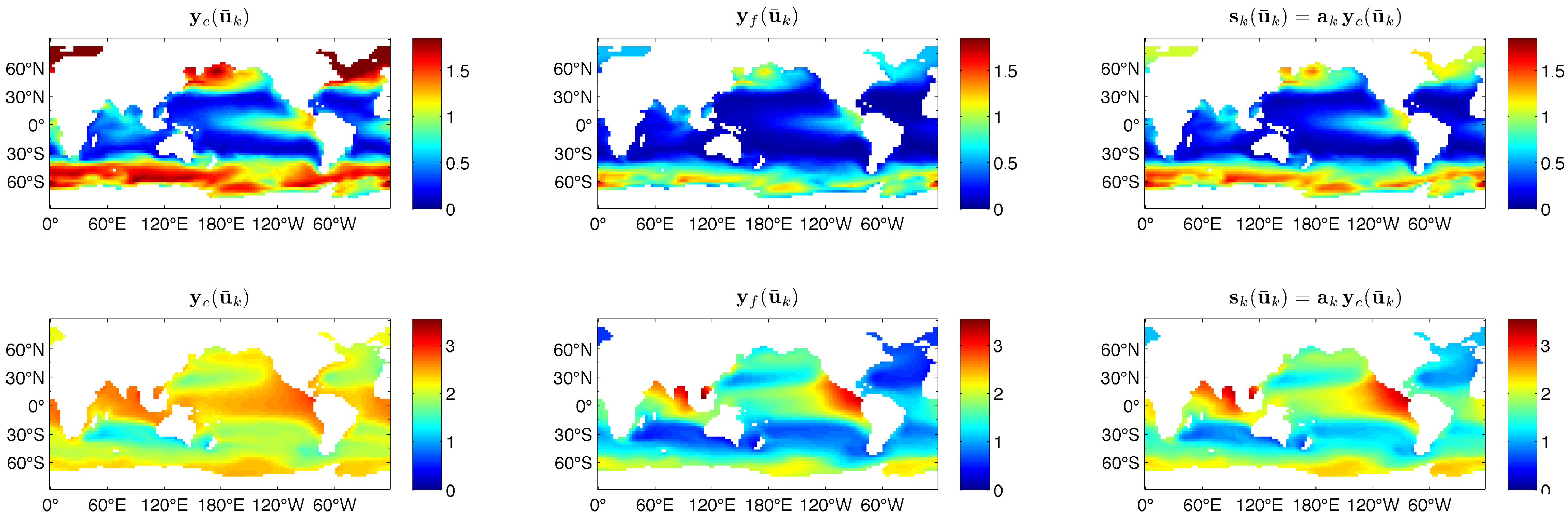


(a) $x = 30.9375°\mathrm{W}, y = 30.9375°\mathrm{N}$.

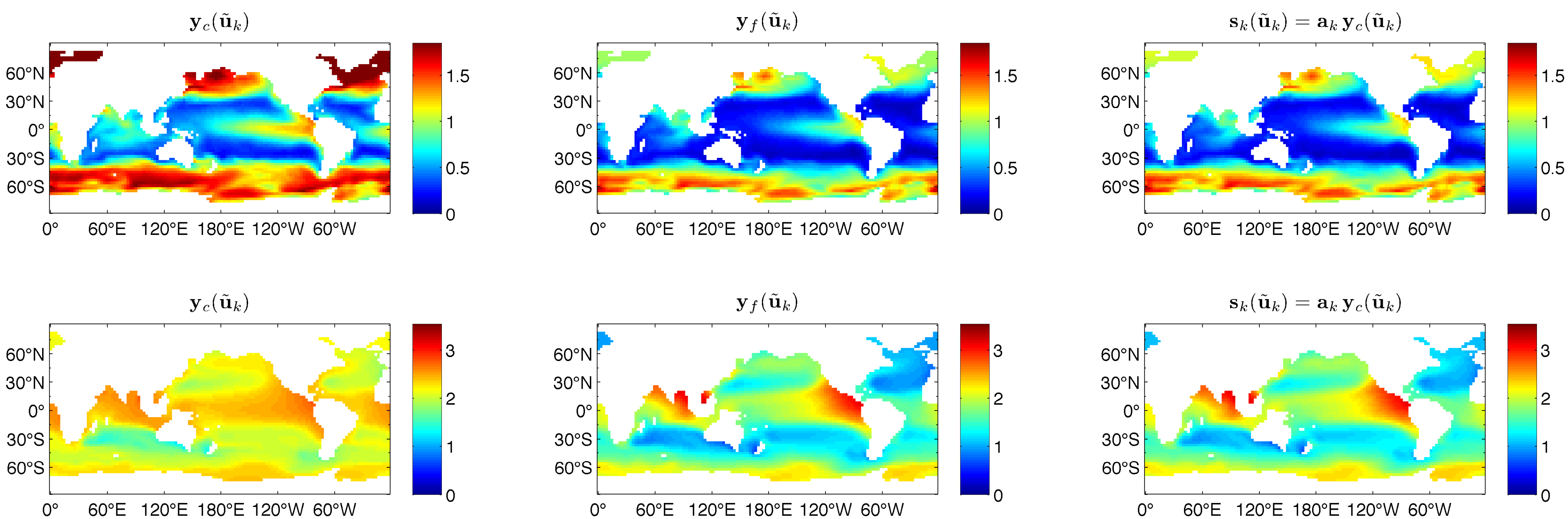
(b) $x = 90°\mathrm{E}, y = 0°$.

Figure 5: Upper: Difference in fine and coarse model responses. Lower: Responses at distinct spatial locations

3D N-DOP Model - Generalization Capability



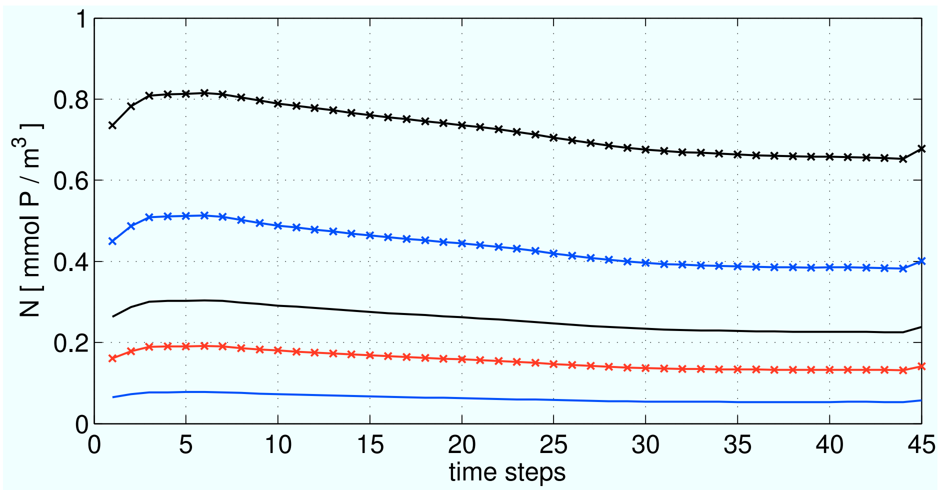
(a) Responses at neighboring point $\bar{\mathbf{u}}_k$.



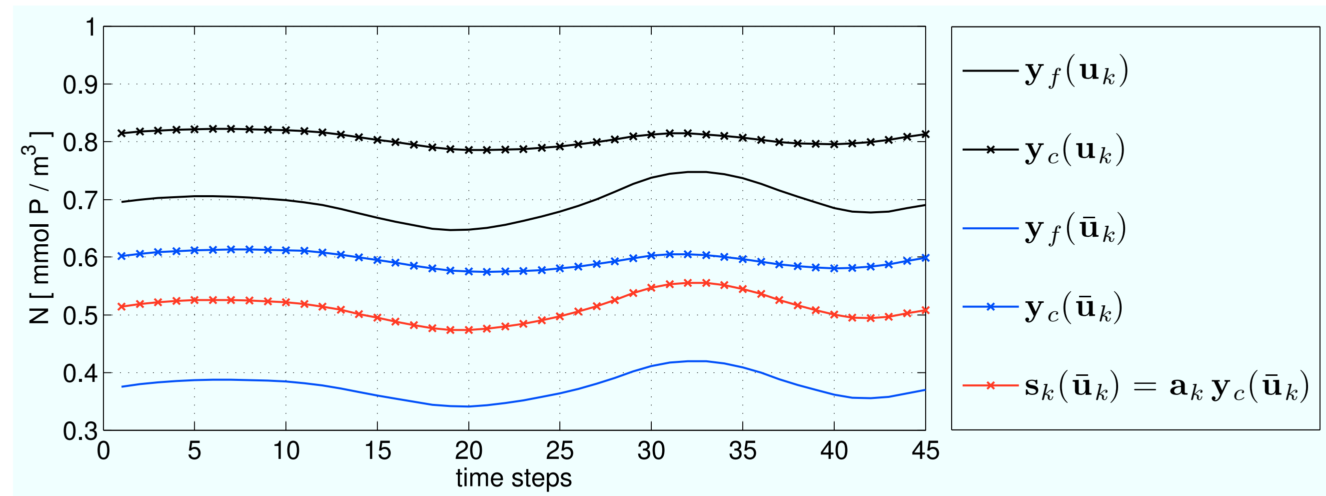
(b) Responses at closer point $\tilde{\mathbf{u}}_k$.

Figure 6: Coarse, fine and surrogate‘s response at some neighboring point (upper) and in an even closer vicinity (lower)

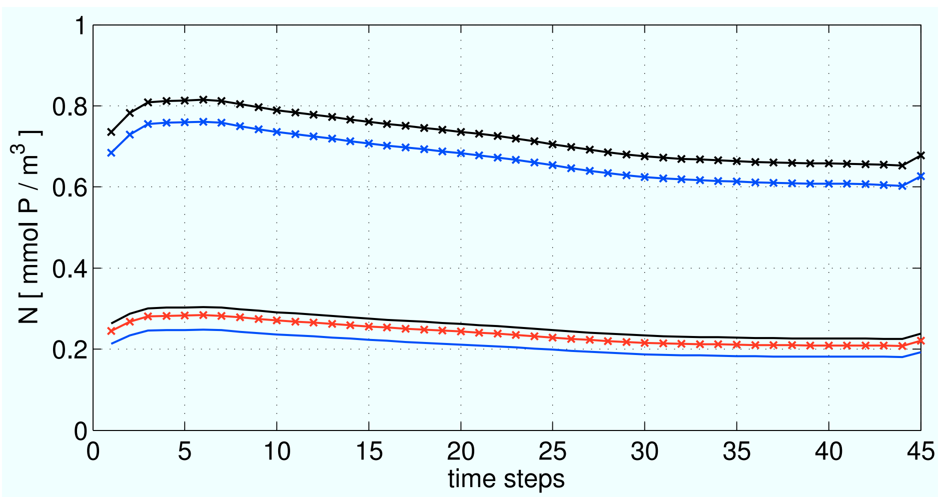
3D N-DOP Model - Generalization Capability



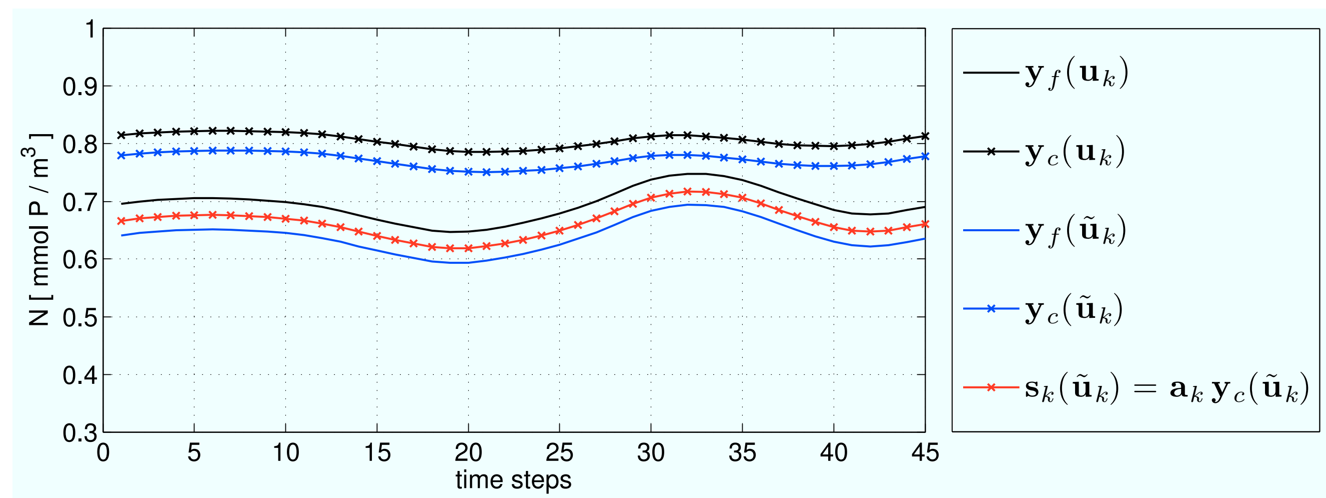
(a) $x = 30.9375°\text{W}, y = 30.9375°\text{N}$.



(b) $x = 90°\text{E}, y = 0°$.



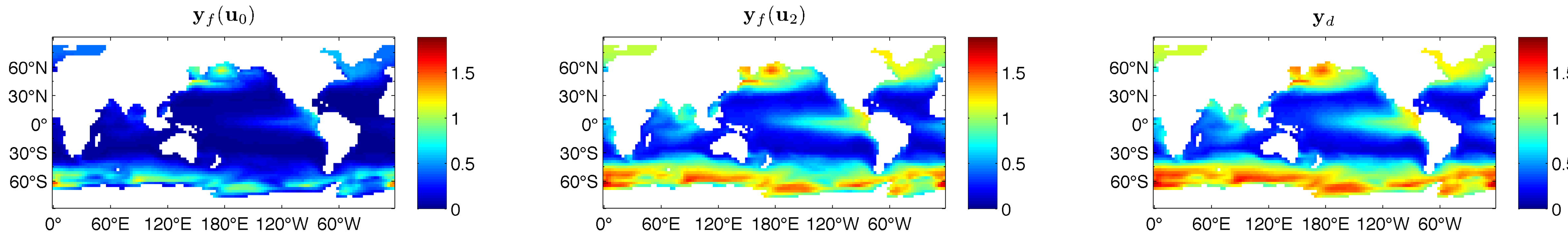
(c) $x = 30.9375°\text{W}, y = 30.9375°\text{N}$.



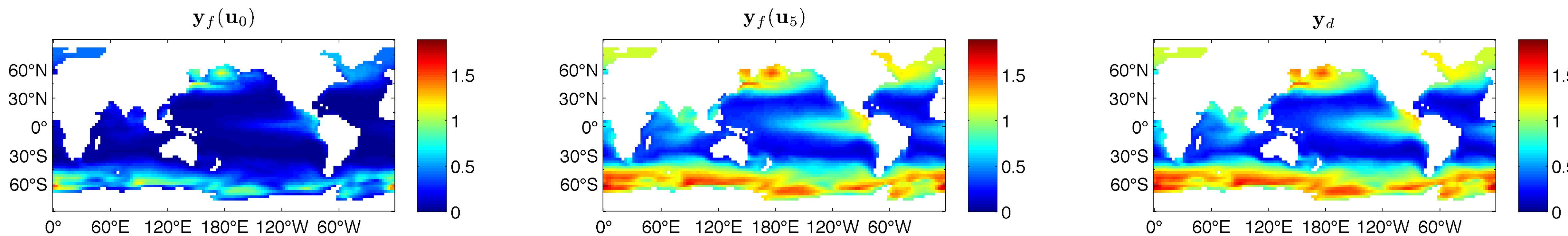
(d) $x = 90°\text{E}, y = 0°$.

Figure 7: Same as in Figure 6, here, for two distinct spatial locations.

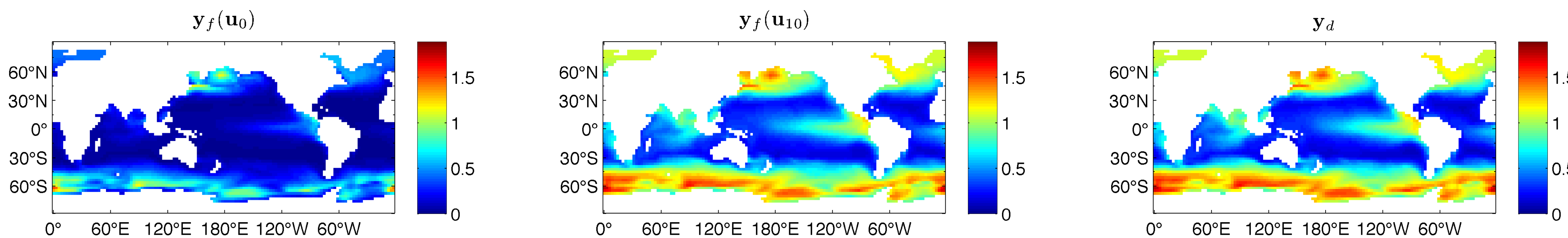
3D N-DOP Model - Illustrative SBO Run



(a) $\gamma = 10^{-1}$.



(b) $\gamma = 10^{-2}$.



(c) $\gamma = 10^{-4}$.

Figure 8: Results of an illustrative optimization run. (a)-(c) correspond to different stopping criteria (here, step size).

3D N-DOP Model - Illustrative SBO Run

future ocean
KIEL MARINE SCIENCES

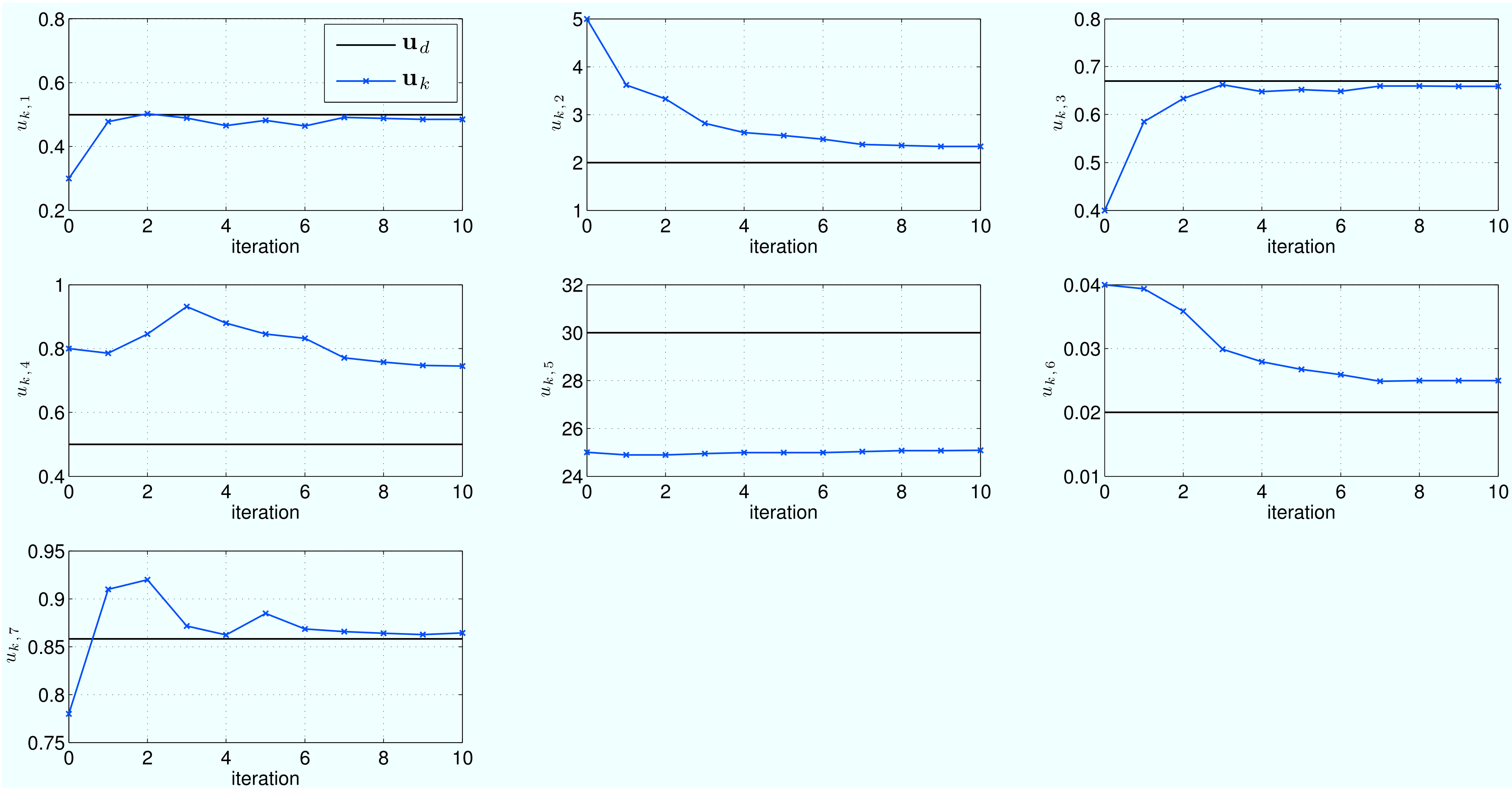


Figure 9: Optimization history of the parameters for the SBO run.

- Assuming approximately 30 minutes for a single fine mode evaluation on a 48-processor cluster, a *direct optimization approach could require about 15 days.*

- On the other hand, the whole *SBO run requires approximately 4 to 23 hours* (depending on the required accuracy and hence number of iterations performed).

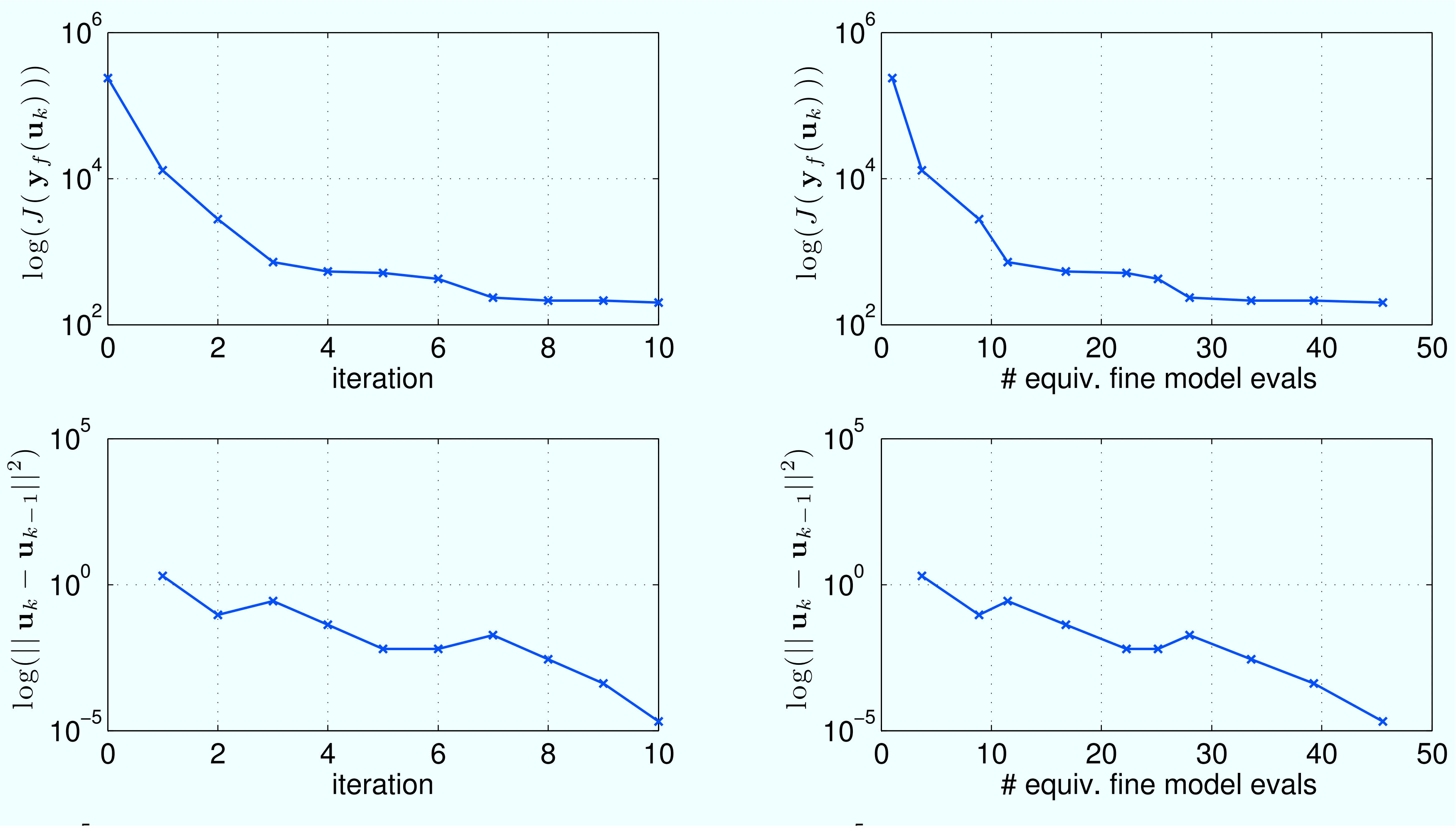


Figure 10: Optimization history of the cost function values and step size norm for the SBO run.

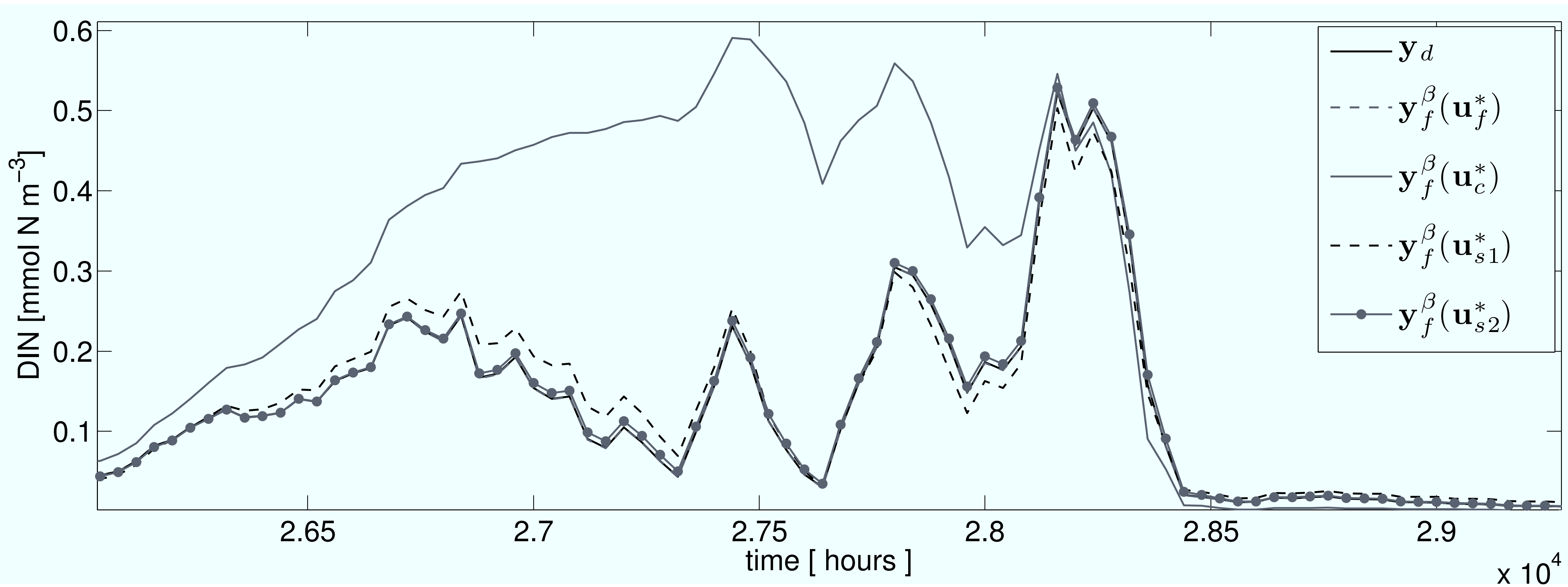
Summary

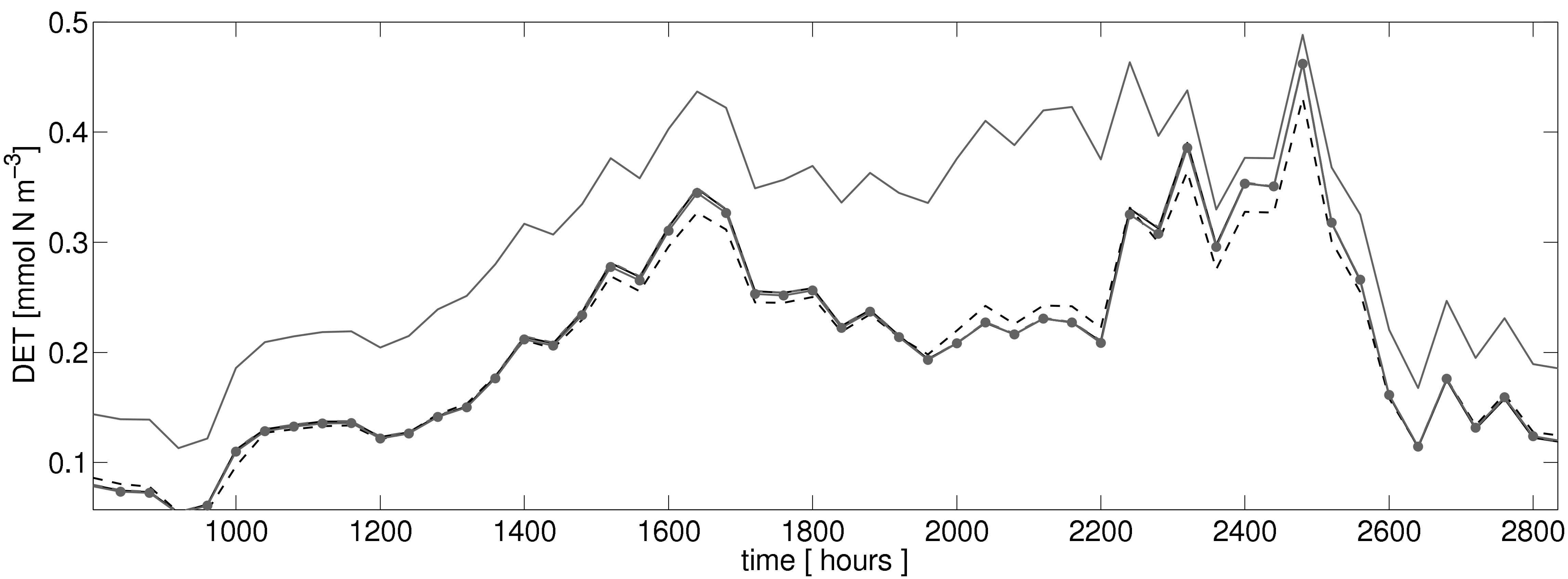
- We presented an *efficient optimization methodlogy* for computationally heave (nonlinear) optimization problems
- Presented SBO exploits *physics-based low-fidelity (or coarse) models*
  - Coarser mesh discretization (1D NPZD model)
  - Relaxed convergence criterion (3D N-DOP model using TMM)
- Coarse model *accuracy is typically not sufficient* to directly exploit them in the optimization loop in lieu of the fine model
- Introduced two popular correction approaches: *Space Mapping, Response Correction*
- MRC:
  - rather *„intuitive“* and *straightforward* RC approach
  - *yet, very powerful*
- SBO with MRC yields a sufficiently accurate solution at a cost of a few fine model evaluations
- *Cost savings are significant*, about *84% and more* when compared to a direct fine model optimization

Acknowledgements & References

- Prof. Slawomir Koziel - *Engineering Optimization & Modeling Center, School of Science and Engineering*, Reykjavik University (koziel@ru.is)
- Prof. Andreas Oschlies - *IFM-GEOMAR*, Kiel (aoschlies@ifm-geomar.de)

- N. V. Queipo, R. T. Haftka, W. Shyy, T. Goel, R. Vaidyanathan, and P. K. Tucker, “Surrogate-based analysis and optimization,” Prog. Aerosp. Sci., vol. 41, no. 1, pp. 1 – 28, 2005
- A. I. Forrester and A. J. Keane, “Recent advances in surrogate-based optimization,” Prog. Aerosp. Sci., vol. 45, no. 1-3, pp. 50 – 79, 2009
- J. W. B, Q. S. Cheng, S. A. Dakroury, A. S. Mohamed, M. H. Bakr, K. Madsen, and J. Søndergaard, “Space mapping: The state of the art,” IEEE T. Microw. Theory., vol. 52, no. 1, 2004
- S. Koziel, J. Bandler, and Q. Cheng, “Robust trust-region space-mapping algorithms for microwave design optimization,” IEEE T. Microw. Theory., vol. 58, pp. 2166 –2174, Aug. 2010.
- Kriest, I., Khatiwala, S., Oschlies, A., 2010. Towards an assessment of simple global marine biogeochemical models of different complexity. Progress In Oceanography 86, 337–360.
- M. Priess, T. Slawig, and S. Koziel, “Improved surrogate-based optimization of climate model parameters using response correction,” *Int. Conf. Simulation and Modeling Methodologies, Technologies and Appl., SIMULTECH 2011*, Noordwijkerhout, The Netherlands, July 29-31, pp. 449-457, 2011
- Parekh, P., Follows, M.J., Boyle, E.A., 2005. Decoupling iron and phosphate in the global ocean. Global Biogeochemical Cycles 19.
- A. Oschlies and V. Garcon, “An eddy-permitting coupled physical-biological model of the north atlantic. 1. sensitivity to advection numerics and mixed layer physics,” Global Biogeochem. Cy., vol. 13, pp. 135– 160, 1999.


DIN [mmol N m^{-3}]
time [hours]
x 10^4
$\mathbf{y}_d$
$\mathbf{y}_f^\beta(\mathbf{u}_f^*)$
$\mathbf{y}_f^\beta(\mathbf{u}_c^*)$
$\mathbf{y}_f^\beta(\mathbf{u}_{s1}^*)$
$\mathbf{y}_f^\beta(\mathbf{u}_{s2}^*)$


DET [mmol N m^{-3}]
time [hours]