Algorithmic Optimal Control - CO_2 Uptake of the Ocean

Junior Research Group A3

Surrogate-Based Optimization of Climate Model Parameters

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

Prof. Thomas Slawig

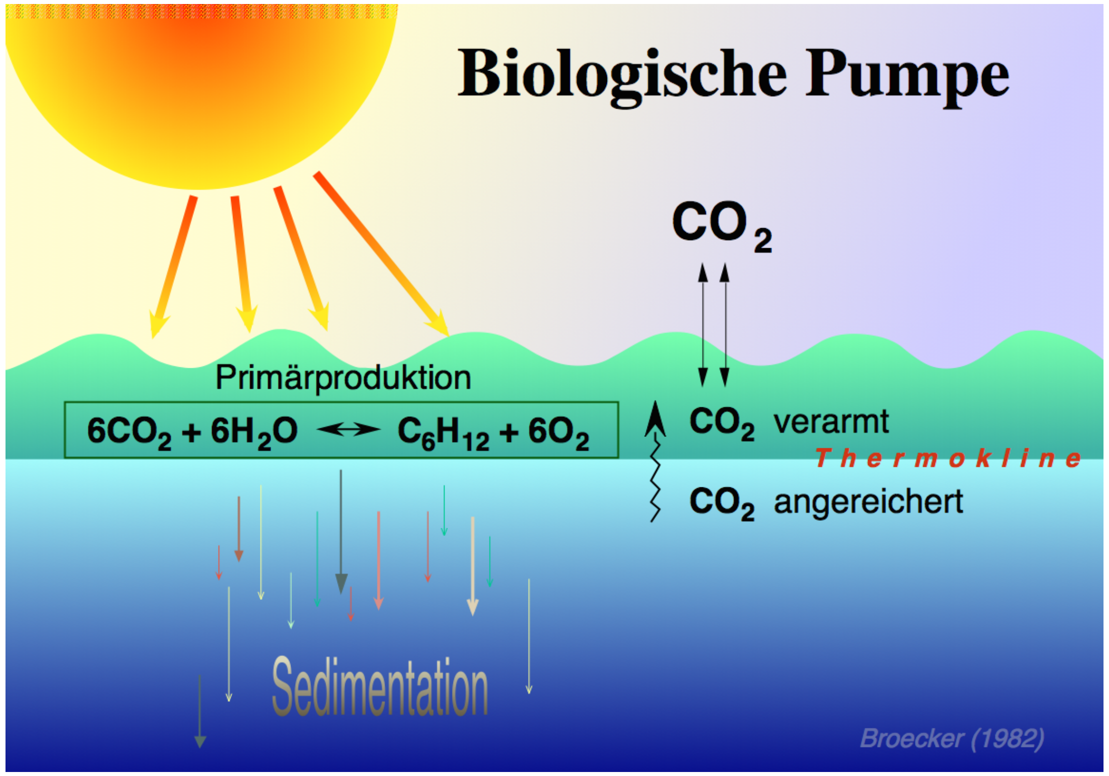
Mustapha El Jarbi, Anna Heinle,
Claudia Kratzenstein, Henrike Mütze,
Jana Petersen, Jaroslaw Piwonski,
Malte Prieß, Johannes Rückelt

PIs: Andreas Oschlies IFM GEOMAR
Anand Srivastav CAU Kiel Institut für Informatik

OP11 - SIAM Conference on Optimization

16-19/05/2011 - Darmstadt

Overview



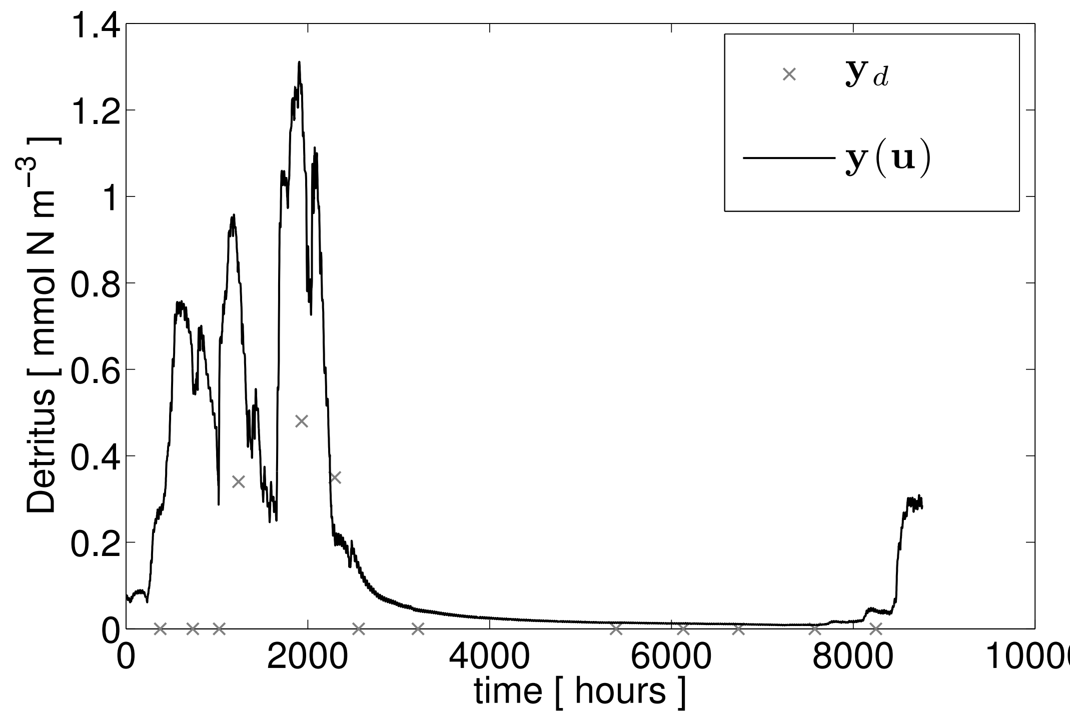
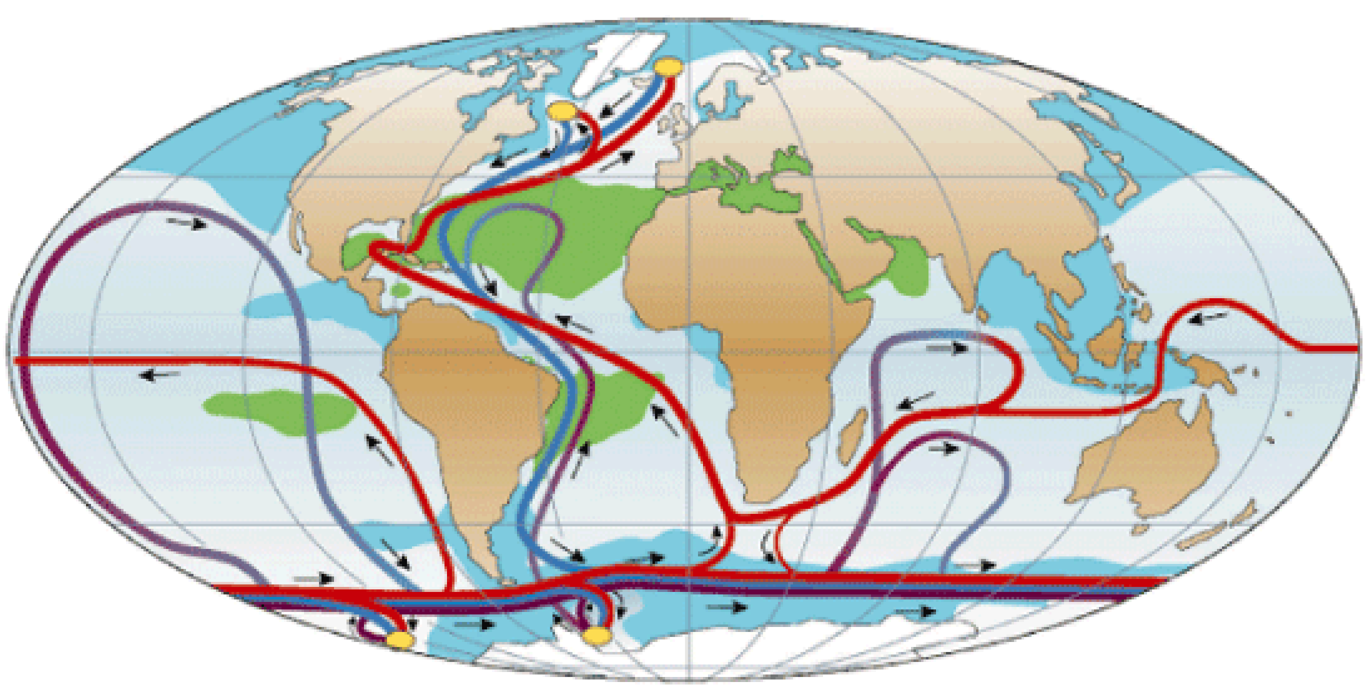
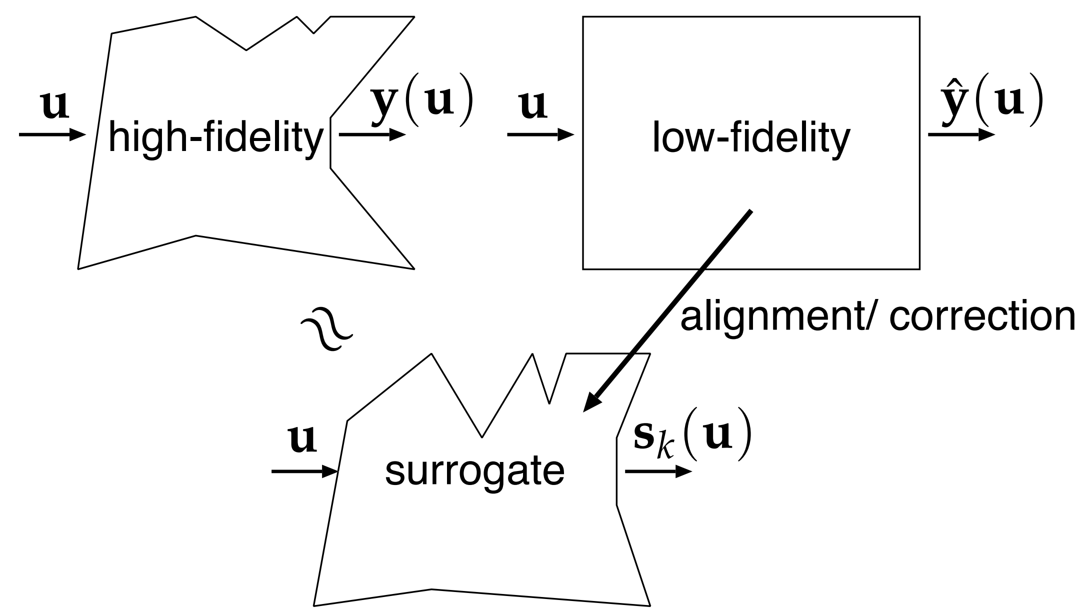


Figure 1: Model output $\boldsymbol{y}^{(D)}$ (detritus) and target *data* $\boldsymbol{y}_d$ *for one year at depth z ≃ −25 m.*





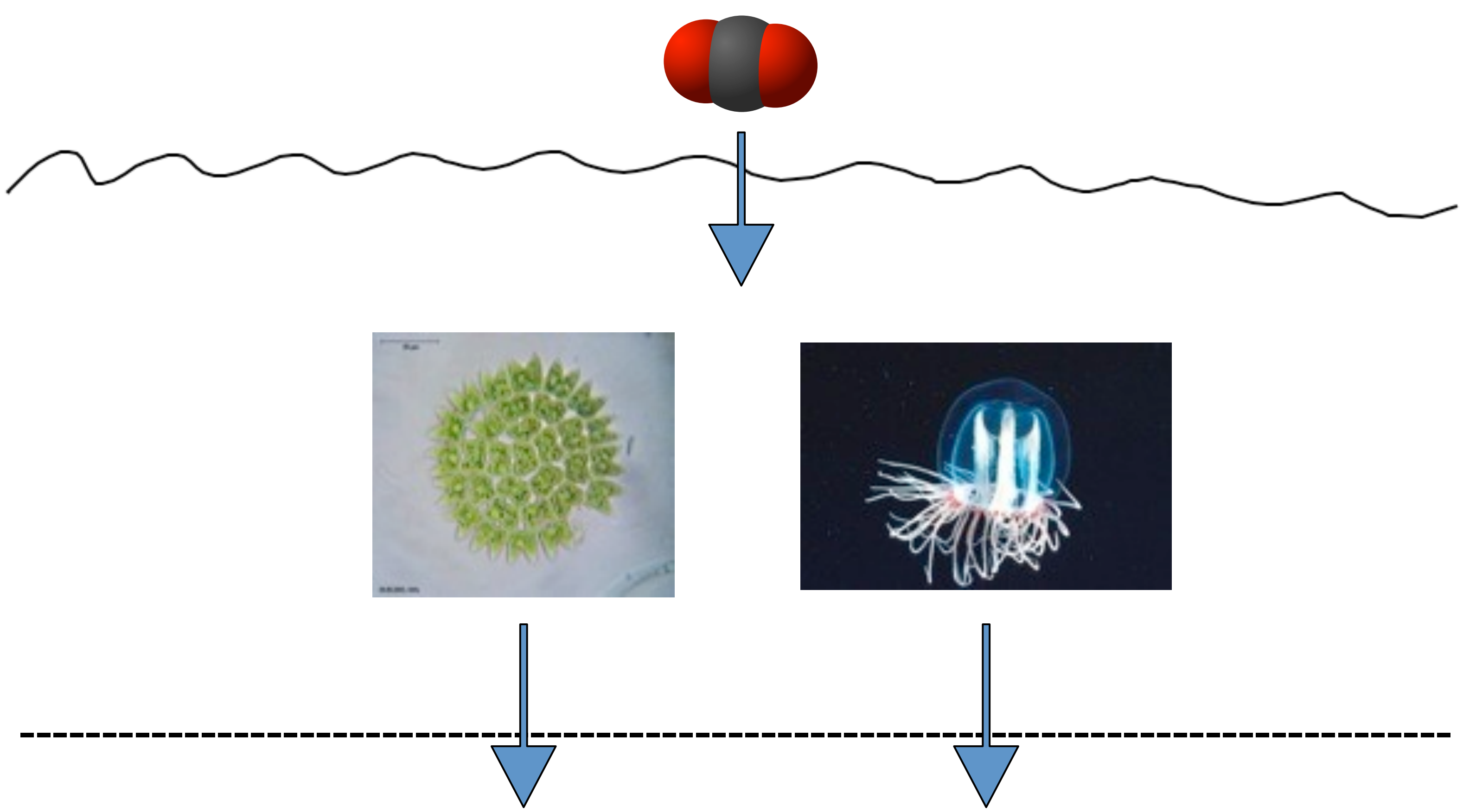
Climate Models - A General Formulation

- Initial boundary value problem (*IBVP*) for a system of time-dependent partial differential or differential algebraic equations (*PDEs/DAEs*) of the following form:

$$
\begin{aligned}
E\,\frac{\partial y}{\partial t} &= f(y,u) && \text{in } \Omega \times (0,T) \\
y(x,0) &= y_{init}(x) && \text{in } \Omega \\
y(x,t) &= y_{bdr}(x,t) && \text{on } \partial\Omega \times (0,T)
\end{aligned}
$$

Marine Ecosystem Models - Some Motivations

- Used for example to compute and predict the *oceanic uptake of CO_2* as part of the *global carbon cycle*
- The uptake is determined by the solution of CO_2 in the water via the ocean surface ...
- ... and *physical and biogeochemical processes* in the water, i.e.
  - *Ocean circulation*
  - *Photosynthesis*, consumption by *zooplankton, sinking of dead material* (which exports the carbon to the deeper ocean)



Marine Ecosystem Models - One Representative Example

- Although *one-dimensional*, the following example illustrates the general formulation of this type of models and *actually provides the basis for many marine ecoystem models (also 3D)*

- Model is of so-called *NPZD type*:

  Concentrations of the tracers *dissolved inorganic nitrogen* ***N****, phytoplankton* ***P****, zooplankton* ***Z****, and detritus (i.e., dead material)* ***D*** are simulated in a water column, $y = (y^{(l)})_{l=N,P,Z,D}$

The Optimization Problem

- Adjust/identify model parameters ***u*** such that *given measurement data* $\mathbf{y}_d$ *is matched by the model output* $\mathbf{y}(\mathbf{u})$

- The mathematical task thus can be classified as a *least-squares type optimization or inverse problem*

- The opt. process requires a *substantial number of (typically expensive) function evaluations*

- Methods that aim at *reducing the optimization cost* (e.g. surrogate-based optimization), are *highly desirable*

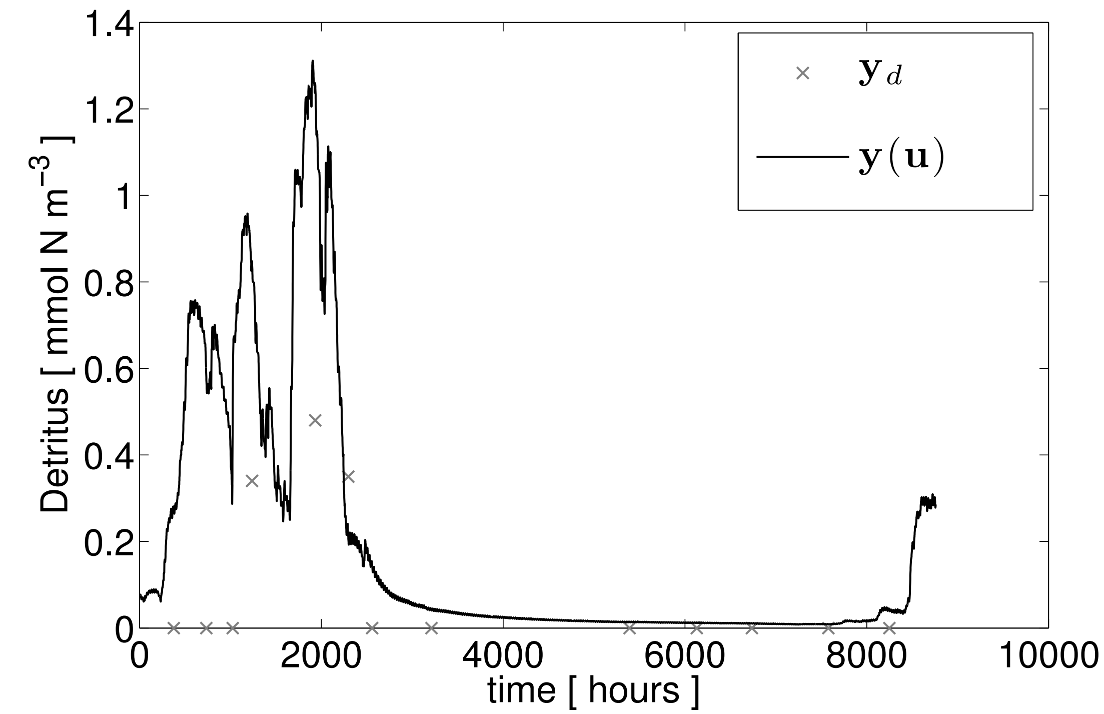


Figure 1: Model output $\mathbf{y}^{(D)}$ (detritus) and target *data* $\mathbf{y}_d$ *for one year at depth* $z \simeq -25$ *m.*

$$\min_{\mathbf{u} \in U_{ad}} J(\mathbf{y}(\mathbf{u})) \tag{1}$$

$$J(\mathbf{y}) := \frac{1}{2} \|\mathbf{y} - \mathbf{y}_d\|_Y^2 \quad , \quad U_{ad} := \{\mathbf{u} \in \mathbb{R}^n : \mathbf{b}_l \leq \mathbf{u} \leq \mathbf{b}_u\} \quad , \quad \mathbf{b}_l, \mathbf{b}_u \in \mathbb{R}^n \quad , \quad \mathbf{b}_l < \mathbf{b}_u .$$

Surrogate-Based Optimization (SBO)

- *Complex* (so-called *high-fidelity*) *models* often require *substantial computational effort* already for a model evaluation

- As a consequence, opt. and *control problems for such models are often still beyond the capability of modern numerical algorithms* and computer power

- Idea: *replace the high-fidelity in focus by a computationally cheaper and yet reasonably accurate representation*, so-called surrogate

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

Physically based:

Constructed from physical low-fidelity model (with *suitable correction/alignment*)

Pro:

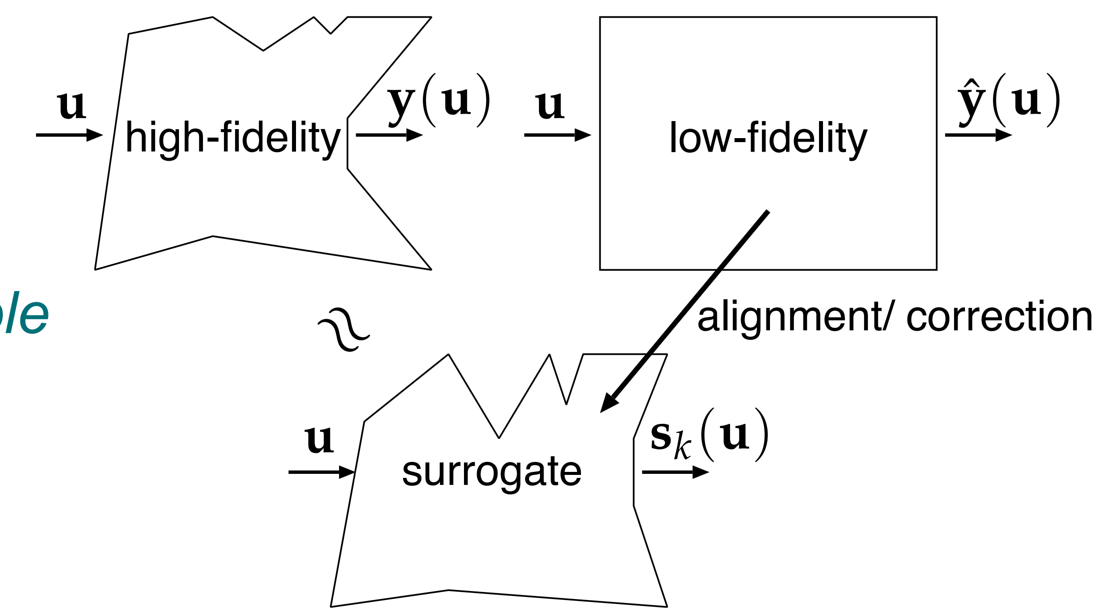
Inherits more characteristics of the system

Contra:

Dedicated (reuse is rare)

Typically more expensive

Low-fidelity model must be available



Popular techniques:

Response correction, Space Mapping

How to obtain the low-fidelity model?

- Using simplified physics (e.g., ignoring second order effects)
- *Coarse discretization*
- Using analytical formulas if available

Low-Fidelity Model - One Example

- Discretized model equation of our *high-fidelity model*:

$$\underbrace{\left[I - \tau A_j^{\text{diff}}\right]}_{:=B_j^{\text{diff}}} \mathbf{y}_{j+1} = \underbrace{\left[I + \tau A^{\text{sink}}\right]}_{:=B^{\text{sink}}} B_j^Q \circ B_j^Q \circ B_j^Q \circ B_j^Q(\mathbf{y}_j),$$

$$B_j^Q(\mathbf{y}_j) \quad := \quad \left[\mathbf{y}_j + \frac{\tau}{4} Q_j(\mathbf{y}_j)\right] \qquad \mathbf{y}_j = (y_{ji})_{i=1,\dots,I}, \quad j = 1, \dots, M$$

(M = # of discrete temporal points of the fine model, I = # of discrete spatial points)

Raw and Smoothed Model Responses

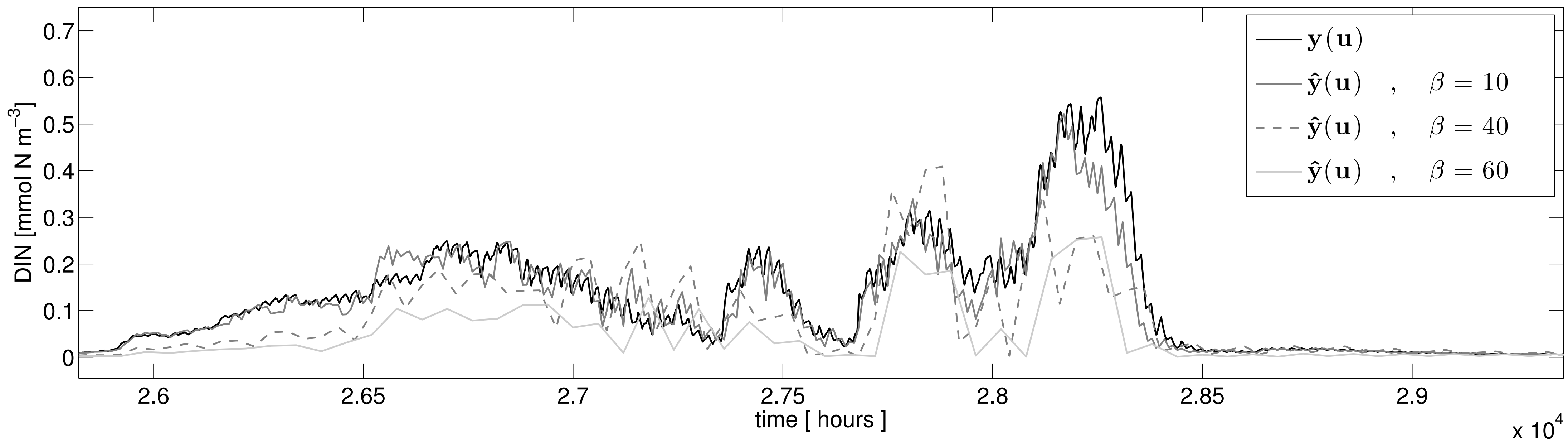


Figure 2: *High- and low-fidelity model output* $\mathbf{y}$, $\hat{\mathbf{y}}$, *respectively, for the state dissolved inorganic nitrogen at depth* $z \approx -2.68$ *m for different values of the coarsening factor* β *and the same randomly chosen parameter vector* $\mathbf{u}$.

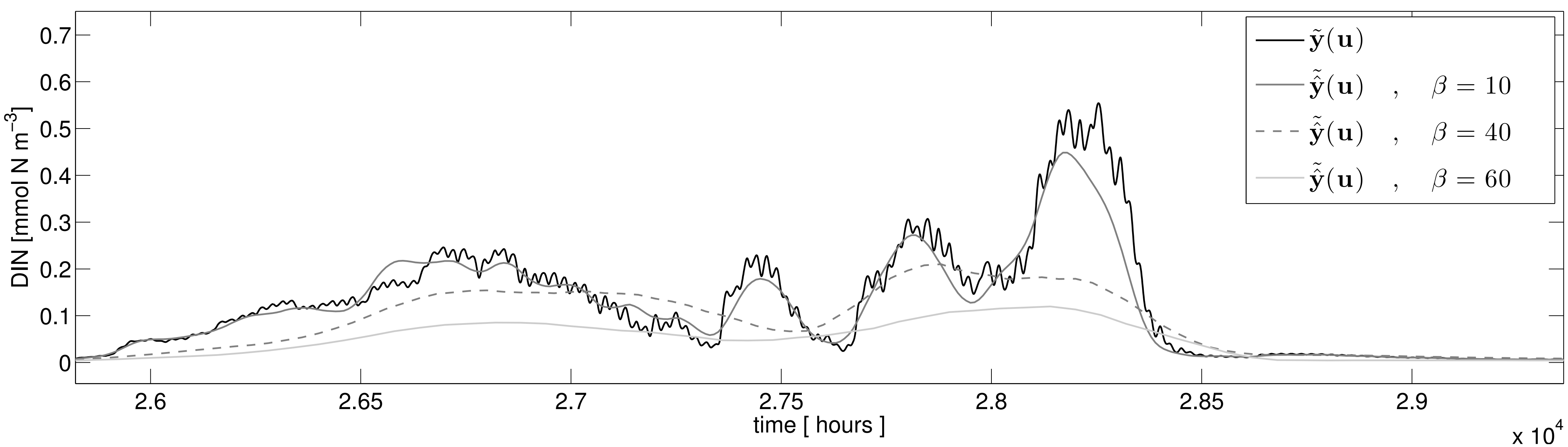


Figure 3: *Same as in Figure 2 but now using smoothing for both the coarse and the fine model. Smoothing helps removing the numerical noise in the model outputs so that the optimization process is able to identify and track relevant changes of the traces of interest.*

Response Correction - Initial Approach

- *Elemental (multiplicative) response correction of (smoothed) coarse model* at iteration k

$$\left.\begin{aligned} s_{kji}(\mathbf{u}) &:= A_{kji}\,\tilde{\hat{y}}_{ji}(\mathbf{u}), \\ A_{kji} &:= \frac{\tilde{y}^{\beta}_{ji}(\mathbf{u}_k)}{\tilde{\hat{y}}_{ji}(\mathbf{u}_k)} \end{aligned}\right\} \quad \begin{aligned} & k = 1, 2, \ldots, \\ & j = 1, \ldots, \hat{M}, \quad i = 1, \ldots, I, \end{aligned}$$

($\hat{M}$ = # of discrete temporal points of coarse model, I = # of discrete spatial points)

Generalization Capability

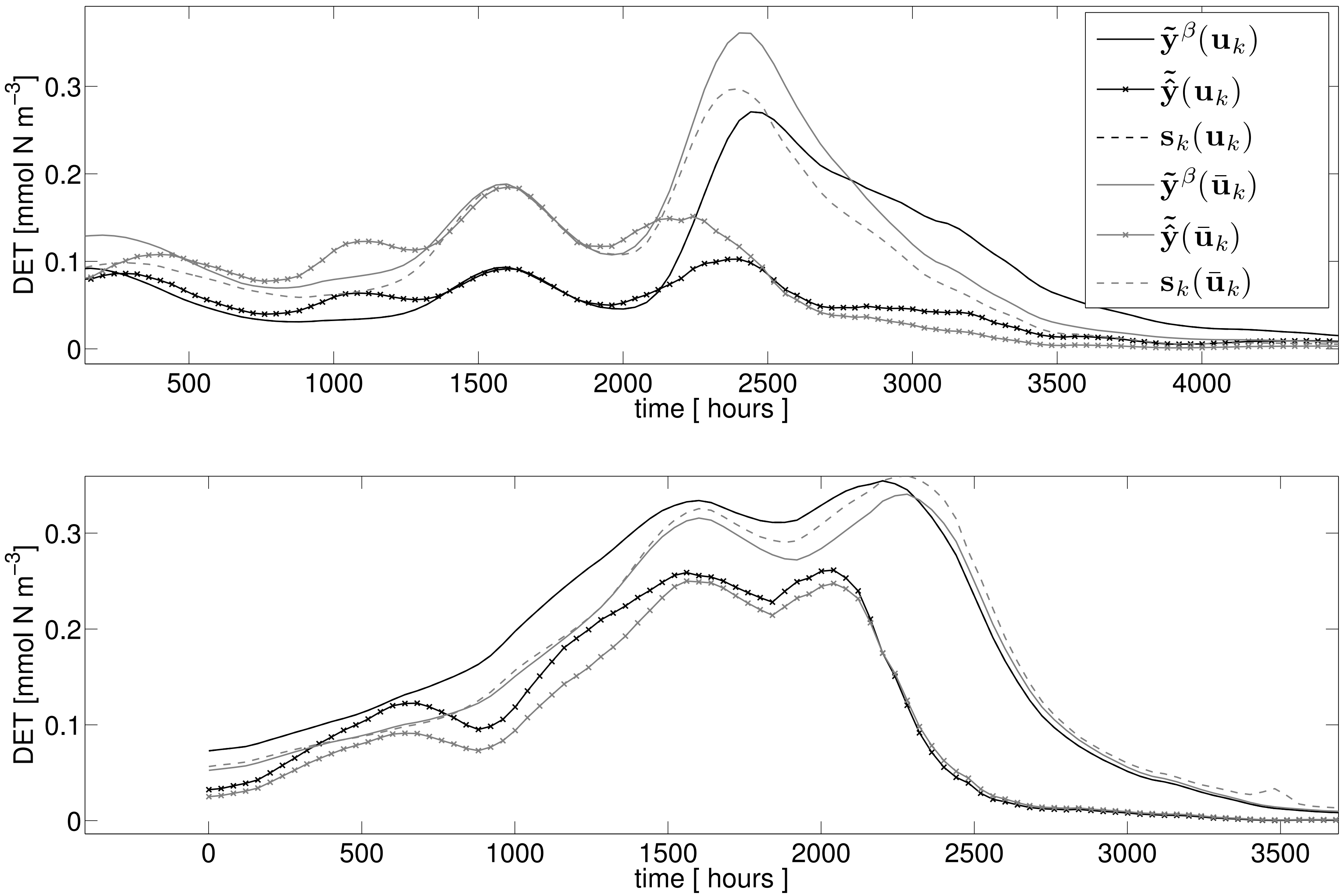
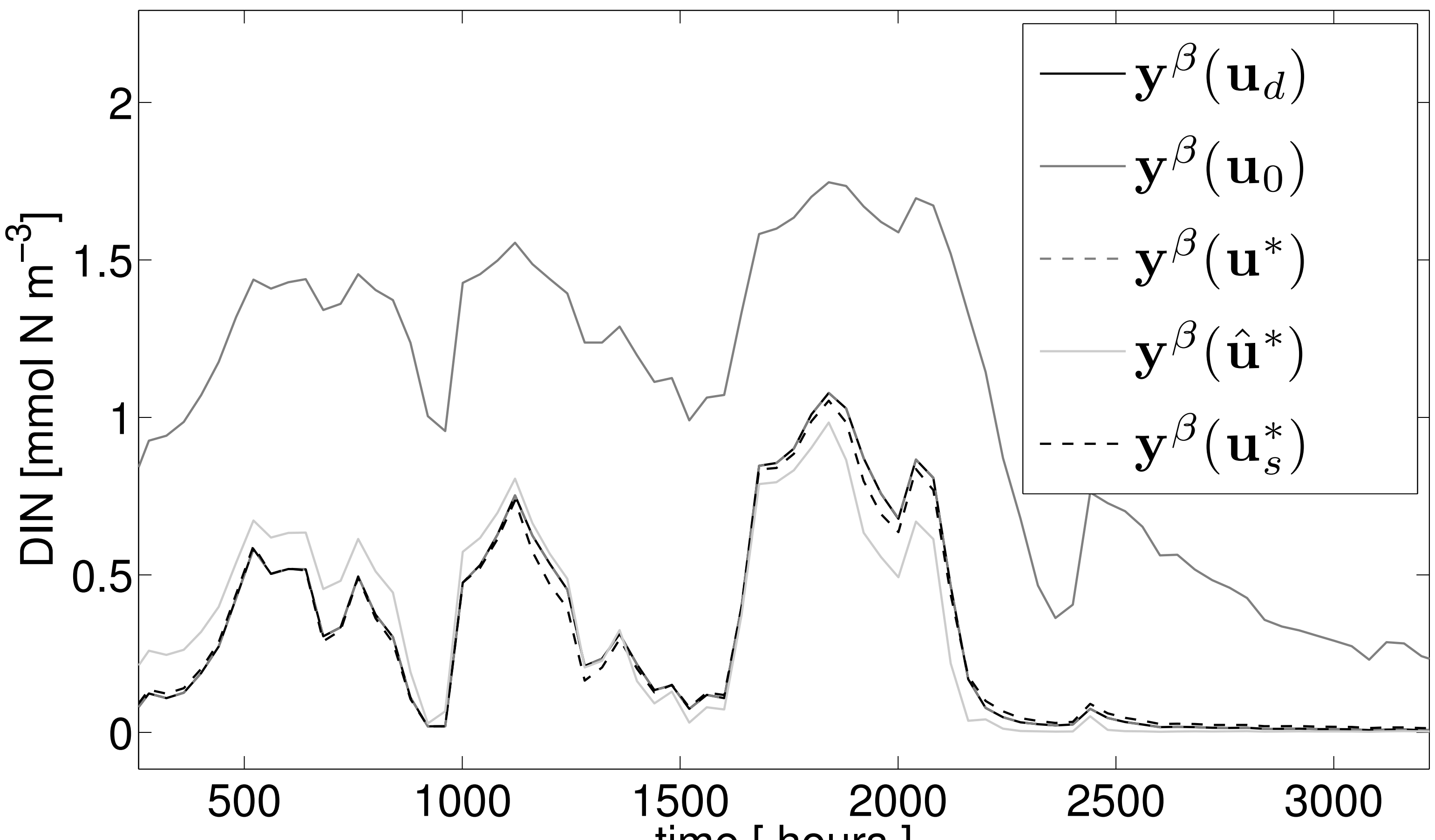


Figure 5: *Surrogate's, fine (down-sampled) and coarse model output $\mathbf{y}^{\beta}$, $\hat{\mathbf{y}}$, $\mathbf{s}_k$ for the state detritus at depth $z \approx -2.68$ m and at two iterates $\mathbf{u}_k$ and in a neighbourhood $\bar{\mathbf{u}}_k$. The surrogate obviously provides a reasonable approximation of the fine model at the point and in the neighborhood. Shown are the smoothed model outputs and for illustration only for some representative tracers and a part of the whole time interval only.*

Numerical Results



iterate	$J(\mathbf{y}^{\beta}(\mathbf{u}))$	C_i
$\mathbf{u}_0$	6.609e+04	
$\mathbf{u}^*$	1.267e-02	983
$\hat{\mathbf{u}}^*$	2.96e+03	11.275
$\mathbf{u}_s^*$	48.527	59.575
$\mathbf{u}_d$	$\sim$ 84% reduction	

Figure 6:

(left) Fine model output $\mathbf{y}^{\beta}$ (down-sampled) for dissolved inorganic nitrogen at depth $z \approx 2.68$ m. Shown are, in the legend from top to bottom: (i) Synthetic target data, i.e., fine model output $\mathbf{y}^{\beta}$ at randomly chosen parameters $\mathbf{u}_d$, (ii) fine model output at the initial value $\mathbf{u}_0$, (iii) at the result of the direct fine model optimization $\mathbf{u}^*$, (iv) at the coarse model optimum $\hat{\mathbf{u}}^*$ and (v) at the result $\mathbf{u}_s^*$ of a SBO run based on a multiplicative response correction.

(right) Cost function values J, computational costs C_i (in terms of number of equivalent fine model evaluations) at the initial parameter value $\mathbf{u}_0$, the fine model optimum $\mathbf{u}^*$, the coarse model optimum $\hat{\mathbf{u}}^*$ and at the solution $\mathbf{u}_s^*$ of a SBO run. Cost savings, when using SBO, are about 84% when compared to the direct fine model optimization.

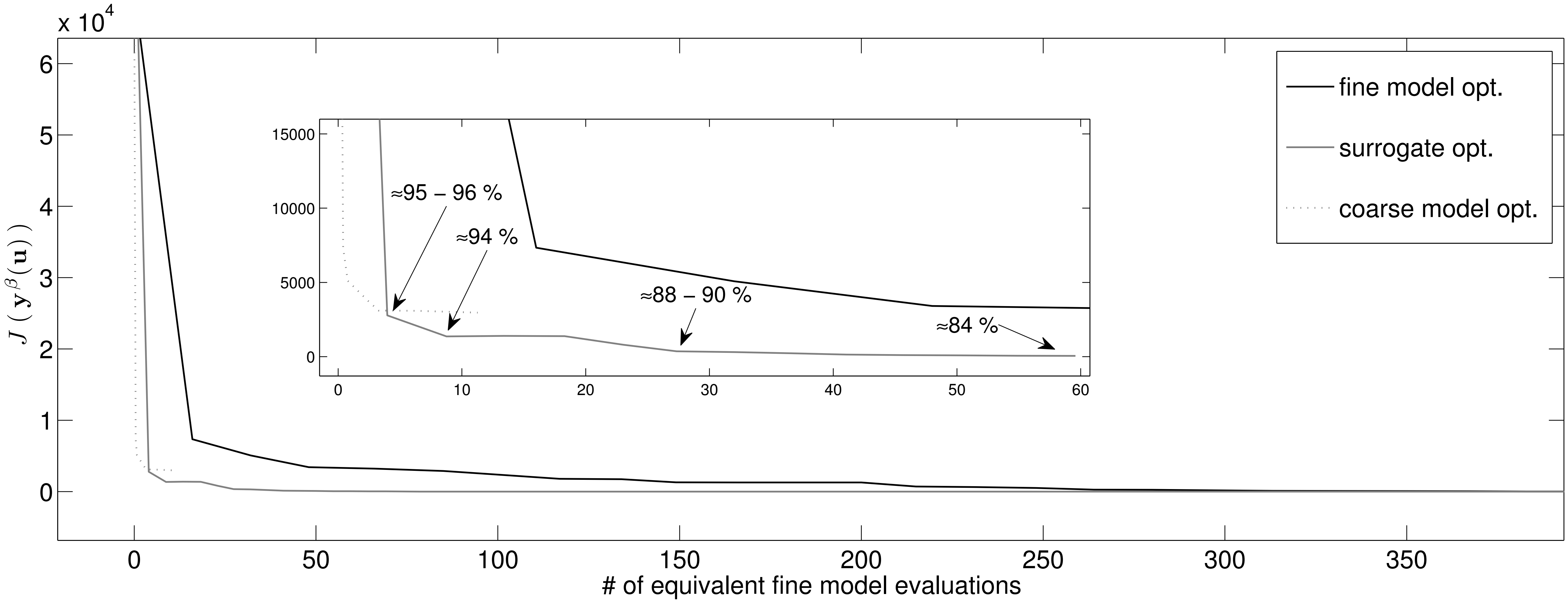


Figure 7: *The values of the cost function J versus the equivalent number of fine model evaluations for the fine, coarse and the surrogate-based optimization run. Several points corresponding to various values of the relative reduction in the total optimization cost (surrogate-based optimization versus straightforward fine model optimization) are also indicated. Results of fine model and surrogate optimization given in Figure 6 (left) correspond to the point marked as ~84%*

Summary

- We presented an efficient optimization methology for the optimization of climate model parameters
- We use a one-dimensional marine ecosystem model as a representative of this class of models
- Our approach is based on a coarser discretized low-fidelity model which is corrected by a multiplicative response correction
- The optimization process requires only one high-fidelity model evaluation per iteration
- It turned out that even without sensitivity information this approach is able to yield a very reasonable solution at the cost of a few high-fidelity evaluations only
- The robustness of the algorihtm can be further improved by using fine model sensitivity information

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.
- L. Leifsson and S. Koziel, "Multi-fidelity design optimization of transonic airfoils using physics-based surrogate modeling and shape-preserving response prediction," Journal of Computational Science, vol. 1, no. 2, pp. 98 – 106, 2010
- J. Sarmiento and N. Gruber, Ocean Biogeochemical Dynamics. Princeton University Press, 2006
- 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.

Smoothing

- Due to *numerical noise* (cf. Figure 2), it is *reasonable to smoothen* the coarse model output
- It was observed by visual inspection of the model outputs that this procedure allows us to *remove the numerical noise and identify the main characteristics* of the traces of interest
- For the smoothing we use a *walking average with span ±n* given as:

$$\tilde{\hat{y}}_{ji} \quad := \quad \frac{1}{2n+1} \sum_{m=j-n}^{j+n} \left[\frac{1}{2n+1} \sum_{p=m-n}^{m+n} \hat{y}_{pi} \right] \qquad j \quad = \quad 1, \ldots, \hat{M}, \quad i = 1, \ldots, I$$

- It turns out, also by visual inspection, that a value of n = 3 and “double” smoothing are suitable for the considered problem

Numerical Instability

- It is important to keep in mind that *choosing β too large could lead to a numerically unstable scheme*
- The condition of stability is dependent on the ratio h / v and the nonlinear coupling term Q (h = spatial step-size, v = here, sinking velocity)

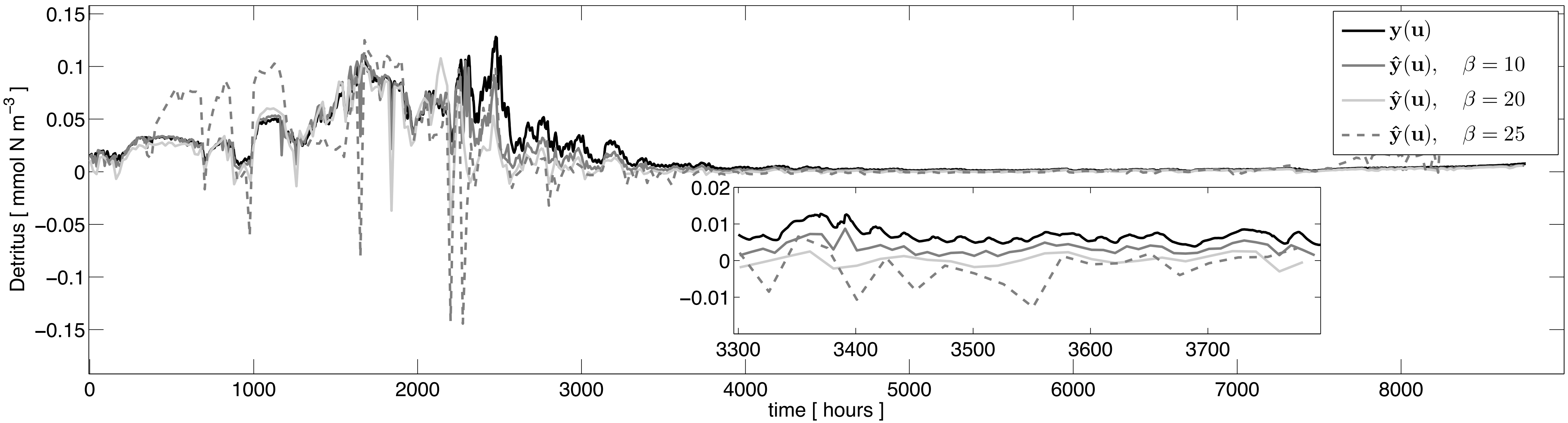


Figure 4: *The figure shows one year of the fine model output **y**(**u**) and of the coarse model output **ŷ** for the state detritus at depth z ≃ 25 m for different values of the coarsening factor β and at some fixed parameters **u**.*

Another Approach: Aggressive Space Mapping (ASM)

- *Aggressive Space Mapping* (firstly developed by John W. Bandler et., 1994) is based on:

$$\mathbf{s}_k(\mathbf{u}) := \hat{\mathbf{y}}\left(\mathbf{p}_k(\mathbf{u})\right), \quad \mathbf{p}_k(\mathbf{u}) = \mathbf{p}(\mathbf{u}_k) + \mathbf{p}'(\mathbf{u}_k)\left(\mathbf{u} - \mathbf{u}_k\right),$$

$$\hat{\mathbf{u}}_k = \mathbf{p}(\mathbf{u}_k) := \underset{\mathbf{u}\in U}{\operatorname{argmin}} \,||\,\hat{\mathbf{y}}(\mathbf{u}) - \mathbf{y}(\mathbf{u}_k)\,||_Y^2 \,.$$

- If either the fine model nearly matches the data in an optimum or if *both models* are similar near their respective optima we obtain, using (5), so-called perfect mapping

$$\mathbf{p}(\mathbf{u}^*) = \underset{\mathbf{u}\in U}{\operatorname{argmin}} \;||\,\hat{\mathbf{y}}(\mathbf{u}) - \mathbf{y}(\mathbf{u}^*)\,||_Y^2 \;\approx\; \underset{\mathbf{u}\in U}{\operatorname{argmin}} \;||\,\hat{\mathbf{y}}(\mathbf{u}) - \mathbf{y}_d\,||_Y^2 = \hat{\mathbf{u}}^*.$$

- This motivates to solve for

$$\mathbf{F}(\bar{\mathbf{u}}) := \mathbf{p}(\bar{\mathbf{u}}) - \hat{\mathbf{u}}^* = 0 \qquad \hat{\mathbf{u}}^* := \underset{\mathbf{u}\in U}{\operatorname{argmin}} \; J\left(\,\hat{\mathbf{y}}(\mathbf{u})\,\right)$$

- Under certain conditions ASM is equivalent to use surrogate given above in a SBO algorithm

$$\bar{\mathbf{u}}_s = \underset{\mathbf{u}\in U}{\operatorname{argmin}} \; J\left(\,\hat{\mathbf{y}}(\mathbf{p}(\mathbf{u}))\,\right)$$

Generalization Capability

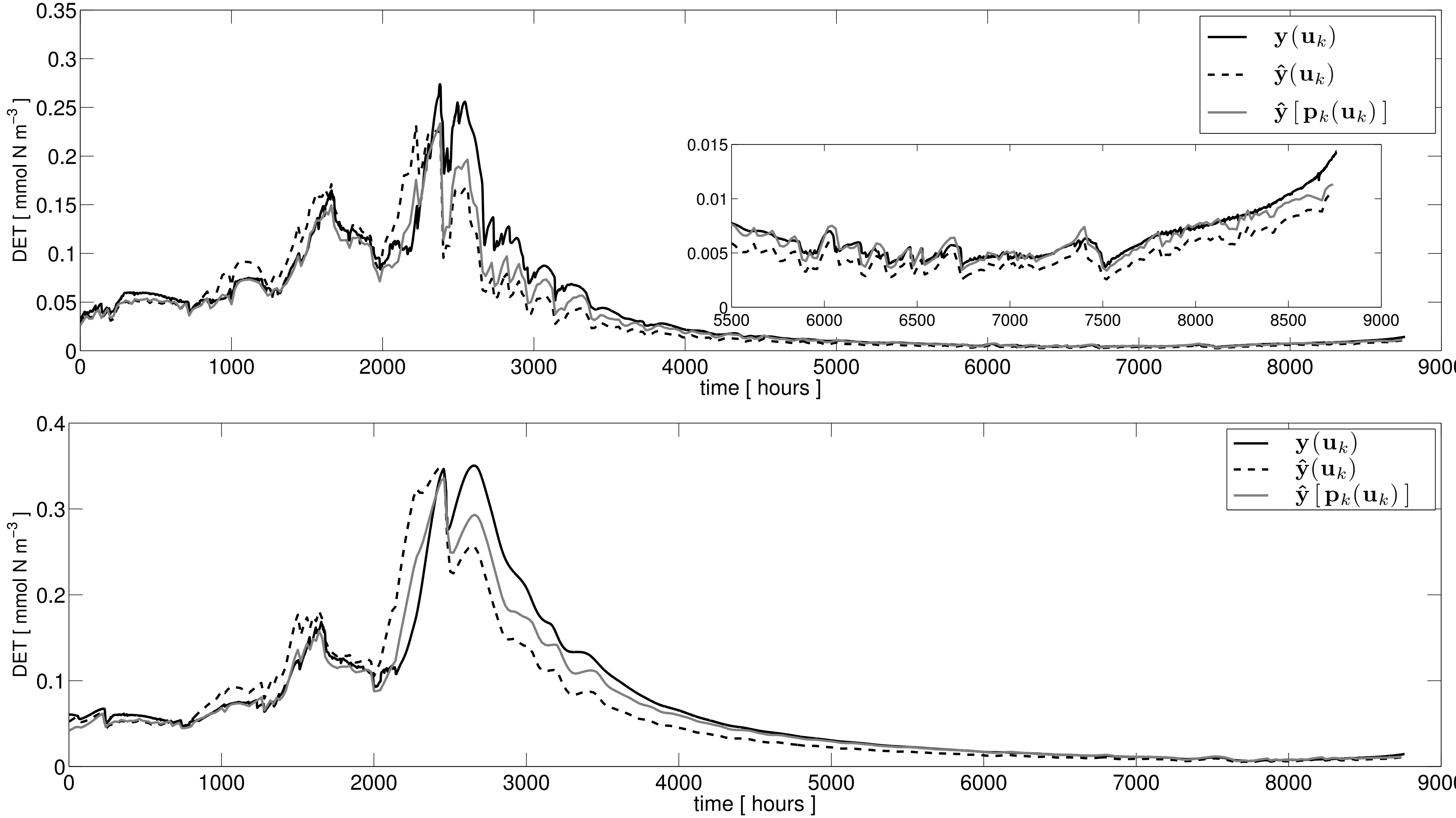
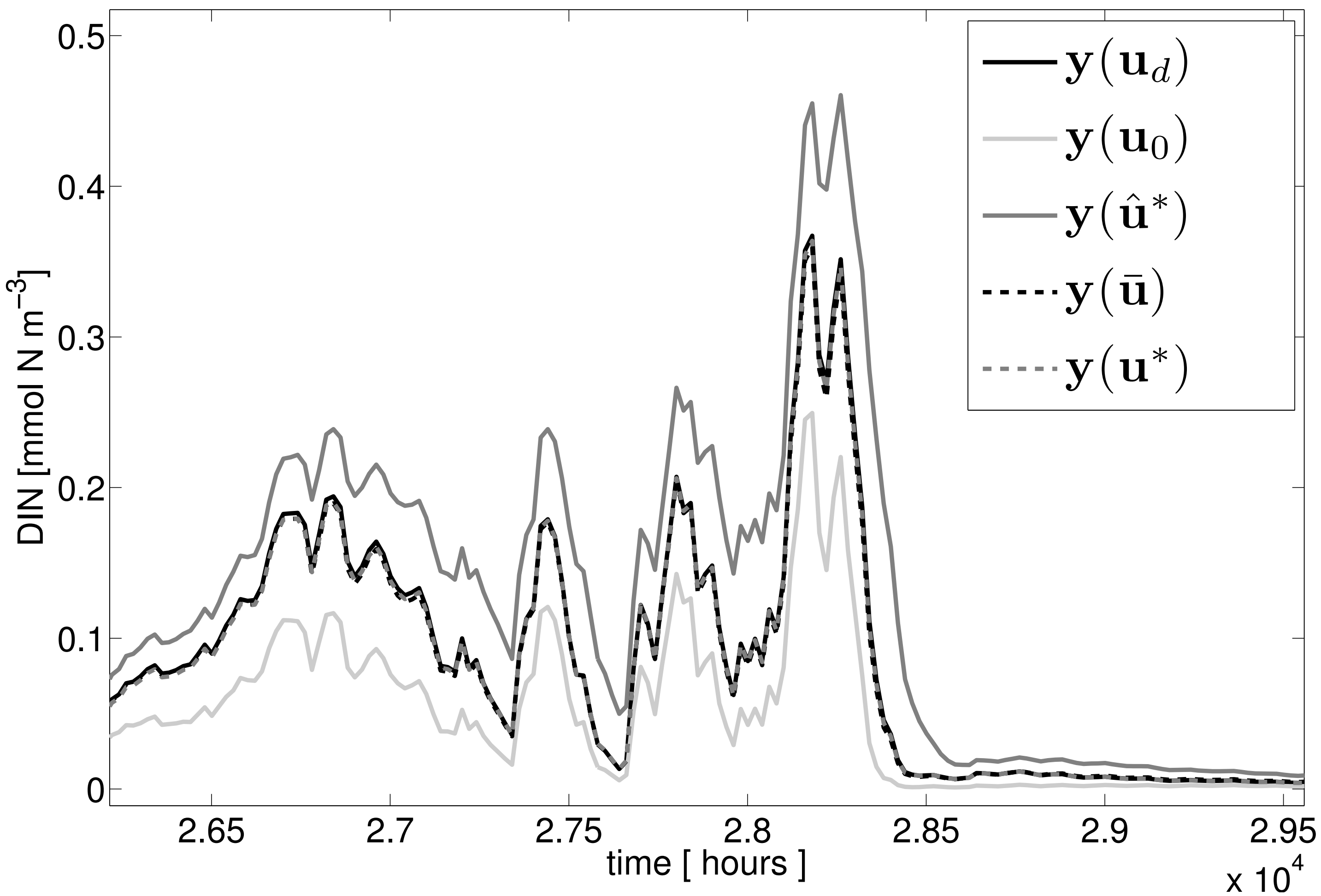


Figure 8: *Fine and coarse model output* ***y***, $\hat{\mathbf{y}}$ *as well as the aligned surrogate* $\mathbf{s}_k(\mathbf{u}_k) = \hat{\mathbf{y}}(\mathbf{p}_k(\mathbf{u}_k))$ *for the state detritus, at the same randomly chosen parameter vector* $\mathbf{u}_k$*, at depths z ≈ 25m (top) and z ≈ 60 m (bottom). The surrogate model provides a reasonable approximation of the fine model while lying closer than the coarse model itself.*



	J	C_i
$\mathbf{u}_0$	5.9e-03	
$\mathbf{u}^*$	1.6e-05	281
$\hat{\mathbf{u}}^*$	1.8e-03	19.95
$\bar{\mathbf{u}}$	5.0e-05	80.25
$\mathbf{u}_d$	57.54% reduction	

Figure 9:

(left) Fine model output ***y*** *for dissolved inorganic nitrogen at depth z ≈ 2.68 m. Shown are, in the legend from top to bottom: (i) Synthetic target data, i.e., fine model output* ***y*** *at randomly chosen parameters* $\mathbf{u}_d$*, (ii) fine model output at the initial value* $\mathbf{u}_0$*, (iii) at the coarse model optimum* $\hat{\mathbf{u}}^*$*, (iv) at the result of the ASM algorithm* $\bar{\mathbf{u}}$*, and (v) at the result of the direct fine model optimization* $\mathbf{u}^*$*.*

(right) Cost function values J, computational costs C_i *(in terms of number of equivalent fine model evaluations) at the initial parameter value* $\mathbf{u}_0$*, the fine model optimum* $\mathbf{u}^*$*, the coarse model optimum* $\hat{\mathbf{u}}^*$ *and the solution* $\bar{\mathbf{u}}$ *of the ASM algorithm. Cost savings, when using ASM, are about 57% when compared to the direct fine model optimization.*