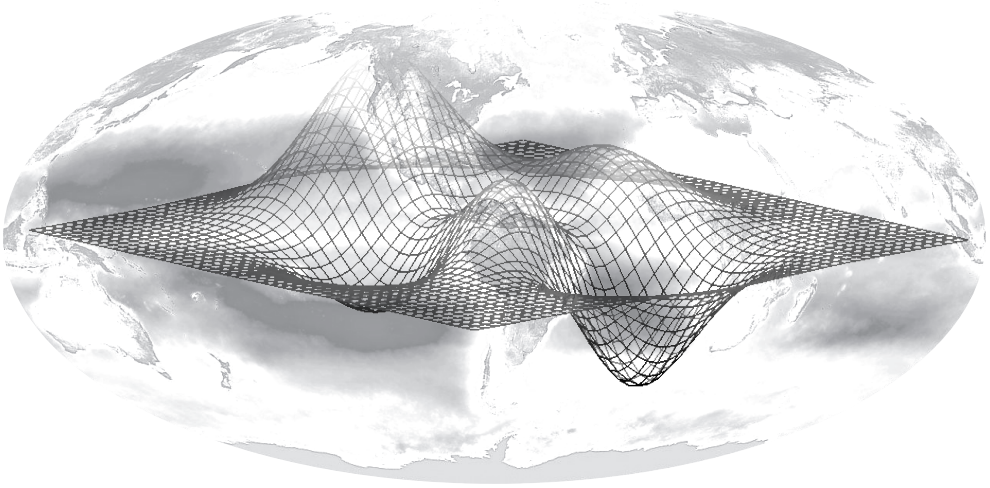
Surrogate-Based Optimization for Marine Ecosystem Models



Dissertation

zur Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Malte Prieß
Kiel, 2011

Referent:	Prof. Dr. Thomas Slawig
1. Koreferent:	Prof. Dr. Andreas Oschlies
2. Koreferent:	Prof. Slawomir Koziel, Ph.D.
Tag der mündlichen Prüfung: ..	31.01.2012
Zum Druck genehmigt:	31.01.2012

gez. Prof. Dr. Lutz Kipp, Dekan

To my family and friends

Abstract

Marine ecosystem models are of great importance for understanding the oceanic uptake of carbon dioxide and for projections of the marine ecosystem's responses to climate change. The applicability of a marine ecosystem model for prognostic simulations crucially depends on its ability to resemble the actually observed physical and biogeochemical processes. An assessment of the quality of a given model is typically based on its calibration against observed quantities. This calibration or optimization process is intrinsically linked to an adjustment of typically poorly known model parameters.

Straightforward calibration attempts by direct adjustment of the model parameters using conventional optimization algorithms are often tedious or even beyond the capabilities of modern computer power as they normally require a large number of simulations. This typically results in prohibitively high computational cost, particularly if already a single model evaluation involves time-consuming computer simulations. The optimization of coupled hydrodynamical marine ecosystem models simulating biogeochemical processes in the ocean is here a representative example. Computing times of hours up to several days already for a single model evaluation are not uncommon. A computationally efficient optimization of expensive simulation models can be realized using for example surrogate-based optimization. Therein, the optimization of the expensive, so-called high-fidelity (or fine) model is carried out by means of a surrogate – a fine model's fast but yet reasonably accurate representation.

This work comprises an investigation and application of surrogate-based optimization methodologies employing physics-based low-fidelity (or coarse) models. Seeking a computationally efficient calibration of marine ecosystem models serves as the fundamental aim. As a case study, two illustrative marine ecosystem models are considered. Here, coarse models obtained by a coarser temporal resolution and by a truncated model spin-up are investigated. The accuracy of these computationally cheaper coarse models is typically not sufficient to directly exploit them in the optimization loop in lieu of the fine model. I investigate suitable correction techniques to ensure that the corrected coarse model (the surrogate) provides a reliable prediction of the fine model optimum. Firstly, I focus on Aggressive Space Mapping as one of the original Space Mapping approaches. It will be shown that this optimization method allows to achieve a reasonable reduction in the optimization costs, provided that the considered coarse and fine model are sufficiently "similar". A multiplicative response correction approach, subsequently investigated, turned out to be very suitable for the considered marine ecosystem models. A reliable surrogate can be obtained. Exploiting the latter in a surrogate-based optimization algorithm, a computationally cheap but yet accurate solution is achieved. The optimization costs can be significantly reduced compared to what is achieved by the Aggressive Space Mapping algorithm.

The proposed methodologies, particularly the multiplicative response correction approach, serve as initial parts of a set of tools for a computationally efficient calibration of marine ecosystem models. The investigation of further enhancements of the presented algorithms as well as other promising approaches in the framework of surrogate-based optimization will be highly valuable.

Zusammenfassung

Marine Ökosystem-Modelle sind von großer Bedeutung, um die ozeanische Aufnahme von Kohlendioxid zu verstehen sowie Vorhersagen über die Reaktionen des marinen Ökosystems auf den Klimawandel treffen zu können. Die Anwendbarkeit eines marinen Ökosystem-Modells für prognostische Simulationen hängt entscheidend von seiner Fähigkeit ab, die tatsächlich beobachteten physikalischen und biogeochemischen Prozesse wiederzugeben. Um die Qualität von verschiedenen Modellen zu validieren, werden diese typischerweise an vorhandene Beobachtungsdaten angeglichen. Diese Validierung (oder Parameter-Identifikation) erfordert die Anpassungen von in der Regel wenig bekannten Modellparametern.

Die direkte Kalibrierung des Modells mit Hilfe konventioneller Optimierungsalgorithmen ist üblicherweise ein langwieriger Prozess, der gegebenenfalls sogar jenseits verfügbarer Rechenressourcen liegt. Ein Grund dafür ist die meist große Zahl erforderlicher Modellsimulationen. Dies führt insbesondere dann zu einem erheblichen Rechenaufwand, wenn bereits eine einzelne Modellauswertung teure Computersimulationen notwendig macht. Ein Beispiel hierfür ist die Kalibrierung gekoppelter mariner Ökosystem-Modelle. Rechenzeiten von Stunden bis hin zu mehreren Tagen für eine einzelne Modellauswertung sind nicht unüblich. Eine effiziente Optimierung von teuren Computermodellen lässt sich beispielsweise mit Hilfe von surrogat-basierten Optimierungsverfahren realisieren. Ein Surrogat – eine schnelle aber dennoch ausreichend genaue Approximation des sogenannten feinen Modells – ermöglicht hierbei dessen Optimierung.

Diese Arbeit umfasst die Untersuchung und Anwendung von Verfahren im Rahmen surrogat-basierter Optimierungsalgorithmen, bei denen die Surrogate auf sogenannten physikalischen groben Modellen beruhen. Übergreifendes Ziel ist eine effiziente und schnelle Kalibrierung von marinen Ökosystem-Modellen. Es werden zwei illustrative Modelle betrachtet. Die dazugehörigen groben Modelle werden beispielhaft durch grobe zeitliche Diskretisierung sowie durch einen verkürzten Modell-Spin-Up gewonnen. In der Regel sind solche groben Modelle nicht genau genug, um sie in der Optimierung direkt als Ersatz der feinen Modelle zu verwenden. Mit Hilfe geeigneter Techniken zur Korrektur der groben Modelle konstruiere ich daher ausreichend genaue Surrogate. Zuerst nutze ich hierfür Aggressive Space Mapping, einen der ursprünglichen Space Mapping-Algorithmen. Es wird gezeigt, dass dieses Optimierungsverfahren eine hinreichende Reduktion der Optimierungskosten erzielen kann, vorausgesetzt, das grobe und feine Modell stimmen ausreichend überein. Anschließend betrachte ich eine multiplikative Korrektur. Wie gezeigt wird, ist dieser Ansatz für die betrachteten Modelle gut geeignet. Zusätzlich ist die Optimierung der damit konstruierten Surrogate kostengünstig, erzielt aber dennoch eine ausreichend präzise Lösung. Die Optimierungskosten lassen sich hierbei deutlich gegenüber dem Aggressive Space Mapping-Algorithmus senken.

Die vorgestellten Verfahren, insbesondere die multiplikative Korrektur, stellen erste Teile einer Sammlung von Tools für eine effiziente Kalibrierung mariner Ökosystem-Modelle dar. Die Untersuchung weiterer Verbesserungen der betrachteten Methoden sowie anderer möglicher Ansätze im Rahmen surrogat-basierter Optimierung ist vielversprechend.

CONTENTS

INTRODUCTION

Understanding the earth's climate system and forecasting its future behavior is of great significance for assessing and designing adequate climate change policies and one of the most challenging tasks in climate science, applied mathematics and other scientific disciplines.

Why Predictions of the Earth's Climate?

The average temperature of the climate system, which includes the atmosphere and the oceans, increased by about 0.8°C within the last 100 years (IPCC, 2007b). The identified continuous global warming can no longer be primarily credited to natural factors. Warming of the climate system is unmistakable and its cause by increasing concentrations of greenhouse gases produced by human activities is hardly scientifically doubted (IPCC, 2007b). According to scientific evidence in reports of the *Intergovernmental Panel on Climate Change* (IPCC), which summarizes and discusses the scientific state of knowledge about global warming, a rising global temperature will lead to increasingly extreme weather conditions, such as droughts or heavy rainfalls. Other likely effects are sea level rises, desertification, the extinction of certain species as well as changes in agricultural yields (Braman et al., 2010; IPCC, 2007b).

Carbon dioxide (CO_2) is – among other greenhouse gases such as methane and nitrous oxide – one of the main contributors to global warming. Its concentration in the atmosphere has dramatically increased since the beginning of the Industrial Revolution predominantly due to the burning of fossil fuels (IPCC, 2007b). In 2010, global CO_2 emissions exceeded the most pessimistic forecasts of the IPCC (Boden and Blasing, 2011). Based on the findings of the IPCC, a warming of 1.1°C to 6.4°C is expected by the end

of the century (IPCC, 2007a). This depends on future greenhouse gas emissions and the actual response of the climate system. However, assuming an increase in the global mean surface temperature of 4.0°C would very likely exceed the adaptation capabilities of humans and the climate system (Warren, 2011). According to current scientific conclusions, mean temperature increases of up to 2.0°C could avoid at least extremely dangerous climate change and the worst risks involved (Hansjürgens, 2009; O'Neill and Oppenheimer, 2002).

Regarding the international law sphere, the *United Nations Framework Convention on Climate Change* (UNFCCC) and its *Kyoto Protocol* are the main instruments relating to climate change (UN, 1992, 1998). The UNFCCC – as an overall framework convention – aims to achieve a stabilization of greenhouse gas concentrations in the atmosphere at a level that would prevent dangerous anthropogenic interference in the climate system (Art. 2 UNFCCC). Further substantiating this aim, the Kyoto Protocol is currently the only international legally binding regime which addresses greenhouse gas emissions. 192 nations ratified this agreement to coordinate international efforts for preventing dangerous climate change (UN, 2011). The Kyoto Protocol, however, has two major limitations. Not included are, amongst others, the United States of America, India and China, the biggest emitters of CO_2 worldwide (Drieschner, 2011). Additionally, the Kyoto Protocol is going to run out by the end of 2012 (Art. 3 para. 1 Kyoto Protocol). So far, there doesn't exist any follow-up agreement to limit emissions in a legally binding way. At least, at the recently completed conference on climate change in Durban, member states agreed that the next UN conference in Qatar in 2012 enacts an extension of the Kyoto-Protocol beyond the year 2012. Furthermore, the 194 member states of the UNFCCC could agree upon a general roadmap for a new global climate treaty. By the year 2015, a legally binding agreement is supposed to be developed in order to limit global warming to 2.0°C above pre-industrial level, which then eventually includes the USA, China and India. This aim has already been formulated during the last conference in Cancún in 2010 (UN, 2010). However, the specifics of a new treaty remain unclear as of 2011, making projections of future emission trajectories rather impossible. Moreover, though the compromise at the conference in Durban initiates the development of a legally binding agreement, this treaty is only supposed to become effective by 2020. This illustrates the slow process of international negotiations concerning climate change.

Clearly, of primary importance for limiting global warming is a drastic reduction of anthropogenic emissions of greenhouse gases. However, according to the prevailing scientific understanding about global warming, it is very unlikely that even if a steep reduction in global greenhouse gas emissions can be achieved (e.g., fossil fuel related carbon emissions), one could avoid global warming in excess of the 2.0°C level (European Envi-

ronment Agency, 2005). Additional efforts originated in the past years – so-called *Climate Engineering* – with the primary aim to artificially remove greenhouse gases from the atmosphere (*Carbon Dioxide Removal*) or reflect incoming solar radiation back to space (*Solar Radiation Management*). Such approaches have the potential to reduce the effects of climate change but, at the same time, they are highly controversial (Rickels et al., 2011). Apart from human efforts for the mitigation of global warming, there exist various natural "sinks" in the climate system which contribute to a *natural removal* of atmospheric CO_2. These sinks are part of the *global carbon cycle* – therein the globally available carbon is exchanged among the biosphere (global sum of all ecosystems), pedosphere (soil), geosphere (inside of the earth), hydrosphere (combined mass of water), and the atmosphere of the earth (see here and the following, e.g., Fasham, 2003; Raven and Falkowski, 1999; Sarmiento and Gruber, 2002, 2006).

Those sinks include for example the reduction of dissolved inorganic carbon through the biogeochemical cycle among carbon and the ocean biota as part of the oceanic ecosystems (*marine carbon cycle*). According to the IPCC, over half of the anthropogenic concentrations of carbon dioxide is therewith removed from the atmosphere within a century (Meehl et al., 2007). Through photosynthesis within phytoplankton, more specifically by the light-independent or dark reactions, the dissolved inorganic carbon as well as nutrients such as dissolved inorganic phosphorus and nitrogen are converted to organic compounds. Through the food chain from zooplankton to higher trophic levels, a large part of the organic carbon is transformed back to dissolved inorganic carbon. This incorporates the processes of respiration and remineralization which take place in the upper 500m of the ocean under the influence of bacteria and oxygen. Parts of the dissolved inorganic carbon in the upper ocean is released back to the atmosphere again in the form of carbon dioxide. The organic carbon compounds incorporated in dead material (both zoo- and phytoplankton), which have neither been consumed nor remineralized, as well as organic excretions, either sink down to the deeper ocean or are transported to the latter by the ocean circulation (see also Denman, 2008; Körtzinger and Wallace, 2002; Lampitt et al., 2008). Apart from the portion which is released again by remineralization, a small fraction is deposited to the ocean ground, where it can be solidified by geological processes. Due to the latter and, moreover, due to the long dynamic adjustment time scales of the large ocean currents, it is assumed that the organic carbon is prospectively "stored" in the deeper ocean layers for longer periods. Time scales of several hundred to millions of years are currently under discussion (Denman et al., 1996; Lampitt et al., 2008). However, still rather unclear are the exact amounts of organic and inorganic material which persist in the deeper ocean and, furthermore, how much of the organic carbon is finally captured in sediments on the ocean ground.

Iron Fertilization, one of the discussed carbon dioxide removal techniques as mentioned above, is related to these natural processes involved in the marine carbon cycle. Iron serves as a minor nutrient for photosynthesis and thus its availability limits the amount of consumed CO_2. The principal idea of Iron Fertilization is to enhance the activity in photosynthesis in certain parts of the oceans and hence to intensify the oceanic uptake of CO_2. However, the specific changes in the dynamics of the oceanic ecosystem when fertilizing with iron are still not fully understood. Thus, at least from the scientific point of view, this approach has yet to be comprehensively analyzed (Boyd et al., 2007; Denman, 2008; Fasham, 2003; Güssow et al., 2010; Lampitt et al., 2008; Oschlies et al., 2010; Rickels et al., 2011).

From the presented results about global warming, it is unequivocal that an understanding of the relevant processes in the earth's climate system and, moreover, of its responses to human impact, is of great importance. Furthermore, projections of future dynamics within the earth's climate are clearly indispensable, last but not least for political decisions in response to global warming.

Climate Models and Computer Simulations

Analyzing the earth's climate system and predicting future dynamics is intrinsically linked to an understanding of the underlying physical principles describing the processes under consideration.

Here, *climate models* can be employed in the effort to obtain a reliable representation of the past and today's observed climate and thus to be appropriate and valid for prognostic simulations. More specifically, a climate model represents processes within the different components of the earth's climate system, i.e., atmosphere, sea and land ice, oceans and other water masses and the different terrestrial and marine ecosystems (see, e.g., McGuffie and Henderson-Sellers, 2005; Storch et al., 1999). The processes to be modeled range from fluid dynamics (in atmosphere and oceans), thermodynamics, radiative transfer to biochemical interactions, e.g., in marine or other type of ecosystems. Moreover, the spatial and temporal scales can be quite different. On the one hand, there can be non-negligible effects on very small scales (e.g., influence of clouds in the atmosphere and turbulence in the ocean), but, on the other hand, some processes might take very long time, such as the large-scale currents in the ocean (see here and the following, McGuffie and Henderson-Sellers, 2005; Storch et al., 1999).

Often, a climate model is formulated as a coupled system of non-linear, partial differential and algebraic equations. The underlying equations are typically too complex for analytical solutions, i.e., in other words, to be solved by "hand". Instead, one seeks

approximate solutions of the model equations by a suitable computer implementation, typically referred to as *model simulation* or *numerical model* (see, e.g., Golub and Ortega, 1995; Storch et al., 1999). In principle, to obtain the numerical model, the underlying model equations – defined *continuously* in time and space – have to be transferred into their *discrete* counterparts (*discretization*), since computers operate in discrete steps. As one simple example, for time dependent functions involved in the model, one introduces a temporal "mesh" on which this function is evaluated. The same is done in all incorporated spatial dimensions. Moreover, differential operators – such as those describing diffusive and advective processes in the ocean – have to be discretized. These operators include first- and second-order derivatives which have to be approximated appropriately. For the latter, typically, standard numerical discretization schemes are exploited. A simple example is the approximation of the first-order derivative of a real scalar function by the well-known finite-difference approximation (see, e.g., Bärwolff, 2007; Griffies, 2004; Golub and Ortega, 1995; Hackbusch, 2010; Marchuk, 1982).

Simulations of the earth's climate system through computer programs can be computationally very expensive and typically require high-performance computing, especially if the system under consideration comprises processes in more than one spatial dimension and with a high temporal and spatial resolution. Due to many interactions between the different components, since many important processes are non-linear, and the human impact, the complexity of the climate system is huge. This makes the development of a sufficiently accurate but yet computationally affordable climate simulation very difficult. Clearly, the primary aim is to include as many processes as possible. However, due to limitations in knowledge and understanding of the climate system and, moreover, due to existing constraints on available computer resources, simplifications are inevitable.

Principally, two different approaches can be made to yield simplifications of the included processes and relevant interactions (see for both, e.g., McGuffie and Henderson-Sellers, 2005; Storch et al., 1999). The first one is related to the representation of the processes themselves. It might be possible to reduce the complexity of their *parametrizations*, i.e., their mathematical descriptions within the model. This can lead to a reduction of the accuracy of the simulated processes and the related computational effort. These parametrizations are typically based on either physical principles or are given by semi-empirical laws. A widely used example in many models is to prescribe observed averages through so-called *climatological specifications*. The second common approach involves the development and use of models with a lower temporal and/or spatial resolution. Typically, it is assumed that a finer resolution yields a higher accuracy of the simulation. Reducing the resolution of a model is one of rather straightforward approaches to obtain a less accurate but computationally fast simulation of the desired processes. In principle, simplifications are

particularly reasonable for processes that might have little effect in the considered system or can be treated sufficiently accurate by simpler formulations. However, the growth in available computer resources and the accelerated progress in new computer technologies allows to exploit computationally even heavier simulations and to continuously increase the complexity of the underlying models.

Nowadays, climate simulations are indispensable and rather standard tools exploited in various application areas, from high-resolution weather forecasting, predictive simulations of global warming trends, coastal research and, last but not least, the simulation of the dynamical evolutions of marine ecosystems. The corresponding marine ecosystem models are of primary importance for understanding and simulating the oceanic uptake of carbon dioxide as well as for projections of the oceanic ecosystem's dynamics and their responses to climate change. As an example, in Sarmiento et al. (2004), the responses of marine ecosystems to climate warming have been investigated and projections were made until the year 2050 by different climate model simulations. Moreover, assessing the risks and the potential of human interventions to global warming such as Iron Fertilization, as motivated above, provides yet another example. This assessment can be significantly supported by appropriate numerical simulations of the marine ecosystem's response to such external impacts, as has been repeatedly investigated in the past (see, e.g., Denman, 2008; Oschlies, 2004; Oschlies et al., 2010).

Marine Ecosystem Models

Generally, marine ecosystem models are formulated as time-dependent systems of equations modeling the transport, the interactions and biogeochemistry among ocean biota. The modeled processes comprise the marine biogeochemical cycles among carbon and the major nutrients – therein the marine carbon cycle (see, e.g., Fasham, 2003; Fennel and Neumann, 2004; Sarmiento and Gruber, 2006).

The transport of the ocean biota (or "tracers") incorporates so-called advective and diffusive transport processes, i.e., in short, the movement and the mixing of tracer concentrations. The terms describing these transport processes include differential operations, such as the derivative of the tracer concentrations with respect to space. The diffusion – as one example – is driven by the concentration gradient (i.e., the vector whose components are the derivatives with respect to the distinct spatial directions). The biochemical interaction terms involve predator-prey relations, where the nutrients, as one example, serve as food for phytoplankton. Some incorporated biogeochemical processes are described by "local" terms, since they are calculated from the tracer concentrations at one spatial point only, such as the mortality of zooplankton. Others are dependent on quantities

at distinct spatial locations, such as the light limited growth of phytoplankton which is typically dependent on the vertically integrated available light intensity. Moreover, many terms describing the biochemical interactions are non-linear. Since the sunlight serves as a major driving force in the system, many models obey annual, seasonal and/or monthly temporal variations. Furthermore, the spatial and temporal scales can be quite different. As noted before, this is due to the incorporation of effects on very small scales on the one hand and, on the other, processes on long time scales such as the transport of the tracer concentrations within the deeper ocean layers by the large-scale ocean currents.

Apart from the well-understood physical principles and governing equations for the ocean circulation, the complex and partly unknown biogeochemical cycles are challenging when formulating a comprehensive and reliable marine ecosystem model. There exists a wide range of models with varying complexity determined for example by the number of incorporated nutrients and plankton types (see, again, Fennel and Neumann, 2004; Sarmiento and Gruber, 2006). However, "(...) there is no general consensus on how complex a biogeochemical model should be in order to faithfully represent the interplay between ocean biota, ocean physics and the marine biogeochemical cycles of carbon and the major nutrients" (see here and the following, Kriest et al., 2010). Clearly, the development of suitably complex ecosystem models, reliable parametrizations therein as well as an assessment of their quality are an indispensable part of current research (see also, Evans and Garçon, 1997).

Simulation

A marine ecosystem model is typically coupled with a hydrodynamic model describing the ocean circulation including the temperature and salinity distribution acting as the main drivers (see, e.g., Gill, 1982; Griffies, 2004; McGuffie and Henderson-Sellers, 2005; Storch et al., 1999). On the one hand, the tracers are advected by the ocean circulation and their diffusion is dominated by the turbulent mixing of water masses in the ocean. Vice versa, a tracer concentration may on the other hand also effect the ocean circulation.

For a fully coupled simulation, i.e., where a coupling between the hydrodynamic and the marine ecosystem model is considered in both directions, the two systems must be simulated simultaneously. This can be computationally expensive; a single model evaluation in three space dimensions typically requires high-performance computers. Therefore, and since the influence of the ocean biota on the circulation (including temperature and salinity distribution) is assumed to be negligible (*passive tracers*) and thus is often omitted, the coupling is mostly regarded as a one-way coupling, so-called *off-line* mode/computation. See for example Mahlman and Moxim (1978) where such an off-line computation has been thoroughly described and investigated for an atmospheric model. In an off-line mode,

quantities such as velocity, temperature fields and the turbulent mixing coefficients are computed beforehand by an ocean circulation model and used as so-called *forcing data* for the biogeochemical simulations (i.e., the simulations of the transport and biogeochemical interactions of the ocean biota). This approach becomes clearly attractive since the amount of computation can be significantly reduced. In principle, there are two ways to make use of the precomputed ocean circulation data in an off-line computation. The data can be stored directly and afterwards used for assembling the system matrices for the differential operators describing the diffusive and advective transport in the marine ecosystem model. Another approach is the *transport matrix approach*, which allows to obtain a considerably more efficient off-line tracer simulation for passive tracers. Therein, the ocean model that precomputes the data is used to generate the necessary, so-called transport matrices which can be subsequently employed to compute the distribution of the biogeochemical tracers. This has been originated and investigated as so-called *Transport Matrix Method* by Khatiwala et al. (2005).

Subsequently, the temporal evolution of the tracer concentrations is calculated. Here – regardless which of the approaches mentioned above is used – the underlying time-dependent system of partial differential equations is typically solved exploiting some *time integration scheme*. A popular and rather simple technique is the Euler time-stepping scheme (see, e.g., Bärwolff, 2007; Golub and Ortega, 1995; Stoer and Bulirsch, 2002). Typically, a time integration is either performed with constantly distributed or a-priori estimated initial tracer distributions. Thus, and due to the possibly long adjustment time scales of the incorporated processes, marine ecosystem models usually have to be run at first (so-called *spin-up*) in order to obtain a physically meaningful simulation.

Generally, for given time-dependent but non-periodic forcing data, one may consider a simple *transient run* for a required time interval. However, *a steady annual cycle*, i.e., a *periodic solution*, might be desired if annual periodic forcings are considered. To "solve" for a periodic solution usually involves many repetitive model simulations in the effort to achieve periodicity of the simulated processes. This procedure corresponds to a classical *fixed point iteration* (Bärwolff, 2007; Stoer and Bulirsch, 2002). Each single model simulation within this iteration is performed, again, using some time integration scheme. Yet another approach to solve for a periodic solution (such as a steady annual cycle) is the so-called *Newton Krylov* method (Khatiwala, 2008). It is based on the classical Newton's method and evolves from the fact that the problem to find such a periodic solution can be translated into solving for a root of a corresponding system of equations.

Model Calibration

The applicability of a marine ecosystem model for prognostic simulations crucially depends on its ability to resemble the actually observed physical and biogeochemical processes. Generally, before a simulation is reasonable, a marine ecosystem model has to be *calibrated*. This involves the *identification* of relevant parameters in order to estimate the model's capability in representing the observed data and thus to be appropriate to simulate the processes under consideration. The parameters to be identified are typically poorly known quantities in the biogeochemical coupling terms. One reason is that the incorporated model tracers are typically composed of various species which are differently affected by changes in their environment (Fennel et al., 2001).

Another typical bottleneck is the right trade-off between a high model complexity (usually, corresponding to a large number of parameters) and a rather simplified model formulation with a few parameters only. A complex model might be not reliably constrained by the data, whereas, a simple model might not correctly represent the data (see, e.g., Kriest et al., 2010; Ward et al., 2010). This is also known as the problem of *over-* and *underdetermination* (Tarantola, 2005). Accordingly, an assessment of the quality of the different models highly depends on their validation against observed quantities. A "manual" model validation process is intrinsically tied to adjustments of the model parameters rather individually and most likely requires repetitive model simulations. However, this rather non-systematically fitting process strongly relies on the experience on the utilized model and involved processes and is a tedious process which does not guarantee optimal results, even for an ecosystem model with a few parameters only.

Therefore, straightforward automation attempts employing systematic and more sophisticated so-called *optimization* procedures – which might even guarantee *convergence* to an optimal solution – are highly desirable. The development and generalization of optimization theory and techniques comprises a large field of applied mathematics and computational science.

Optimization

Mathematical optimization (or *mathematical programming*) generally denotes the process of systematically finding optimal *parameters* (or *designs*) from within an *allowed set* such that a so-called *objective function* (or *cost function*) is maximized or minimized. The objective function typically measures a misfit between some simulated quantities and some desired specifications. These quantities are either directly given by the model response itself or are calculated from the latter. Typically, *constraints* on the parameters and/or on the simulated quantities are furthermore considered. The optimization task is typically

formulated as a minimization problem of the form

$$\min_{\mathbf{u}} J(\mathbf{y}(\mathbf{u})) \quad \text{s.t. constraints,} \tag{1}$$

where J denotes the objective function, $\mathbf{u}$ the parameters subject to the optimization and $\mathbf{y}$ denotes the model response (with *state variable* $\mathbf{y}$) which is dependent on $\mathbf{u}$. The model state $\mathbf{y}$ is normally evaluated through a computationally possibly expensive numerical simulation and will be referred to as the *high-fidelity* (or *fine*) model in the following. The optimization (1) typically involves repetitive *iterations* until a user-defined *termination condition* is satisfied. This condition can be based on certain convergence criteria (e.g., a threshold for the step size), an assumed level of the objective function value or a specific number of iterations (particularly if the computational budget of the optimization process is limited).

Examples for optimization problems can be found in various application fields, for example within the area of economics. Here, the objective function to be maximized mostly represents the profit or turnover of a company; the parameters to be optimized are raw materials, labor, machine usage, prices, etc. (see, e.g., Schmedders, 2008). Another example evolved from the growing development and usage of computational fluid dynamics solvers in the design and analysis of various engineering systems, such as aircraft, turbo machinery, ships, and automotives (see, e.g., Braembussche, 2008; Dumas, 2008; Hicks and Henne, 1978; Leifsson and Koziel, 2010; Percival et al., 2001). Probably, their most common application in engineering design is for adjustment of designable parameters of the fluid system under consideration to make it satisfy prescribed performance requirements. This can be for example the direct adjustment of geometry shape parameters of airfoils and wings to optimize their performance in terms of maximum lift and minimum drag (Jameson, 1988).

Another approach – also known as *inverse approach* or *parameter identification* – is to define a specific target to be attained by the model simulation by identifying the relevant model parameters (Tarantola, 2005). Typically, this is the fundamental aim of calibration and validation of marine ecosystem models. Here, the objective function to be minimized is often formulated as the squared difference of the simulated and observed quantities (so-called *least-squares* type optimization problem). Typically, some *weighting factors* for the distinct quantities are furthermore taken into account. To ensure a physically reasonable solution, the parameters which are subject to the optimization are usually constrained by simple upper and lower parameter bounds. Moreover, additional constraints on the model state variables might be given. The "optimal" parameters are those which yield a minimal misfit between the simulated model data and the measurements. Parameter identification so far has been widely implemented in meteorological and ocean circulation

studies (see, e.g., Ghil and Malanotte-Rizzoli, 1991; Malanotte-Rizzoli, 1996) and further extended to coupled marine ecosystem models (see details further below).

Some practical problems, mostly in business and economics and for some engineering applications, fall into the area of *linear* optimization, i.e., where both the objective function and the constraints are linear. Linear optimization is a specific type of *convex* minimization which means that each "local" minimum of the objective function is automatically equal to the "global" optimal solution of the considered problem. The exact solution and analysis of linear optimization problems is rather straightforward – exploiting well known linear programming solutions – and has been thoroughly investigated in the past. Popular methods include the *Simplex algorithm*, developed by Georg B. Dantzig in 1947, as well as various, typically more efficient, *Interior point methods* which have been invented by John von Neumann and further advanced by the work of Narendra Karmarkar in the year 1984 (Luenberger, 2008; Nocedal and Wright, 2000). These two main algorithms are still subject to current research. Nowadays, many linear problems with hundreds to thousands of variables and inequalities can be optimally solved within comparably short times.

However, more complicated is the case of *non-linear* optimization where the objective function and/or the constraints are non-linear. Typically, for non-linear optimization problems, many different local minima exist, whereas the global solution might be hard to detect. As an example, parameter identification problems employing marine ecosystem model are typically non-linear (due to the biogeochemical interactions) and non-convex.

Generally, one can distinguish between so-called *descent* (or *downhill*) optimization algorithms which seek a local solution of an optimization problem, also known as *local methods* (see, e.g., Luenberger, 2008; Nocedal and Wright, 2000) and so-called *meta-heuristics* that are usually exploited in the effort of finding the global optimal solution of a given problem (*global* methods). Meta-heuristics methods are typically based on experienced data and some certain systematics and can usually search very large parameter spaces (see here and the following, Koziel and Yang, 2011; Glover and Kochenberger, 2003; Talbi, 2009; Weicker, 2007). Many meta-heuristics implement some form of stochastic optimization. However, seeking the global solution is much more challenging. Practically, there exist no method that can ensure convergence to the global minimum of a (non-convex) optimization problem. This is in contrast to local descent methods as will be further addressed below.

Descent methods are seeking a (local) optimal solution by exploiting gradient information – or suitable approximations – of the objective function. One example and very popular descent method for the local solution of non-linear optimization problems is the so-called *Method of Steepest Descent* which exploits a "direction" proportional to the negative gradient of the objective, thus uses first-order derivative information only (see

here and the following, Kelley, 2003; Luenberger, 2008; Nocedal and Wright, 2000). Moreover, in optimization, seeking a minimum or maximum of an objective function can be translated into finding a root of its first-order derivative. Accordingly, *Newton's method*, which is thus based on both first- and second-order derivatives, can be exploited. These two basic approaches are fast but leak of *robustness* which in this context means that the convergence to a local minimum of the objective requires that the starting point lies "sufficiently" close to the optimum. Convergence independent of the starting point is not guaranteed. This is also known as *global convergence*. Obviously, the term "global" in this context must not be mistaken with the capability of the algorithm to find the global optimal solution but characterizes the independence of the local convergences of the initially chosen starting point (see again, Luenberger, 2008; Nocedal and Wright, 2000).

The key idea of globally convergent optimization algorithms is to ensure a continually descent in the objective. The first popular strategy for ensuring global convergence is to exploit a *line search*, where the algorithm searches along a single dimension ("line"). Another increasingly popular methods are so-called *trust-region approaches* (see here and the following, e.g., Conn et al., 2000; Koziel and Yang, 2011; Luenberger, 2008; Nocedal and Wright, 2000). Here, in principle, a truncated Taylor expansion (often, in a quadratic form) is exploited and a local solution in the ball around the current parameter/design with a given *model-trust radius* is computed. Common to both approaches is that, in case no or not sufficient descent in the objective is found, the search range, i.e., the line or the ball, is subsequently decreased according to some given strategy until a "successful" iterate is attained. Both strategies are widely used in modern methods of non-linear optimization.

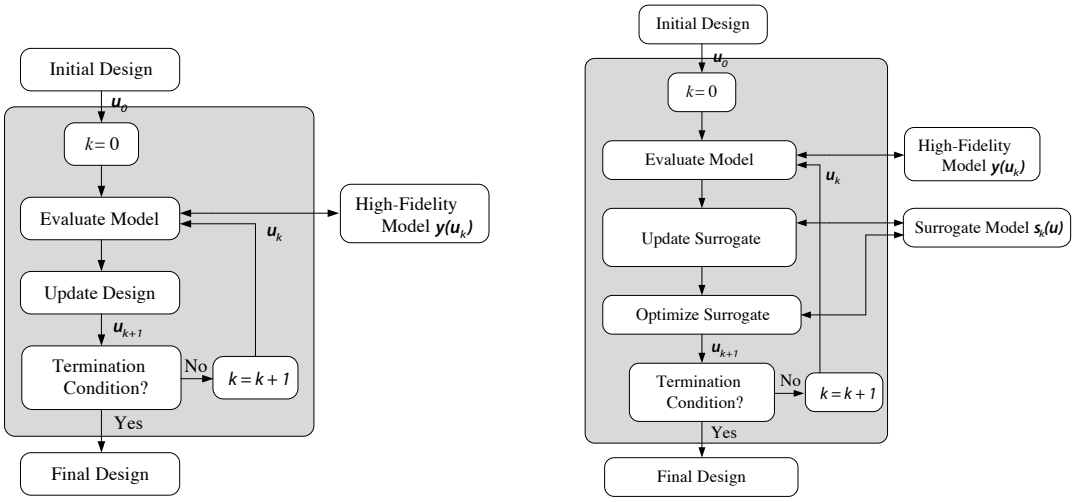
For many optimization problems, the objective's gradient (or approximations of it) are computationally expensive or even infeasible, particularly if many model parameters are involved. *Gradient-free* descent methods address this issue, where, by local evaluations of the objective function and suitable techniques, a descent direction is approximated. On the one hand, such methods have comparably slow convergence when compared to gradient-based ones, but, on the other hand, are rather easy to implement and computationally cheap. Moreover, these methods might be sufficient if only a rather inaccurate solution for a given optimization problem is desired. Popular and commonly used gradient-free methods are the *Nelder-Mead Method*, also known as *Downhill Simplex Method* (Kelley, 1999), as well as so-called *Pattern-Search* algorithms (see, e.g., Koziel, 2010a; Koziel and Yang, 2011).

Besides these convergent iterative descent methods, most conventional global algorithms do not use any gradient information nor any other techniques to obtain a direction of descent in the objective. They are based upon heuristics strategies which can be either deterministic or stochastic. A common element in global methods is the

repetitive search for local optimal solutions according to some given systematics. Among the global optimization algorithms, widely utilized stochastic meta-heuristic optimization approaches are so-called *Evolutionary Algorithms*, with *Genetic Algorithms* as the most popular type (see here and the following Koziel and Yang, 2011; Glover and Kochenberger, 2003; Talbi, 2009; Weicker, 2007). An Evolutionary Algorithm is based on principles from biological evolution, more specifically, reproduction, mutation, recombination and selection. As an example, evolutionary approaches have been widely used for the calibration of marine ecosystem models (see, e.g., Rückelt et al., 2010; Schartau, 2001; Schartau and Oschlies, 2003; Ward, 2009; Ward et al., 2010). Another meta-heuristic algorithm which is based upon probabilistic variables that are characterized by means and variances is the so-called *Simulated Annealing* which is also very popular in this application area (Armstrong et al., 1995; Hurtt and Armstrong, 1999; Matear, 1995). Typically, a global optimizer is computationally expensive and much slower than advanced local optimizers. Clearly, convergence to the global optimum of the given problem can't be guaranteed. Often, a more efficient global optimizer can be constructed by starting suitable and efficient local optimization algorithms from different (e.g., random) starting points. As an example, a mixed evolutionary and deterministic gradient-based approach has been proposed for a one-dimensional nitrogen-budget marine ecosystem model by Rückelt et al. (2010), employing a so-called *Quantum Evolutionary Algorithms* with a line search strategy and an efficient local optimization method.

Another approach in mathematical optimization is the so-called *Lagrange multiplier method* which provides yet another strategy to obtain an optimal solution of a constrained optimization problem (see, e.g., Nocedal and Wright, 2000). Therein, the so-called *adjoint* of the model – which runs backward in time – directly provides the gradient of the cost function with respect to the parameters subject to optimization. This *adjoint sensitivity* information can subsequently be exploited in some gradient-based algorithm. However, the most challenging aspect of this method is the development of the adjoint model formulation. The adjoint approach has become very popular in meteorology and oceanography and has been widely used in the context of parameter identification for marine ecosystem models (see with further references, Schartau et al., 2001; Ward, 2009; Ward et al., 2010; Lawson et al., 1996; Spitz et al., 1998).

Generally, solving non-linear optimization problems where computation of the objective function involves time consuming computer simulations may be quite challenging. One of the fundamental bottlenecks is that most of conventional optimization algorithms, whether gradient-based or some meta-heuristics, typically require a large number of objective function evaluations. This can result in prohibitively high computational costs, particularly if already a single model evaluation involves time-consuming computer simu-



(a) Direct optimization **(b)** Typical surrogate-based optimization

Figure 1: In a direct optimization (a), the complex high-fidelity or fine model under consideration is directly used in an optimization loop using conventional optimization approaches. In a surrogate-based approach (b), a computationally cheaper representation is exploited in lieu of the fine model in iteration k. In the figures, $\mathbf{u}_k$ denotes the parameter/design vector at iteration k.[1]

lations. Straightforward attempts employing the fine model under consideration directly in an optimization loop (*direct optimization*) is thus often a tedious or even infeasible process (cf. (1) and Figure 1a). As an example, the simulation of coupled high-resolution hydrodynamical marine ecosystem models still requires evaluation times of hours up to several days already for a single model evaluation. This is mainly due to the long dynamical adjustment timescales of incorporated processes such as the deep ocean currents. These typically require a sufficiently long integration time within the model simulation (see, e.g., Khatiwala et al., 2005). Consequently, sufficiently accurate simulations of three-dimensional coupled models – not to mention full optimization runs employing those models – are nowadays still computationally very expensive or even beyond the capabilities of modern computer power. Generally, as a consequence of a high-computational burden, only a comparably rather limited number of computationally complex marine ecosystem models have been assessed so far by means of exhaustive optimization runs.

The problem of hardly manageable high-computational complexity has been long recognized. Different solutions have been proposed to address this problem. One solution is to exploit the transport matrix approach for passive tracers (Transport Matrix Method) as has been introduced before (Khatiwala et al., 2005). Another approach – suitable for steady-state rather than for transient tracer simulations – is to exploit a coarse-resolution

[1]Source: cf. Koziel and Yang (2011), p. 196–198.

model initially in the effort to seek a model state already close to the desired periodic solution. This model state can subsequently be utilized as initial condition for the high-resolution model simulation, respectively (see again, with further references, Khatiwala et al., 2005). Yet another, rather straightforward strategy to obtain a computationally cheaper *low-fidelity* (or *coarse*) model, is to exploit a coarser temporal and/or spatial resolution of the model under consideration (while, typically, using the same simulation tool as for the fine model).

Generally, such approaches pursue the primary objective to reduce the computational expense of a single model evaluation. However, computationally cheaper coarse models – e.g., obtained by a coarser discretization – are usually not sufficiently accurate to directly exploit them in a classical optimization loop in lieu of the original fine model. The optimal solution obtained by coarse model optimization might provide a rather inaccurate approximation of the (typically local) fine model optimum only. Most likely, a subsequent and usually expensive fine model optimization is required to locate the fine model optimum more precisely. Therefore, although a coarse model is used initially, the overall optimization costs can be still comparably high. Thus, clearly desirable is the investigation of computationally efficient optimization methods that would allow to exploit a computationally cheap but inaccurate coarse model in an optimization in an efficient way to obtain a sufficiently accurate solution. This serves as the fundamental motivation for the studies presented in this work.

Surrogate-Based Optimization

Computationally efficient optimization of expensive simulation models can be realized using *surrogate-based optimization*, a widely and very successfully used methodology in engineering sciences (see, e.g., Bandler et al., 2004; Forrester and Keane, 2009; Koziel and Yang, 2011; Leifsson and Koziel, 2010; Queipo et al., 2005). Surrogate-based optimization shifts the computational burden from the accurate and expensive fine model to its fast and yet reasonably accurate *surrogate*. As the surrogate models are typically computationally much cheaper than the fine ones and only a few iterations are usually required, the cost of the optimization process could be greatly reduced.

More specifically, the idea of surrogate-based optimization is to replace the fine model in the optimization process by its surrogate in the sense of providing predictions of the fine model optimum (cf. Figure 1b). In particular, the surrogate at the iterate $\mathbf{u}_k$, in the following denoted by $\mathbf{s}_k(\mathbf{u})$, is constructed typically using available fine model data from the current and possibly also from previous iterates, $\mathbf{u}_k$ and $(\mathbf{u}_i)_{i=0,\ldots,k-1}$. The next iterate $\mathbf{u}_{k+1}$ in a surrogate-based scheme is usually obtained by optimizing the surrogate

$\mathbf{s}_k$, i.e.,

$$\begin{aligned} &\mathbf{u}_{k+1} = \underset{\mathbf{u}}{\operatorname{argmin}}\, J\left(\mathbf{s}_k(\mathbf{u})\right) \quad \text{s.t. constraints,} \\ &k = 1, 2, \ldots . \end{aligned} \tag{2}$$

where J denotes the cost function, typically the same as in (1). The process (2) of updating the surrogate and subsequent optimization is usually *iterated* until a user-defined termination condition is satisfied. As for the fine model optimization, this condition can be based on certain convergence criteria (e.g., a threshold for the step size), an assumed level of the objective function value or a specific number of iterations (particularly if the computational budget of the optimization process is limited). A well-performing surrogate-based algorithm is capable of yielding a reasonably accurate solution at a low computational cost, typically corresponding to a few evaluations of the fine model only.

Function-Approximation Surrogates

One possibility to create a surrogate is by approximating sampled fine model data using suitable techniques, e.g., *Polynomial Regression* (Queipo et al., 2005), *Kriging* (Simpson et al., 2001) or *Support-Vector Regression* (Smola and Schölkopf, 2004). These so-called *function-approximation surrogates* (or *functional surrogates*) are constructed without any particular knowledge of the system and are thus rather easily transferable to other application areas. Since these surrogates are based on sampled data only, they do not inherit any physical characteristics of the original fine model and normally require a substantial amount of fine model data samples to ensure a sufficiently accurate approximation. Thus, their use to ad-hoc optimization may be questionable.

Physics-Based Surrogates

Another possibility, explored in this work, is to construct the surrogate from a *physics-based* coarse model, a usually computationally much cheaper but – on the other hand – less accurate representation of the fine model. These type of models are referred to as *physics-based surrogates*. Since the accuracy of the coarse model is typically not sufficient to directly replace the fine model in an optimization loop, it is then necessary to use suitable alignment/correction techniques to reduce the misalignment between the coarse and fine model responses and to ensure that the corrected coarse model (the surrogate) provides a reliable prediction of the fine model optimum.

There exist several methods to construct the surrogate from a physics-based coarse model. *Space Mapping*, as one of the probably most recognized surrogate-based optimization techniques exploiting physics-based coarse models, has been initially introduced by

Bandler et al. (1994) in the context of circuit optimization as a fundamental new theory. The concept therein is a *mapping* relating the fine and coarse model parameters that is exploited to correct a physics-based coarse model. This mapping involves a so-called *parameter extraction* process which typically requires a non-linear optimization process itself. The bulk of the computational effort is involved in the parameter extraction which is carried out in the coarse model space. For a well-performing algorithm, a rapidly improved design can typically be achieved following each evaluation of the fine model (Bandler et al., 2004). Many different formulations for the parameter extraction and yet other more advanced Space Mapping methodologies have been proposed (see again, with further references Bandler et al., 2004). Apart from Space Mapping, other algorithms in the context of surrogate-based optimization originated. They include various *Response Correction* techniques (Søndergaard, 2003), *Manifold Mapping* (Echeverria and Hemker, 2008), and *Shape-Preserving Response Prediction* (Koziel, 2010b; Leifsson and Koziel, 2010). The selection of the underlying coarse model and an appropriate response correction technique is usually problem-specific.

The coarse model can be created in various ways. The simple and straightforward approaches include the use of a *coarser discretization* in time and/or space or by a *relaxed convergence* criterion used in the solution of the model state. Typically, for both approaches, the same simulation tool as for the fine model can be employed. Moreover, *simplified physics* or different ways of describing the same physical phenomenon or even *analytical formulas*, if available, can be used. However, the development of such simplifications most likely requires greater knowledge about the physical system under consideration.

Physics-based surrogates inherit relevant physical characteristics of the original fine model so that only a few fine model data is necessary to ensure a sufficient accuracy. Also, the generalization capability of the physics-based surrogates, i.e., the quality of their approximations of the fine model in the current iterate and in some vicinity, is typically much better than for functional ones. As a result, surrogate-based schemes exploiting physics-based surrogates normally require a small number of fine model evaluations to yield a satisfactory solution. However, the reuse of physics-based surrogates across different problems is typically less straightforward compared to functional ones. Key prerequisites to ensure a good performance of a surrogate-based optimization algorithm employing physics-based surrogates, both in terms of low computational complexity and the quality of the final solution, are a cheap and yet reasonably accurate coarse model as well as a properly selected and "low-cost" alignment procedure (i.e., using a limited number of fine model evaluations per iteration, preferably just one).

Consistency Conditions and Convergence

For a well-performing SBO algorithm, the optimization of the surrogate in one iteration will lead to new parameters ensuring a closer agreement of the desired quantities with the given specifications. If the optimization problem is formulated as minimization, this corresponds to a decrease in the objective function value. More precisely, the surrogate-based scheme (2) is provably convergent to at least a local optimum of the original fine model optimization problem (1), provided that the surrogate $\mathbf{s}_k$ satisfies so-called *zero-* and *first-order consistency* conditions with the original fine model response $\mathbf{y}$ at the current iteration point $\mathbf{u}_k$ (i.e., agreement between function values and the first-order derivative). This can be mathematically written as

$$\mathbf{s}_k(\mathbf{u}_k) = \mathbf{y}(\mathbf{u}_k), \quad \mathbf{s}_k'(\mathbf{u}_k) = \mathbf{y}'(\mathbf{u}_k)\,. \tag{3}$$

Furthermore required for convergence is that the surrogate-based algorithm (2) is embedded in the trust-region approach as well as that mild conditions regarding the coarse and fine model smoothness are ensured (Conn et al., 2000; Koziel et al., 2010). To briefly recall, employing a trust region means restricting the parameters $\mathbf{u}$ in the optimization loop (2) to some model-trust radius. This radius is updated after each iteration, i.e., decreased if a reduction in the objective function value can not be achieved or is not sufficient, and increased otherwise. Typically, standard updating rules can be employed (see again, Conn et al., 2000; Koziel et al., 2010).

In (3), $\mathbf{y}'$ and $\mathbf{s}_k'$ denote the derivatives of the fine model and surrogate's response with respect to the parameter vector $\mathbf{u}$ and at the point $\mathbf{u}_k$, i.e., generally given as

$$\mathbf{y}'(\mathbf{u}_k) := \left.\frac{\mathrm{d}\,\mathbf{y}}{\mathrm{d}\,\mathbf{u}}\right|_{\mathbf{u}=\mathbf{u}_k}. \tag{4}$$

However, in practice, exact sensitivity information may not be obtainable, e.g., if the derivatives are calculated using finite differentiation. In such cases, the first-order consistency condition in (3) only holds approximately. While this may not be sufficient for "theoretical" convergence, the use of the trust-region approach and even approximate sensitivity can substantially improve the algorithm performance (see, e.g., Koziel et al., 2010).

STUDY DESIGN AND RESULTS

This work comprises the investigation and application of surrogate-based optimization methodologies employing physics-based coarse models. Seeking a computationally efficient calibration of marine ecosystem models serves as the fundamental aim.

Overview

As a first case study, I consider a rather simple but nevertheless widely used nitrogen-budget ecosystem model simulating the dynamical evolutions of four tracers, dissolved inorganic nitrogen, phytoplankton, zooplankton, and dead material (detritus). The therefore termed *NPZD model* – developed by Oschlies and Garçon (1999) – simulates the tracer concentrations in one water column at a given horizontal position. This is motivated by the fact that there have been special time series studies, more specifically the *Bermuda Atlantic Time-Series Study* (BATS), located at $31°N$, $64°W$ (see, e.g., Schartau and Oschlies, 2003). Furthermore, a constant vertical sinking velocity of detritus is incorporated in the model. The NPZD model is coupled with a high-resolution ocean circulation model (Modular Ocean Model, version 1.1) which originates from the Geophysical Fluid Dynamics Laboratory (Pacanowski et al., 1991). For the numerical solution of the underlying time-dependent transport-reaction equations, a standard Euler time-stepping scheme and time-dependent (but, non-periodic) forcing data are exploited (see, e.g., Bärwolff, 2007; Stoer and Bulirsch, 2002). Exhaustive optimization runs by using both local, gradient-based and global, genetic algorithms have been previously performed for the considered NPZD model (see, e.g., Schartau and Oschlies, 2003; Rückelt et al., 2010; Schartau, 2001).

NPZD models are widely used in various spatial resolutions from vertically integrated

and one-dimensional models to coupled hydrodynamical three-dimensional versions (see, with further references, Fennel et al., 2001). Clearly, the computational effort of a one-dimensional simulation is significantly smaller than of three-dimensional high-resolution models. However, the complexity of the response of the selected NPZD model is comparably high. Moreover, biochemistry mainly happens locally in space. Thus, this model serves as a suitable test example, before investigating computationally more expensive three-dimensional models (compare with, e.g., Oschlies and Garçon, 1999).

As one possible basis to create a surrogate, I exploit a coarser temporal mesh discretization of the 1D NPZD model. This is one of rather straightforward ways to introduce a physics-based coarse model for such time-dependent systems. Moreover, compared to other strategies such as replacing the specific model equations by simplified descriptions, this approach has the key advantage to be more versatile and thus probably suitable for yet other transient marine ecosystem models. Accordingly, the NPZD model with the originally employed time discretization serves as the fine model here.

As a second case study, I focus on a three-dimensional coupled hydrodynamic marine ecosystem model. More specifically, I choose an *N-DOP model*, simulating the transport and biogeochemical cycles among nitrogen and dissolved organic phosphorus (see, e.g., Kriest et al., 2010). The model is coupled with a general ocean circulation model in an off-line mode. The transport of the biogeochemical tracers is provided through transport matrices as briefly explained in the introduction (Khatiwala et al., 2005). For the N-DOP model, a steady annual cycle – here, using a classical fixed point iteration – is simulated. The model simulation is implemented as part of the simulation package of *Metos3D* (Marine Ecosystem Toolkit for Simulation and Optimization in 3-D), developed by Piwonski and Slawig (2011). The choice of this specific model has been motivated by the fact that it is a rather simple representative of the class of global marine ecosystem models. Moreover, the N-DOP model serves as a suitable basis for models with a yet higher complexity of the biogeochemical coupling (see again, Kriest et al., 2010). Thus, this model provides a suitable first test case here. For the N-DOP model, I investigate yet another approach to obtain a physics-based coarse model. I exploit a truncated spin-up, which is realized by simply cutting down the number of fixed point iterations, or, equivalently, by weakening the corresponding termination condition. For this model, I accordingly treat the model solution which is converged up to a pre-defined and sufficient accuracy as the reference fine model.

For the selected marine ecosystem models, I will principally investigate two correction approaches – *Aggressive Space Mapping* and *Multiplicative Response Correction* – in the effort to reduce the misalignment between the chosen coarse and the corresponding fine models. It will be shown that the corrected coarse models, i.e., the physics-based

surrogates, provide reliable approximations of the corresponding fine models. Subsequently, their application in a surrogate-based scheme as in (2) will be investigated (cf. Figure 1b). Moreover, I will assess the applicability of the corresponding coarse models as direct replacement of the fine models in the optimization (cf. Figure 1a). The quality of the solution obtained by surrogate-based optimization as well as the algorithms' performance in terms of the computational costs will be compared to what can be achieved by direct coarse and fine model optimization (in case the latter is computationally affordable).

The focus of this work clearly is demonstrating the applicability of the proposed methodologies to the parameter optimization of the considered models and not the actual interpretation of the obtained results in the biogeochemical context. Principally, it will be investigated, whether the presented surrogate-based optimization methodologies allow to yield a reasonable solution at low optimization costs.

Aggressive Space Mapping

In the first part, I focus on one of the original Space Mapping approaches – Aggressive Space Mapping (see, e.g., Bandler et al., 2004) in Paper 1 (Appenix A.1). I investigate the applicability of this approach to achieve a computationally efficient parameter optimization of the 1-D NPZD model. The Aggressive Space Mapping algorithm evolves from the original formulation of Space Mapping (Bandler et al., 1994, 2004). Therein, some parameter mapping, $\mathbf{p} : \mathbf{u} \rightarrow \hat{\mathbf{u}}$, from the fine to the coarse model parameters ($\mathbf{u}$ and $\hat{\mathbf{u}}$, respectively) is proposed as

$$\mathbf{p}(\mathbf{u}) = \hat{\mathbf{u}}, \tag{5}$$

such that the mapped coarse model – the surrogate – provides an approximation of the fine model $\mathbf{y}$, i.e.,

$$\mathbf{y}(\mathbf{u}) \approx \hat{\mathbf{y}}\left(\mathbf{p}(\mathbf{u})\right) \tag{6}$$

in a region of interest, with $\hat{\mathbf{y}}$ denoting the coarse model response. The idea of the original Space Mapping approach is to first solve for a coarse model optimum $\hat{\mathbf{u}}^*$ and to subsequently seek a solution $\bar{\mathbf{u}}$ of the following non-linear system of equations

$$\mathbf{p}(\bar{\mathbf{u}}) = \hat{\mathbf{u}}^* \,. \tag{7}$$

Since the mapping $\mathbf{p}$ is defined such that the similarity (6) holds, solving the above equation (7) is equivalent to solving for

$$\mathbf{y}(\bar{\mathbf{u}}) \approx \hat{\mathbf{y}}(\hat{\mathbf{u}}^*) \,. \tag{8}$$

This can be realized by, e.g., minimizing the difference in a least-squares type objective. Thus, the solution $\bar{\mathbf{u}}$ of (7) is close to the fine model optimum $\mathbf{u}^*$, provided that the fine and the coarse model responses are close to each other at their respective optima, i.e.,

$$\mathbf{y}(\mathbf{u}^*) \approx \hat{\mathbf{y}}(\hat{\mathbf{u}}^*)\,. \tag{9}$$

In this case, the mapping $\mathbf{p}$ satisfies

$$\mathbf{p}(\mathbf{u}^*) \approx \hat{\mathbf{u}}^*,$$

which is sometimes referred to as *perfect mapping* (see, e.g., Echeverría and Hemker, 2005). The connection of this original Space Mapping approach to the introduced strategy of surrogate-based optimization is as follows: The solution of (7) is equivalent to exploiting the mapped coarse model as a surrogate, i.e,

$$\mathbf{s}_k(\mathbf{u}) := \hat{\mathbf{y}}(\mathbf{p}_k(\mathbf{u})),$$

in a surrogate-based scheme as in (2), provided that the mapping is injective and the coarse model optimum $\hat{\mathbf{u}}^*$ is unique (see here and the following, e.g., Bandler et al., 2004; Echeverría and Hemker, 2005). Here, $\mathbf{p}_k(\mathbf{u})$ denotes some approximation of the mapping $\mathbf{p}$ in the point $\mathbf{u}_k$.

The Aggressive Space Mapping as a particular approach now solves for a solution of the non-linear system of equations given in (7), using a first-order Taylor approximation of the mapping, i.e.,

$$\mathbf{p}_k(\mathbf{u}) := \mathbf{p}(\mathbf{u}_k) + \mathbf{p}'(\mathbf{u}_k)(\mathbf{u} - \mathbf{u}_k)$$

and exploiting a Quasi-Newton iteration (Kosmol, 1993; Nocedal and Wright, 2000). Furthermore, a Broyden rank-one approximation is used to approximate the derivative $\mathbf{p}'$ of the mapping (Bandler et al., 2004; Broyden, 1965). Since the standard Quasi-Newton algorithm may suffer from local convergence, it is typically reasonable to use some globalization strategies such as a trust-region or a line-search approach. In this work, I will exploit the latter.

For the considered temporally coarser discretized NPZD model, the required similarity in (9) is clearly related to the fineness of the considered coarser temporal resolution. Roughly speaking, the coarse model has to be accurate enough, i.e., the coarse temporal resolution has to be fine enough, such that its response contains the relevant characteristics of the corresponding fine one. This is supported by the optimization results obtained for the chosen discretization which has been concretely utilized for this approach. More

specifically, when applying the Aggressive Space Mapping to the calibration of the NPZD model, I demonstrate that this approach allows to obtain a solution fairly close to the fine model optimum. A reasonable reduction in the optimization costs of about 35% can be achieved, when compared to a direct fine model optimization. The performance of this algorithm is verified using model-generated, i.e., attainable, measurement data.

Multiplicative Response Correction

A second approach for the coarse model alignment comprises a multiplicative response correction technique. Here, the surrogate $\mathbf{s}_k$ in iteration k of the optimization loop is generated through a *multiplicative correction* of the coarse model response $\hat{\mathbf{y}}$, which can be briefly formulated as

$$\mathbf{s}_k(\mathbf{u}) := \mathbf{a}_k \, \hat{\mathbf{y}}(\mathbf{u}) \,. \tag{10}$$

I define the correction term $\mathbf{a}_k$ as the point-wise division of the fine by the coarse model response at the point $\mathbf{u}_k$, i.e.,

$$\mathbf{a}_k := \frac{\mathbf{y}(\mathbf{u}_k)}{\hat{\mathbf{y}}(\mathbf{u}_k)} \,. \tag{11}$$

It turns out that this multiplicative approach is quite suitable for both, the coarser discretized NPZD model as well as for the coarse N-DOP model obtained from a truncated spin-up. The reason is that for both models, the overall "shape" of the coarse model responses resemble that of the fine ones and the relation between the coarse and the fine model responses is rather well preserved while moving from one parameter vector to another. In principle, multiplicative response correction is a convenient way of adjusting the model response level without distorting the responses' relevant characteristics (such as minima and maxima) during the optimization process.

By definition, the surrogate in (10) ensures the zero-order consistency condition in (3) with the fine model response $\mathbf{y}$ in the current iteration point $\mathbf{u}_k$ (i.e., agreement between function values of the surrogate and the fine model). Let me recall that, formally, this is not sufficient to ensure the convergence of a surrogate-based scheme exploiting (10) to a (local) minimum of the corresponding fine model optimization problem. However, since the physics-based surrogate inherits substantial knowledge about the marine ecosystem model under consideration, its derivatives are expected to be at least similar to those of the fine model. Furthermore, because of being constructed from a physics-based coarse model, the surrogate exhibits quite good generalization capability. This means that it provides a reliable approximation of the fine model when moving from one parameter

vector to another. Thus, for physics-based surrogates, even an approximate agreement in the derivatives can be sufficient to allow for a good optimization performance.

However, if also exact first-order consistency is desired, an additive term E_k including fine and coarse model sensitivity information can be subsequently employed in the surrogate's construction as

$$\mathbf{s}_k(\mathbf{u}) := \mathbf{a}_k \, \hat{\mathbf{y}}(\mathbf{u}) + E_k(\mathbf{u} - \mathbf{u}_k) \, . \tag{12}$$

Clearly, on the one hand, employing the typically expensive fine model sensitivity information increases the computational costs for the surrogate's construction. But, on the other hand, employing an enhanced surrogate as in (12) and embedding the surrogate-based algorithm in the trust-region approach can allow to locate the fine model optimum even more precisely, if desired. Convergence of the algorithm to a local minimum of the fine model optimization problem can be attained, typically, even if sensitivity information can only be obtained approximately. In order to assess the algorithm's performance, the trade-offs between the solution's accuracy and the extra computational overhead related to sensitivity calculation will have to be addressed for the specific optimization problem.

1-D NPZD Model

In a first step, a surrogate as proposed in (10) is exploited in conjunction with the initially proposed temporally coarse NPZD model in Paper 2 (Appenix A.2). Due to the larger time step employed in the numerical solution of the coarse model, its response is rather inaccurate. Also both, the fine and coarse model responses, contain numerical noise which is misleading while performing the coarse model alignment as in (10). It will been demonstrated that an adequate alignment should be based on the main characteristics of the fine model response, which can be extracted by "smoothing". Inaccurate, rather unphysical peaks in the coarse model responses will be removed. In order to obtain a commensurable fine model response in (11), smoothing is also applied to this response.

Again, as a first test scenario, the performance of the algorithm is investigated by considering model-generated measurement data. As will be demonstrated, a zero-order consistent surrogate as in (10) is sufficient for good performance of the surrogate-based optimization. More specifically, I will demonstrate by illustrative optimization runs that a solution close to the one obtained by fine model optimization can be obtained at a few fine model evaluations only. On the other hand, coarse model optimization is cheap but its solution is rather inaccurate. Using surrogate-based optimization, the optimization costs can be further reduced down to 15% and therefore significantly decreased compared to what has been obtained by the Aggressive Space Mapping algorithm.

In an additional part, I demonstrate that by employing certain modifications of the original correction scheme as given in (11), one can further improve the surrogate's accuracy as well as reduce the computational cost of the optimization process. The optimization costs can be further reduced down to 5% and thus the performance of the algorithm can be improved compared to the original correction scheme. This work is covered in Paper 3 (Appenix A.3)

Furthermore, the application of this surrogate-based approach to real, measurement data is outlined in Paper 4 (Appenix A.4). Experiments carried out for this problem reveal that implementation of the first-order consistency in (3) is important to ensure sufficient performance of the algorithm. Thus, the surrogate model in (10) is enhanced by using fine and coarse model sensitivity information as proposed in (12). Furthermore, the algorithm is embedded in the trust-region approach as explained before. The enhancements of the surrogate model and the optimization algorithm (i.e., the use of fine model sensitivity and the trust-region approach) are essential to calibrate the selected NPZD model against the real data. Specifically, they ensure a sufficiently accurate solution of the surrogate-based optimization while retaining the high computational savings when compared to a direct fine model optimization. The trade-offs between the solution's accuracy and additional computational expenses due to fine model sensitivity calculations will be addressed.

3-D N-DOP Model

In the final part of my work, which is covered in Paper 5 (Appenix A.5), I employ the successful multiplicative response correction approach to parameter optimization of the selected three-dimensional coupled marine ecosystem model.

I demonstrate that the proposed response correction approach can reduce the misalignment between the coarse (here, obtained by a truncated spin-up) and fine model leading to a reliable approximation of the original, fine ecosystem model. For the considered coarse N-DOP model, smoothing seems to be not necessary which is mainly due to the fact that the discretization is not coarsened in this case. I subsequently present the results of an illustrative surrogate-based optimization run using model-generated measurement data to initially verify this approach. As for the 1-D NPZD model, I start with the computationally less expensive surrogate as proposed in (10) since this might ensure sufficient performance of the algorithm. Solutions again demonstrate that exploiting this surrogate in a surrogate-based optimization run allows to obtain a remarkably accurate solution at the cost of a few evaluations of the fine model only. Whereas a direct optimization run using the fine model in a classical optimization loop (cf. Figure 1a) would typically require several weeks on a 48-processor cluster, the computational cost using such a surrogate can be significantly reduced down to a few hours.

Numerical Stability

Another investigation, which is not an integral part of this work, is related to the *numerical stability* of the considered NPZD model.

Numerical stability is a property of numerical algorithms. Roughly speaking, its definition is related to the accuracy of the algorithm (see here and the following, e.g., Bärwolff, 2007; Golub and Ortega, 1995; Großmann and Roos, 2005; Marchuk, 1982; Stoer and Bulirsch, 2002). Typically, in numerical mathematics, an algorithm is denoted as numerically stable, if it is insensitive with respect to small perturbations/errors. These approximation errors can be round-off or truncation errors due to the internal representation of numeric values in digital computers (*computer number format*). Dependent on the specific numerical method, these errors might be magnified, which causes the error to grow and hence the numerical solution to be invalid.

For marine ecosystem models, the chosen spatial and/or temporal mesh discretization of the model and, in the case of a transient simulation, also the formulation of the underlying time integration schemes, strongly determine the numerical stability of the model simulation. Numerical noise in the model responses is misleading during the optimization process. It can slow down the algorithm's convergence and make the optimum more difficult to locate. Thus, numerical stability of a temporally coarser discretized marine ecosystem model becomes crucial to ensure a physically reasonable model response and a well performing optimization process when exploiting the coarse model for the model calibration, no matter whether "purely" or as corrected/aligned surrogate. This, for example, was taken into account for the considered NPZD model in this work.

It has been investigated that only the numerical method used for the vertical advective transport (here, an *explicit* Euler time-stepping scheme) seems to be actually critical in terms of instabilities in the NPZD model response. This has been primarily concluded from visual inspections of the coarse model responses and from various numerical experiments. The reason is that the utilized implicit scheme for the diffusive processes does not cause numerical instabilities. Furthermore, the biogeochemical coupling does not seem to measurably affect the numerical stability. More specifically, a rather standard stability condition for time-dependent "advection-type" transport models (see, e.g., Großmann and Roos, 2005, p. 48) can ensure a numerically stable model response and, most importantly, a reasonable optimization performance. This condition in principle incorporates a certain relation between the spatial and temporal discretization as well as the vertical velocity. These results were taken into account in all subsequent investigations and optimization runs where a temporally coarser discretization of the NPZD model has been considered to create a suitable coarse model.

In terms of yielding a rather flexible model, the independence of the numerical stability

of quantities in the numerical model such as the mesh discretization is clearly desirable. Most importantly in the context of surrogate-based optimization, this would allow to exploit an even coarser temporal resolution to create a physically yet reasonable coarse model. As a consequence, the computational expense of a corresponding coarse model evaluation and, accordingly, of surrogate-based optimization runs exploiting these models, could be further significantly decreased. In order to address this issue, I furthermore investigated a modification of the originally exploited explicit time integration approach for the vertical advection by exploiting an *implicit* Euler scheme instead. It turned out that this enhancement allows to obtain a numerically stable solution without restrictions to the mesh discretization and the vertical velocity. Again, this was obtained from numerical experiments.

Clearly, a more thorough investigation as well as a more profound mathematical analysis of the numerical stability of the original and improved time integration schemes will be useful to substantiate those results. This will be particularly important in order to demonstrate the full capabilities of the proposed physics-based surrogates where the underlying coarse models are obtained from a temporally and/or spatially coarser discretization. Considerable initial steps in this direction have already been done, in terms of numerical experiments as noted above as well as in terms of a mathematical investigation. Continuing work in the effort to obtain a more complete analysis will be highly valuable.

OUTLOOK

The proposed methodologies in the framework of surrogate-based optimization using physics-based surrogate models, particularly the multiplicative response correction approach, turned out to have great potential for a computationally efficient calibration of marine ecosystem models.

Regarding the optimization of the considered three-dimensional coupled marine ecosystem model, enhancement of the present approach by coarse/fine model sensitivity would probably allow to locate the fine model optimum even more accurately, if desired. Employing those improvements, an investigation of the trade-offs between the solution accuracy and the extra computational overhead related to sensitivity calculation will be useful. Also, considering other ways of reducing the evaluation cost of the coarse model such as coarsening the temporal and/or spatial discretization – possibly also in combination with a truncated spin-up – will be of great interest. Such approaches may further improve the cost savings while still ensuring a sufficiently accurate solution. Moreover, the application to real measurement data will be necessary to demonstrate the full capabilities of the multiplicative response correction for the calibration of the considered coarse N-DOP model.

Furthermore, I expect different modifications of the current implementation of the Aggressive Space Mapping algorithm to bring further enhancement of the optimization performance and corresponding cost savings. One of possible improvements could include smoothing of the rather inaccurate coarse model responses as it was later done for the proposed multiplicative response correction approach. Also, exploiting the *Trust-Region Aggressive Space Mapping* might be suitable, where the algorithm is embedded in a trust-region approach (see here and the following, e.g., Bandler et al., 2004). A more accurate approximation of the jacobian of the mapping by utilizing coarse and fine model sensi-

tivity information might further improve the algorithms' performance. Since parameter extraction is a crucial part of the Aggressive Space Mapping approach, non-uniqueness of its solution can be problematic. Amongst other strategies, *multipoint parameter extraction* and for example *penalty parameter extraction* have been suggested to address this issue (see again, with further references, Bandler et al., 2004). Moreover, the investigation of yet other, more advanced Space Mapping approaches might have great potential.

Possible enhancements by other strategies in the effort to further decrease the overall computational costs will be valuable. These could include the re-use of the computationally expensive fine model sensitivity information (if used) within subsequent iterations. This might be sufficient to allow for a convergent algorithm and accurate solution. Also, employing a simple first-order Taylor approximation of the fine model to create yet another surrogate in the final stage of the algorithm, i.e., in a region close to the fine model optimum, might further decrease the computational complexity of the optimization process. Also, it could allow to locate the fine model optimum even more precisely. Here, initial numerical experiments and optimization runs have already been performed including some of the above enhancements.

As noted, an implicit time-stepping scheme for the integration of the vertical advection processes in the transient 1D NPZD model, can be highly valuable. This would allow to exploit an even coarser temporal resolution to create a physically yet reasonable coarse model. As a direct consequence, the computational expenses of a surrogate-based algorithm exploiting such coarser discretized models could be further reduced compared to the results presented in this work.

Moreover, other ways to construct a physics-based coarse model, such as replacing the model equations by simplified or even analytical formulations of the relevant processes, can be promising. Clearly, such approaches will require substantially more knowledge about the marine ecosystem model and the incorporated processes and thus a larger effort in the coarse model construction initially. However, the computational costs for such coarse models can be much lower compared to those obtained by a coarser mesh discretization, for example. Thus, in conjunction with suitable correction techniques, exploiting such models in a surrogate-based optimization framework might be very promising in terms of a significant speedup of the optimization process.

Last but not least, the application of the proposed methodologies to different marine ecosystem models for exhaustive model-data comparison studies will be useful to demonstrate the full capabilities of the introduced approaches. Particularly, for the so far hardly manageable optimization of coupled marine ecosystem models at a sufficiently high resolution, the proposed approaches might provide first parts of a highly attractive toolkit for a computationally efficient optimization.

APPENDIX

Aggressive space mapping for optimisation of a marine ecosystem model

M. Prieβ* and T. Slawig

Department of Computer Science,
Algorithmic Optimal Control – CO_2 Uptake of the Ocean,
Excellence Cluster The Future Ocean,
Christian-Albrechts-Platz 4, 24118 Kiel, Germany
E-mail: mpr@informatik.uni-kiel.de
E-mail: ts@informatik.uni-kiel.de
*Corresponding author

Abstract: We apply the aggressive space mapping (ASM) algorithm to parameter optimisation of a one-dimensional marine ecosystem model. Such models are important to calculate the global carbon cycle and specifically the oceanic CO_2 uptake. The aim of the optimisation is to find model parameters that minimise the misfit between model output and observational data. ASM first solves for an optimum of a computationally cheaper low-fidelity model that in our case is based on a coarser time discretisation while still guaranteeing numerical stability. Secondly, an approximate optimum of the high-fidelity model is iteratively found by a parameter mapping applying a Quasi-Newton method and Broyden update. The applicability of the ASM technique to the problem is verified by using synthetic target data. Results are compared to those of the direct high-fidelity model optimisation. We show that a very reasonable solution can be obtained while yielding a significant reduction in the total optimisation cost.

Keywords: marine ecosystem models; surrogate-based optimisation; low-fidelity model; aggressive space mapping; ASM; globalised Quasi-Newton; numerical optimisation; inverse problems.

Reference to this paper should be made as follows: Prieβ, M. and Slawig, T. (2012) 'Aggressive space mapping for optimisation of a marine ecosystem model', *Int. J. Mathematical Modelling and Numerical Optimisation*, Vol. 3, Nos. 1/2, pp.98–116.

Biographical notes: M. Prieβ is currently a PhD student in the research group *Algorithmic Optimal Control – CO_2 Uptake of the Ocean* at the Christian-Albrechts-Universität zu Kiel in the Institute for Computer Science. His research interests lie in the field of parameter optimisation in climate models, surrogate-based optimisation and algorithms for non-linear optimisation.

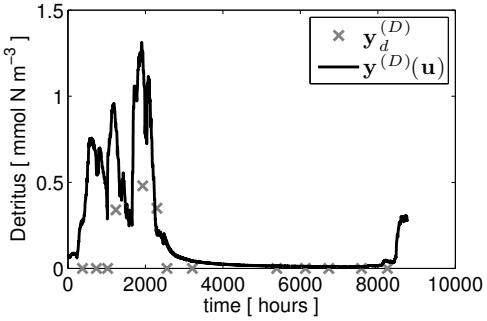
T. Slawig is a Professor at the Christian-Albrechts-Universität zu Kiel in the Institute for Computer Science. He is the Leader of the research group *Algorithmic Optimal Control – CO_2 Uptake of the Ocean* in the Cluster *Future Ocean*. His research interests lie in the field of optimal control for partial differential equations, optimisation and algorithmic differentiation.

1 Introduction

Understanding the oceanic CO_2 uptake ß is of central importance for projections of climate change and oceanic ecosystems. Simulating ocean circulation and biogeochemistry has become a key tool for understanding the ocean carbon cycle and its variability (Sarmiento and Gruber, 2006). The underlying models are governed by coupled systems of parabolic partial differential equations for ocean circulation (ocean models) and transport of biogeochemical state variables or so-called tracers (marine ecosystem models) (Fennel and Neumann, 2004; Sarmiento and Gruber, 2006). The coupling relations between the tracers are more or less empirical, i.e., it is not very clear how the coupling terms look like mathematically, and, moreover, how many tracers have to be taken into account. In ecosystem models many parameters are used which are chosen such that given measurement data are matched, and that the model output remains feasible, i.e., non-negative.

For this purpose, the aim is to minimise a least-squares type cost functional, measuring this misfit (cf., Figure 1), and optionally constrained by inequalities for the parameters and/or state variables (Tarantola, 2005). The optimisation variables are unknown physical/biological parameters in the non-linear coupling terms in the tracer transport equations. This optimisation process requires a lot of – typically expensive – function and optionally sensitivity or gradient evaluations. If the latter are computed by finite difference approximations, the critical quantity regarding the computational cost of optimisation is the one needed for one function evaluation, which is basically one model run. Hence a big issue in order to reduce the overall optimisation cost is to decrease the effort for the function evaluations. This in particular becomes significant for computationally expensive three-dimensional coupled models.

Figure 1 Model output $\mathbf{y}$ and target data $\mathbf{y}_d$ for the state detritus for one year at depth $z \approx -25$ m



One approach to pursue this aim is to replace the original model in focus (also called *high-fidelity* or *fine model*) by a computationally cheaper so-called *surrogate* (Bandler et al., 2004a; Forrester and Keane, 2009; Queipo et al., 2005). To create this surrogate, either approximations of sampled fine model data can be used or a *low-fidelity* or *coarse model*, which is usually less accurate, is introduced and iteratively corrected by suitable methods. For this correction or alignment, only a few evaluations of the fine model and possibly also its derivatives are necessary. Apart from this alignment, the whole optimisation process is performed in the surrogate's model space which could dramatically reduce the overall cost. Surrogate-based optimisation is widely and very successfully used in engineering sciences (Bandler et al., 2004b; Forrester and

Keane, 2009; Queipo et al., 2005). The application on parameter optimisation in climate models (e.g., a marine ecosystem model as used here) is rather new.

In this paper, we focus on the so-called aggressive space mapping (ASM) technique (Bandler et al., 1994, 1995, 2004a). A coarse model, which is based on the same physics as the fine model, is corrected by a parameter mapping, introduced by Bandler et al. (1994), to yield a surrogate. In particular, the ASM algorithm solves a non-linear system of equations which is conditionally equivalent to directly use this type of surrogate in the optimisation run as described above. In this work we analyse the application of the ASM approach to parameter optimisation of a one-dimensional ecosystem model (Oschlies and Garcon, 1999).

The structure of the paper is as follows: the ecosystem model is introduced in Section 2. In Section 3, we describe the corresponding optimisation problem. The basic idea of surrogate-based optimisation is recalled in Section 4. We use a coarsening in the temporal mesh to create the basis of our surrogate, the coarse model, which we briefly present in Section 5. The surrogate obtained from this coarse model through a certain parameter mapping is described in Section 6. The basic idea of the ASM algorithm and the equivalence to using this surrogate in the optimisation run is recalled in Section 7. Also the globalised Quasi-Newton procedure (Kosmol, 1993) we use to compute the ASM solution is presented. We verify this approach by using synthetic target data and by comparing the solutions with those obtained from direct fine model optimisation. Corresponding results are given in Section 8. Section 9 concludes the paper.

2 Model description

A one-dimensional marine ecosystem model that simulates the interaction of dissolved inorganic nitrogen, phytoplankton, zooplankton and detritus (thus also called NPZD model) was developed by Oschlies and Garcon (1999), with the aim of simultaneously reproducing observations at three North Atlantic locations by optimisation of free parameters within credible limits. The model uses the ocean circulation and temperature field in an off-line modus, i.e., no feedback on them is modelled. The model simulates one water column at a given horizontal position which is motivated by the fact that there have been special time series studies at fixed locations.

Marine ecosystem models are coupled PDE systems consisting of time-dependent advection-diffusion-reaction equations with non-linear coupling terms. The velocity and temperature and sometimes also salinity data are either computed simultaneously or in advance by an ocean model. Clearly, the second variant, that is used in this paper, is computationally cheaper but neglects the feedback effects from the biogeochemistry to the ocean circulation and temperature distribution, etc.

In the model used here, the concentrations (in mmol N m^{-3}) of dissolved inorganic nitrogen, phytoplankton, zooplankton, and detritus are summarised in the vector $\mathbf{y} := (y^{(l)})_{l=N,P,Z,D}$ and described by the following coupled PDE system

$$\left.\begin{aligned} \frac{\partial y^{(l)}}{\partial t} &= \frac{\partial}{\partial z}\left(K_\rho \frac{\partial y^{(l)}}{\partial z}\right) + Q^{(l)}(\mathbf{y},\mathbf{u}), && l = N,P,Z \\ \frac{\partial y^{(D)}}{\partial t} &= \frac{\partial}{\partial z}\left(K_\rho \frac{\partial y^{(D)}}{\partial z}\right) + Q^{(D)}(\mathbf{y},\mathbf{u}) - \frac{\partial y^{(D)}}{\partial z} u_{12}, && l = D \end{aligned}\right\} \quad \text{in } (-H,0)\times(0,t_e) \tag{1}$$

with additional appropriate initial values. Here z denotes the vertical spatial coordinate, H the depth of the water column, and t_e the total integration time. The sinking term with sinking velocity u_{12} is only apparent in the equation for detritus. In the one-dimensional model no advection term is used, since a reduction to vertical advection only would make no sense.

The $Q^{(l)}$ are the biogeochemical coupling (or *source-minus-sink*) terms for the four tracers. For simplicity, their dependence on space, time and given temperature data is omitted in the notation. To give an impression of the non-linearities and the meaning of the model parameters that are summarised in the vector $\mathbf{u} = (u_1, \ldots, u_n)$ – with $n = 12$ in this model – and that are subject to optimisation, we briefly present the coupling terms [more details can be found in Oschlies and Garcon (1999)]:

$$\begin{aligned}
Q^{(N)}(\mathbf{y}) &= -\min\left(\mu(y^{(P)}), u_2 c^T \frac{y^{(N)}}{u_{11} + y^{(N)}}\right) y^{(P)} + u_4 y^{(Z)} + u_{10} y^{(D)}, \\
Q^{(P)}(\mathbf{y}) &= \min\left(\mu(y^{(P)}), u_2 c^T \frac{y^{(N)}}{u_{11} + y^{(N)}}\right) y^{(P)} - u_8 y^{(P)} - \frac{u_6 u_7 {y^{(P)}}^2}{u_7 + u_6 {y^{(P)}}^2} y^{(Z)}, \\
Q^{(Z)}(\mathbf{y}) &= u_1 \frac{u_6 u_7 {y^{(P)}}^2}{u_7 + u_6 {y^{(P)}}^2} y^{(Z)} - u_4 y^{(Z)} - u_9 {y^{(Z)}}^2, \\
Q^{(D)}(\mathbf{y}) &= (1 - u_1) \frac{u_6 u_7 {y^{(P)}}^2}{u_7 + u_6 {y^{(P)}}^2} y^{(Z)} + u_8 y^{(P)} + u_9 {y^{(Z)}}^2 - u_{10} y^{(D)}.
\end{aligned} \tag{2}$$

The function μ describes the dependency of photosynthesis on the amount of light at depth z. It depends on the value of phytoplankton $y^{(P)}$ by a non-local (integral) relation over the water column. Here two additional parameters u_3, u_5 are envolved. The circulation data (taken from an ocean model) are the turbulent mixing coefficient $K_\rho = K_\rho(z,t)$ and the temperature $T = T(z,t)$ which is used in the non-linear term c^T where $c = 1.066$ is kept constant. Note that besides the non-linearities there is an additional non-differentiability due to the term 'min $(\ldots)$'.

3 Optimisation problem

The aim is to minimise a least-squares type cost function measuring the misfit between the model output $\mathbf{y} = \mathbf{y}(\mathbf{u})$ and given observational data $\mathbf{y}_d$ (Figure 1). Thus the problem can be written as

$$\begin{aligned}
&\min_{\mathbf{u} \in U} \; J(\mathbf{y}(\mathbf{u})), \quad J(\mathbf{y}) := ||\mathbf{y} - \mathbf{y}_d||_Y^2, \\
&U := \{\mathbf{u} \in \mathbb{R}^n : \mathbf{b}_l \leq \mathbf{u} \leq \mathbf{b}_u\}, \quad J : Y \times U \to \mathbb{R}.
\end{aligned} \tag{3}$$

Note that the optimisation variables are real numbers with lower and upper bounds described by the vectors $\mathbf{b}_l, \mathbf{b}_u \in \mathbb{R}^n$. The functional J may additionally include a regularisation term for the parameters, which was not necessary in our case.

The above formulation is valid if the state or vector of tracers $\mathbf{y}$ is regarded

- either in a continuous setting as an element of an appropriately chosen function space Y
- or after discretisation (cf., Section 5) as a discrete vector $\mathbf{y} \approx (\, y(z_i, t_j)\,)_{i=1,\dots,4K,\, j=1,\dots,M}$, i.e., $\mathbf{y} \in \mathbb{R}^{4KM}$ where $4K$ and M denote the total number of spatial and temporal grid points, respectively.

We will consider the latter formulation from now on.

Constraints on the state variable $\mathbf{y}$ are not treated explicitly in our formulation (3). However, by using appropriate parameter bounds $\mathbf{b}_l$ and $\mathbf{b}_u$, the desired non-negativity of the state/tracer vector can be ensured in the model. This was already observed and used in Rückelt et al. (2010).

4 Surrogate-based optimisation

For many non-linear optimisation problems high computational cost of accurate simulations and derivatives or even the lack of sensitivity information are major drawbacks. The need for a decrease in the computational cost is especially important in the case of parameter optimisation of complex three-dimensional models.

In surrogate-based optimisation, the original fine model output $\mathbf{y}$ is replaced by a surrogate. Surrogates can be either based upon approximations of sampled fine model data (*functional* surrogates) or upon a low-fidelity or coarse model [also known as *physically-based surrogates* (cf., Søndergaard, 2003)]. A functional surrogate is constructed without any particular knowledge of the system and will not be addressed further in this paper. In contrast physically-based surrogates inherit more characteristics of the fine model in focus since the underlying coarse model is based upon the same physics as the fine one (*physical* coarse model). Since the coarse model, in the following denoted by $\hat{\mathbf{y}}$, is usually less accurate it has to be iteratively aligned or corrected by suitable methods. In the kth step of an optimisation algorithm, arrived at optimisation variable iterate $\mathbf{u}_k \in U$, we thus perform a step generally written as

$$\mathbf{y}(\mathbf{u}_k),\ \hat{\mathbf{y}}(\mathbf{u}_k) \quad \mapsto \quad \mathbf{s}_k.$$

The next iterate, $\mathbf{u}_{k+1}$, is obtained by optimising this surrogate $\mathbf{s}_k$. Then the updated surrogate $\mathbf{s}_{k+1}$ is determined by re-aligning the low-fidelity model at $\mathbf{u}_{k+1}$ and optimised again. The process of aligning the coarse model to obtain the surrogate and subsequent optimisation of this surrogate is repeated until a user-defined termination condition is satisfied.

In order to obtain convergence to at least a local minimum of (3) the surrogate $\mathbf{s}_k$ should satisfy so-called 0-order and also 1st-order consistency conditions (Conn et al., 2000; Koziel et al., 2010) with the fine model in the current iterate $\mathbf{u}_k$, i.e.,

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

Key prerequisites for a well performing surrogate algorithm are a cheap function and sensitivity evaluation of its basis, the coarse model, a low cost for the alignment of this coarse model, and a low number of necessary iterations in the surrogate-based optimisation process, since this results in only few evaluations of the fine model.

Space mapping (SM) is one approach for obtaining such a surrogate (Bandler et al., 2004a). SM assumes the existence of a physical coarse model (see the following section)

that describes the same physics as the fine model. This coarse model is less accurate, but much faster to evaluate than the fine one. In the context of SM, a surrogate model is constructed from the coarse model in such a way that it is a suitable distortion and that given matching conditions are satisfied. In this paper we will analyse the application of a conditionally equivalent approach to directly use this surrogate in the optimisation run, the ASM (Bandler et al., 1994, 1995, 2004a) which will be described in Section 7.

For a comprehensive overview on methods and recent approaches in SM technology we refer to Bakr et al. (2000), Bandler et al. (2004a, 2004b) and Koziel et al. (2006).

5 The low-fidelity model

In this paper we use a physical coarse model. Possible ways to create such a coarse model are by a coarser discretisation, by using simplified physics or different ways of describing the same physical phenomenon or even by using analytical formulas if available. In this paper we use a coarse model which is based on a coarser time discretisation. For this purpose, we firstly present the discretisation scheme of the original, i.e., fine model (1) and secondly describe the corresponding discretisation of our coarse model.

5.1 Discretisation scheme of the fine model

The fine model described by (1) and (2) is solved using an operator splitting method (cf., Marchuk, 1982), which is explained in this section.

Let a time step $\tau > 0$ be given. Then, at first the non-linear coupling operators $Q^{(l)}$ [cf., (2)] are computed. Since they depend implicitly on time, i.e., on the temporal index j, we denote by

$$Q_j(\mathbf{y}_j) = \left(Q^{(l)}(\mathbf{y}_j)\right)_{l=N,P,Z,D}$$

the vector of the results of these four coupling terms applied on the discrete state $\mathbf{y}_j \approx (\, y(z_i, t_j)\,)_{i=1,\ldots,4K}$ at time step j. Now four explicit Euler steps with step size $\tau/4$ are performed, each of which is described by the operator

$$B_j^Q(\mathbf{y}_j) := \mathbf{y}_j + \tfrac{\tau}{4} Q_j(\mathbf{y}_j).$$

Then, an explicit Euler step with full step-size τ is performed for the sinking term which is spatially discretised by an upstream scheme (cf., Fletcher, 1991). This step is summarised in a matrix, denoted by B^{sink}. Since the sinking velocity is temporarily constant, this matrix does not depend on the time step j.

Finally, an implicit Euler step for the diffusion operator, discretised with second order central differences, is applied. Due to $K_\rho = K_\rho(z,t)$ the resulting matrix, denoted by B_j^{diff} depends on j and is non-symmetric (cf., Hackbusch, 2010; Section 5). It is tridiagonal, and the system is solved directly by splitting it up into four blocks. Note that $A_j^{\text{diff}}, A^{\text{sink}}$ are 4×4 block-diagonal matrices.

Thus, the operator splitting method can be summarised by the following operator equation:

$$\underbrace{\left[I - \tau A_j^{\text{diff}}\right]}_{:=B_j^{\text{diff}}} \mathbf{y}_{j+1} = \underbrace{\left[I + \tau A^{\text{sink}}\right]}_{:=B^{\text{sink}}} B_j^Q \circ B_j^Q \circ B_j^Q \circ B_j^Q(\mathbf{y}_j),\ j = 1, \ldots, M. \qquad (4)$$

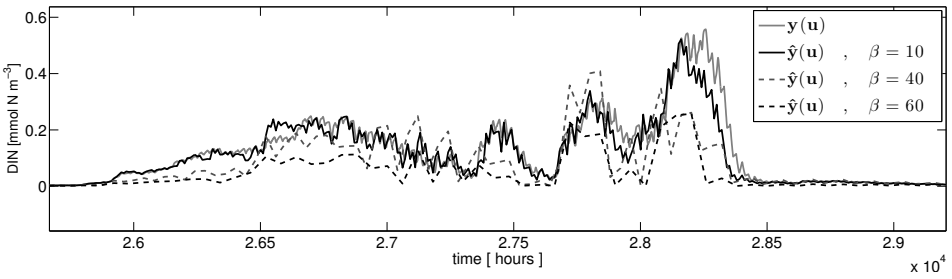
In the original discretisation given by (4) the time step τ is chosen as one hour. In the following this is referred to as the high-fidelity or fine model.

5.2 Coarser time discretisation

The coarse model is obtained by using a coarser time discretisation with $\hat{\tau} = \beta\tau$ applying a *coarsening factor* $\beta \in \mathbb{N} \setminus \{0, 1\}$, while keeping the spatial discretisation fixed. The state variable for this coarser discretised model will be denoted by $\hat{\mathbf{y}}$, the corresponding number of discrete time steps by $\hat{M} = M/\beta$, i.e., $\hat{\mathbf{y}} \in \mathbb{R}^{4K\hat{M}}$. Note that in our case number and type of the parameters $\mathbf{u}$ for the coarse and fine model are equal.

Figure 2 shows the fine and coarse model output $\mathbf{y}$ and $\hat{\mathbf{y}}$, respectively, for the state detritus, for different values of β and at the same randomly chosen parameter vector $\mathbf{u}$. We point out that figures presented in this paper show selected (representative) state variables/tracers for a part of the whole time interval at some distinct depth layers only. The total number of depth layers considered in the optimisation process is 33 and the entire time scale for the fine model is 43,800 so that it is impossible to present a full model output here. We emphasise that the qualitative behaviour of the other tracers and at different times and spatial layers is similar.

Figure 2 Fine and coarse 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 $\mathbf{u}$



Note: For simplicity we skip subscripts in the legends of all figures.

It is important to keep in mind that choosing β too big could lead to a numerically unstable scheme (cf., Fletcher, 1991). The condition on stability is determined by the ratio h/u_{12} (where h denotes the spatial step-size) and properties of the non-linear coupling term Q_j which we do not discuss here. All computations in this paper were performed with parameters that guarantee stability.

6 The surrogate model

The ASM algorithm, as will be described in the next section, is a conditionally equivalent approach to use a surrogate model in the optimisation which is obtained by a SM approach introduced by Bandler et al. (1994). Here a a physical low-fidelity or

coarse model with output $\hat{\mathbf{y}}$ (cf., Section 5) is corrected in the kth optimisation step by a so-called parameter mapping $\mathbf{p}_k$ to obtain a surrogate $\mathbf{s}_k$ for the fine model, in detail

$$\begin{aligned} &\mathbf{s}_k(\mathbf{u}) := \hat{\mathbf{y}}\left(\mathbf{p}_k(\mathbf{u})\right), \quad \mathbf{p}_k(\mathbf{u}) = \mathbf{p}(\mathbf{u}_k) + \mathbf{p}'(\mathbf{u}_k)\left(\mathbf{u} - \mathbf{u}_k\right), \\ &\hat{\mathbf{u}}_k = \mathbf{p}(\mathbf{u}_k) := \operatorname{argmin}_{\mathbf{u}\in U} \left|\left| \hat{\mathbf{y}}(\mathbf{u}) - \mathbf{y}(\mathbf{u}_k) \right|\right|_Y^2 . \end{aligned} \tag{5}$$

The usually non-linear mapping $\mathbf{p}$ is aligning the fine and coarse model and is approximated in the point $\mathbf{u}_k$ using a first-order Taylor expansion.

6.1 0-order consistency

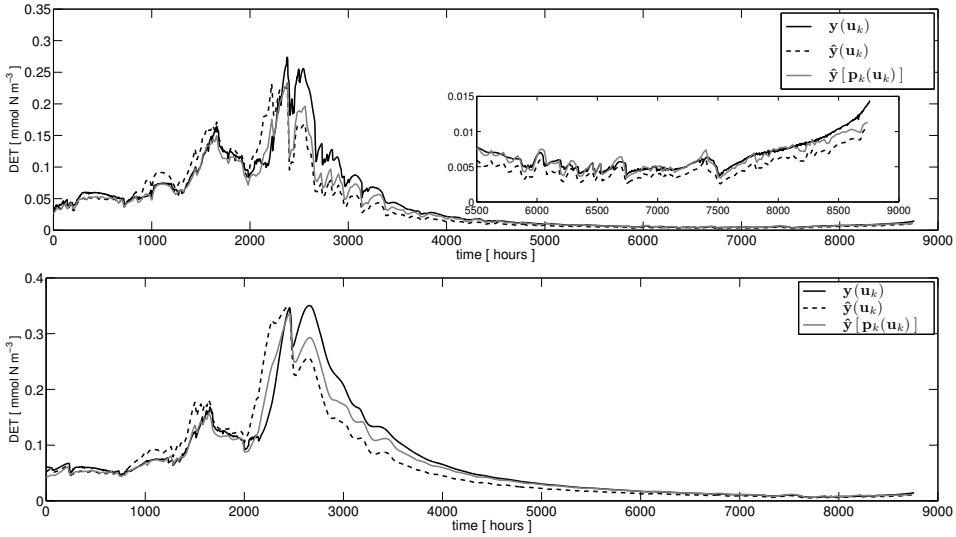
Assuming that the minimisation in (5) actually yields perfect alignment

$$\hat{\mathbf{y}}(\hat{\mathbf{u}}_k) = \mathbf{y}(\mathbf{u}_k),$$

the surrogate exactly satisfies 0-order consistency, i.e., $\mathbf{s}_k(\mathbf{u}_k) = \mathbf{y}(\mathbf{u}_k)$ (cf., Section 4).

If this is not the case, i.e., the minimisation (5) yields a local minimum for which we would have obtained an approximate alignment only, i.e., $\hat{\mathbf{y}}(\hat{\mathbf{u}}_k) \approx \mathbf{y}(\mathbf{u}_k)$ then obviously the surrogate's consistency is only satisfied approximately, i.e., $\mathbf{s}_k(\mathbf{u}_k) \approx \mathbf{y}(\mathbf{u}_k)$.

Figure 3 Fine and coarse model output $\mathbf{y}, \hat{\mathbf{y}}$ as well as the aligned surrogate $\mathbf{s}_k(\mathbf{u}_k) = \hat{\mathbf{y}}\left(\mathbf{p}_k(\mathbf{u}_k)\right)$ for the state detritus, at the same randomly chosen parameter vector $\mathbf{u}_k$, at depths $z \approx 25$ m (top) and $z \approx 60$ m



Note: The surrogate model provides a reasonable approximation of the fine model while lying closer than the coarse model itself.

The 0-order consistency is dependent on how close the alignment of the coarse model can be achieved by $\mathbf{p}$. However, using the definition of the surrogate and the mapping from (5), the surrogate obviously is at least as close to the fine model as the coarse model itself, i.e.,

$$\left\|\mathbf{s}_k(\mathbf{u}_k) - \mathbf{y}(\mathbf{u}_k)\right\| = \left\|\hat{\mathbf{y}}\left[\mathbf{p}(\mathbf{u}_k)\right] - \mathbf{y}(\mathbf{u}_k)\right\| \leq \left\|\hat{\mathbf{y}}(\mathbf{u}_k) - \mathbf{y}(\mathbf{u}_k)\right\| \tag{6}$$

where the second relation is ensured by the minimisation (5). Figure 3 illustrates this property showing the fine and coarse as well as the surrogate model output for the state detritus at a randomly chosen parameter vector $\mathbf{u}_k$. This supports the argumentation above: In the point $\mathbf{u}_k$ the surrogate obviously provides a reasonable approximation for the fine model while being closer to it than the coarse model itself. We will see in the next section that this property is also given in a neighbourhood.

7 Aggressive space mapping

In this section, we will briefly recall the basic idea of the ASM algorithm and present the globalisation strategy as well as the pseudo code of the algorithm we used to obtain the results presented in this paper. The ASM algorithm was firstly developed by Bandler et al. (1994). It firstly solves for an optimum of the coarse model, i.e.,

$$\hat{\mathbf{u}}^* := \operatorname*{argmin}_{\mathbf{u}\in U} J(\hat{\mathbf{y}}(\mathbf{u}))$$

and then iteratively computes a solution $\bar{\mathbf{u}}$ of the non-linear system

$$\mathbf{F}(\bar{\mathbf{u}}) := \mathbf{p}(\bar{\mathbf{u}}) - \hat{\mathbf{u}}^* = 0. \tag{7}$$

using a Quasi-Newton iteration (Kosmol, 1993; Nocedal and Wright, 2000) with a Broyden rank-one approximation (Broyden, 1965) for the Jacobian $B_k \approx \mathbf{p}'(\mathbf{u}_k)$ (see also Bandler et al., 1994, 2004a).

The following results were shown in Echeverrá and Hemker (2005): if either the fine model nearly matches the data in an optimum

$$\mathbf{u}^* := \operatorname*{argmin}_{\mathbf{u}\in U} J(\mathbf{y}(\mathbf{u})), \quad \text{i.e., } \mathbf{y}(\mathbf{u}^*) \approx \mathbf{y}_d,$$

or if both models are similar near their respective optima ($\mathbf{y}(\mathbf{u}^*) \approx \hat{\mathbf{y}}(\hat{\mathbf{u}}^*)$), we obtain [using (5)]

$$\mathbf{p}(\mathbf{u}^*) = \operatorname*{argmin}_{\mathbf{u}\in U} ||\hat{\mathbf{y}}(\mathbf{u}) - \mathbf{y}(\mathbf{u}^*)||_Y^2 \;\approx\; \operatorname*{argmin}_{\mathbf{u}\in U} ||\hat{\mathbf{y}}(\mathbf{u}) - \mathbf{y}_d||_Y^2 = \hat{\mathbf{u}}^*. \tag{8}$$

which is also referred to as a *perfect mapping* and which motivates to solve for a solution of (7).

If in addition to (8) the mapping is *injective* and the coarse model optimum $\hat{\mathbf{u}}^*$ is *unique*, then the solution of the ASM approach, $\bar{\mathbf{u}}$, coincides with the fine model optimum $\mathbf{u}^*$ and the solution $\bar{\mathbf{u}}_s$ obtained by directly optimising the surrogate defined in (5), i.e.,

$$\bar{\mathbf{u}}_s = \operatorname*{argmin}_{\mathbf{u}\in U} J(\hat{\mathbf{y}}(\mathbf{p}(\mathbf{u}))). \tag{9}$$

However, in most real applications these theoretically derived conditions might of course not be exactly satisfied. For a more detailed analysis we also refer to Echeverrá and Hemker (2005).

For the complex model used here, it is not the focus of this paper (and it is not clear if it is possible) to prove those theoretical conditions. Instead, the applicability of the ASM algorithm is verified by using synthetic target data $\mathbf{y}_d = \mathbf{y}(\mathbf{u}_d)$ with known parameters $\mathbf{u}_d$ and by comparing the ASM solution $\bar{\mathbf{u}}$ to those obtained by fine and coarse model optimisation, $\mathbf{u}^*$ and $\hat{\mathbf{u}}^*$, as well as to the known optimal parameters $\mathbf{u}_d$.

7.1 Globalised Quasi-Newton method

Since the standard Quasi-Newton algorithm, as given in, e.g., Kosmol (1993) and Nocedal and Wright (2000), may suffer from local convergence one can additional use a classical line search strategy introducing a merit function $h : U \to \mathbb{R}$ given as (Kosmol, 1993)

$$h(\mathbf{u}) := \frac{1}{2} \,||\, \mathbf{F}(\mathbf{u}) \,||^2 = \frac{1}{2} \,||\, \mathbf{p}(\mathbf{u}) - \hat{\mathbf{u}}^* \,||^2 .$$

If $\mathbf{F}'(\mathbf{u}_k)B_k^{-1}$ is positive-definite, then

$$\nabla h(\mathbf{u}_k)^\top \mathbf{d}_k = \mathbf{F}(\mathbf{u}_k)^\top \mathbf{F}'(\mathbf{u}_k) B_k^{-1} \mathbf{F}(\mathbf{u}_k) \leq 0,$$

i.e., $\mathbf{d}_k$ is a descent direction for h at the point $\mathbf{u}_k$.

Obviously the Newton direction [where B_k is replaced by $\mathbf{F}'(\mathbf{u}_k)$] is always a descent direction for h in $\mathbf{u}_k$, satisfying $\nabla h(\mathbf{u}_k)^\top \mathbf{d}_k = -2h(\mathbf{u}_k)$. Assuming that B_k is a "good" approximation of $\mathbf{F}'(\mathbf{u}_k)$, we use the last relation also in a line search in the Quasi-Newton method. The iteration step in the globalised Quasi-Newton algorithm then takes the following form:

$$\left.\begin{aligned} \text{Find } \sigma \in]0,1[,\ \text{s.t. } h(\mathbf{u}_k + \sigma \mathbf{d}_k) &\leq (1 - 2\sigma\delta)\, h(\mathbf{u}_k) \\ &\approx h(\mathbf{u}_k) + \sigma\delta \left(\nabla h(\mathbf{u}_k)^\top \mathbf{d}_k\right), \\ \text{Update} \quad \mathbf{u}_{k+1} &= \mathbf{u}_k + \sigma \mathbf{d}_k, \\ B_{k+1} &= B_k + \tfrac{(\mathbf{y}_k - \sigma B_k \mathbf{d}_k)\sigma \mathbf{d}_k^\top}{\sigma^2 \mathbf{d}_k^\top \mathbf{d}_k}. \end{aligned}\right\} \tag{10}$$

Here $\delta \in]0,1[$ is a parameter that defines the rate of decrease in the merit function that is desired in the current step, similar as in Armijo's rule (e.g., Nocedal and Wright, 2000). Since we are using an approximation in (10) anyway, we may write it in a simpler form with $2\sigma\delta$ replaced by $C \in]0,1[$ (Kosmol, 1993). The resulting pseudo code can be found in Algorithm 1 at the end of this section.

In general, the Broyden update does *not* guarantee to provide a descent direction, and thus the line search might fail (cf., line 5 and 6 in Algorithm 1). In this case one could use the Newton or steepest descent direction (cf., lines 7 to 17 in Algorithm 1) and apply one of them directly to the merit function, following the idea that minimising h will lead closer to a zero of $\mathbf{F}$.

7.2 Practical issues

For a given optimisation problem one has to carefully consider how many iterations and hence evaluations of the function $\mathbf{F}$ are affordable. The alignment of the coarse model through the parameter mapping $\mathbf{p}$ and hence the evaluation of the function $\mathbf{F}$ is quite expensive due to the minimisation required to obtain $\mathbf{p}(\mathbf{u}_k)$. One evaluation of the function $\mathbf{p}$ requires one expensive evaluation of the fine model plus the minimisation which – depending on the chosen method – might additionally include derivatives or their approximations by more function evaluations of the coarse model.

Algorithm 1 Globalised Quasi-Newton algorithm to solve for the root of $\mathbf{F}(\mathbf{u})$

function $[\ \mathbf{u}_k\ ] =$ glob_quasi_newton ▷ Main programme
  $B_0 = I, \mathbf{u}_0 = \hat{\mathbf{u}}^*, k = 0, \epsilon_h = 1e-8, \epsilon_{LS} = 1.25e-1, \delta = 1e-4, k_{max} = 10$
  $\mathbf{F}_0 = \mathbf{F}(\mathbf{u}_0) \quad , \quad h_0 = h(\mathbf{u}_0)$ ▷ Initialisation
  while $h_k \geq \epsilon_h$ **and** $k < k_{max}$ **do**
    $B_k\, \mathbf{d}_k = -\mathbf{F}_k \quad , \quad p_k = 2 \cdot \delta \cdot h_k$ (**or** $p_k = C$) ▷ Calculate Quasi-Newton step
    $[\ \mathbf{F}_{trial}, h_{trial}, \mathbf{u}_{trial}, flag\ ] =$ line search($\mathbf{u}_k, \mathbf{d}_k, h_k, p_k$)
    if $flag = 0$ **then** ▷ If line search failed ...
      Error('Quasi-Newton step failed: Step length σ below threshold ϵ_{LS}. Broyden matrix might be inaccurate ($B_k \neq \mathbf{F}'_k$) or local minimum of h reached.')
      $B_k = \mathbf{F}'_k$ ▷ Approximate Jacobian
      if $\mathbf{F}'_k$ regular **then**
        $B_k\, \mathbf{d}_k = -\mathbf{F}_k \quad , \quad p_k = 2 \cdot \delta \cdot h_k$ ▷ Newton step
      else
        $\mathbf{d}_k = -B_k^\top\, \mathbf{F}_k \quad , \quad p_k = \delta \cdot ||\, \mathbf{d}_k\, ||^2$ ▷ Steepest Descent step
      end if
      $[\ \mathbf{F}_{trial}, h_{trial}, \mathbf{u}_{trial}, flag\ ] =$ line search($\mathbf{u}_k, \mathbf{d}_k, h_k, p_k$)
    end if
    if $flag = 1$ **then** ▷ Step successful
      $\mathbf{F}_{k+1} = \mathbf{F}_{trial} \quad , \quad h_{k+1} = h_{trial} \quad , \quad \mathbf{u}_{k+1} = \mathbf{u}_{trial} \quad , \quad \mathbf{d}_k = \mathbf{u}_{k+1} - \mathbf{u}_k$
      $B_{k+1} = B_k + \frac{(\mathbf{y}_k - B_k\, \mathbf{d}_k)\, \mathbf{d}_k^\top}{\mathbf{d}_k^\top\, \mathbf{d}_k}$
      $k = k + 1$
    else
      Error('Also Newton/Steepest Descent step failed: Step length below threshold ϵ_{LS}.Local minimum assumed.')
      $k = k_{max}$
    end if
  end while
end function
function $[\ \mathbf{F}_{trial}, h_{trial}, \mathbf{u}_{trial}, flag\ ] =$ line search($\mathbf{u}, \mathbf{d}, \mathbf{h}, p_k$) ▷ line search procedure
  $\alpha = 0 \quad , \quad \sigma = \frac{1}{2}^\alpha \quad , \quad flag = 0$ ▷ Initialisation
  while ($\sigma > \epsilon_{LS}$ **and** $flag = 0$) **do**
    $\mathbf{u}_{trial} = \mathbf{u} + \sigma \cdot \mathbf{d} \quad , \quad \mathbf{F}_{trial} = \mathbf{F}(\mathbf{u}_{trial}) \quad , \quad h_{trial} = h(\mathbf{u}_{trial})$
    if $h_{trial} \leq h - \sigma \cdot p_k$ **then**
      $flag = 1$ ▷ $flag$ indicates success/ failure of line search
    else
      $\alpha = \alpha + 1 \quad , \quad \sigma = \frac{1}{2}^\alpha$
    end if
  end while
end function

Definitions:

$\mathbf{F}_k := \mathbf{F}(\mathbf{u}_k) \quad , \quad h_k := h(\mathbf{u}_k) \quad , \quad \mathbf{y}_k := \mathbf{F}_{k+1} - \mathbf{F}_k$

In order to keep the number of fine model evaluations as low as possible we did not use optional and expensive Newton or steepest descent steps but only the cheaper Quasi-Newton direction. For the results provided in the next section we furthermore used a line search with three iterations at maximum (corresponding to a minimal step length of $||\,\mathbf{d}\,|| = 1.25e-1$). If the line search failed to find a suitable step length, the algorithm was terminated. In this case it either got stuck in a local minimum of h or the Broyden matrix provided an inaccurate approximation resulting in a non-descent direction.

Note that a successful line search in the ASM algorithm also indirectly confirms that the Taylor approximation we use to approximate the mapping $\mathbf{p}$ [cf., (5)] is reasonable in the current point $\mathbf{u}_k$ and also in some neighbourhood $\mathbf{u}_k + \epsilon$, i.e., $\mathbf{p}_k(\mathbf{u}_k + \epsilon) \approx \mathbf{p}(\mathbf{u}_k + \epsilon)$ for sufficiently small perturbation vectors ϵ. Hence the surrogate $\mathbf{s}_k = \hat{\mathbf{y}}\,(\mathbf{p}_k(\mathbf{u}))$ will be a reasonable approximation also in a neighbourhood around $\mathbf{u}_k$. We yield [using (6)]

$$||\,\hat{\mathbf{y}}\,(\mathbf{p}_k(\mathbf{u}_k + \epsilon)) - \mathbf{y}(\mathbf{u}_k + \epsilon)\,||^2 \leq ||\,\hat{\mathbf{y}}(\mathbf{u}_k + \epsilon) - \mathbf{y}(\mathbf{u}_k + \epsilon)\,||^2 .$$

This is also interesting when directly optimising the surrogate [cf., (9)].

8 Results and discussion

We verified the ASM approach by using synthetic target data $\mathbf{y}_d$ at a randomly chosen parameter vector $\mathbf{u}_d$ within admissible bounds $\mathbf{b}_l, \mathbf{b}_u$:

$$\mathbf{y}_d := \mathbf{y}(\mathbf{u}_d), \quad \mathbf{y}_d \in \mathbb{R}^{4KM}, \quad \mathbf{b}_l \leq \mathbf{u}_d \leq \mathbf{b}_u. \tag{11}$$

We used the globalised Quasi-Newton method described in the last section. For the optimisation of the coarse and fine model and for the minimisation required to obtain the parameter mapping $\mathbf{p}$ [cf., (5)] we used the MATLAB[1] function `fmincon`, taking the option for the active-set algorithm.

8.1 Cost function

The cost function used for the fine model optimisation (using a weighted Euclidean vector norm) is given as follows

$$J\,(\,\mathbf{y}\,) = \frac{1}{4KM} \sum_{i=1}^{4K} \sum_{j=1}^{M} \left(y_{ij} - (y_d)_{ij}\right)^2 . \tag{12}$$

Since the coarse model output contains less data points ($\hat{\mathbf{y}} \in \mathbb{R}^{4K\hat{M}}$, cf., Section 5) we accordingly use a sum over $\hat{M}$ discrete time steps and have to consider the sampled target data $(y_d)_{i,\beta j}$ in order to obtain comparable data points.

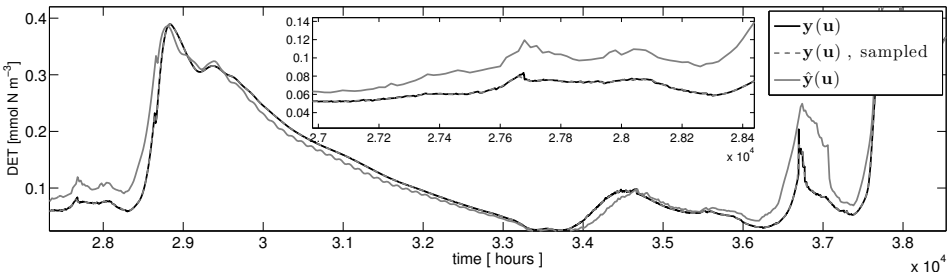
8.2 Parameter mapping

For obtaining the parameter mapping $\mathbf{p}$ we analogously have to consider the sampled fine model output $y_{i,\beta j}(\mathbf{u}_k)$, i.e.,

$$\mathbf{p}(\mathbf{u}_k) = \operatorname*{argmin}_{\mathbf{u}\in U} \left[\frac{1}{4K\hat{M}} \sum_{i=1}^{4K} \sum_{j=1}^{\hat{M}} \left(\hat{y}_{ij}(\mathbf{u}) - y_{i,\beta j}(\mathbf{u}_k)\right)^2 \right]. \tag{13}$$

Figure 4 shows a comparison of the full and sampled fine model and the coarse model output in the 10th vertical level with a sampling rate and coarsening factor of $\beta = 20$ (as in the following results). Obviously, we are still able to capture the main characteristics of the fine model output when using sampling, since the curves for the original and sampled output are very close.

Figure 4 Comparison of the full (i.e., not sampled) fine model output $\mathbf{y}$ for the state detritus, the corresponding sampled output and the coarse model output $\hat{\mathbf{y}}$ at the same randomly chosen parameter vector $\mathbf{u}$, in the 10th vertical layer, using a coarsening factor of β = 20



Note: Curves for original and sampled fine model output are very close.

8.3 ASM vs. direct fine model optimisation

We now compare the results and the computational cost of the ASM approach with the direct optimisation of the fine model using the cost function given in (12). To give a profound illustration of the behaviour of the algorithm, we below consider the following parameter values, model outputs, and respective cost function values:

(1) the target $\mathbf{y}_d$, i.e., the fine model output at a randomly chosen parameter vector $\mathbf{u}_d$

(2) the fine model output at another randomly chosen parameter vector $\mathbf{u}_0$, serving as initial value of the optimisation runs

(3) the model output at the result $\hat{\mathbf{u}}^*$ of a pure coarse model optimisation

(4) the fine model output at the result $\bar{\mathbf{u}}$ of the ASM algorithm

(5) the output of a (rather expensive) fine model optimisation yielding $\mathbf{u}^*$.

Using different optimisation routines might yield different results in (3) to (5), but this will probably not influence the relative reduction in the total optimisation cost using the

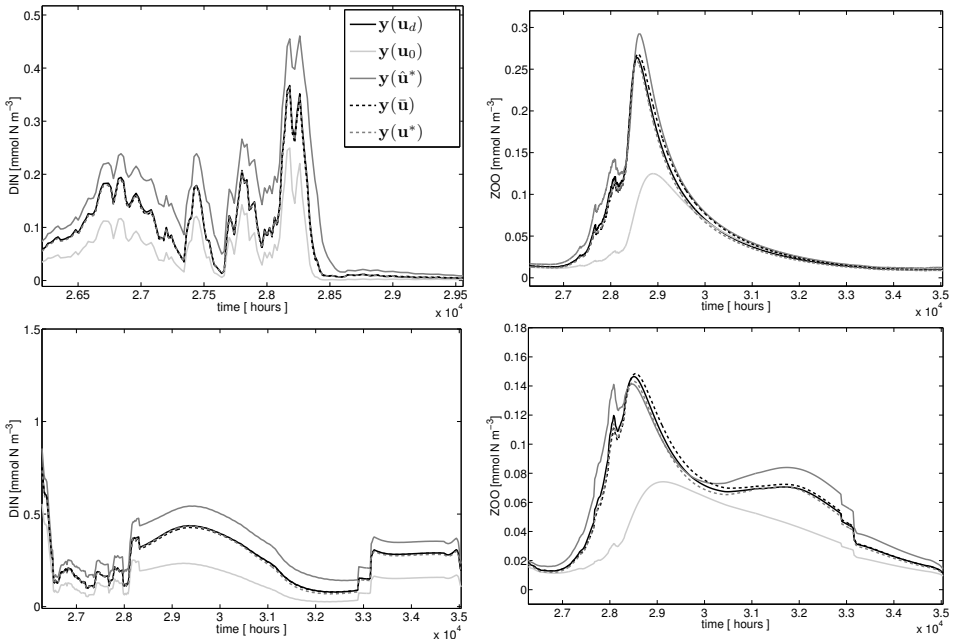
ASM algorithm (see below). For example, in Rückelt et al. (2010) better cost function values were obtained by direct fine model optimisation using a different optimisation method (other than MATLAB's `fmincon`) for the same problem and the same model.

A good agreement between the results of (1) and (5) would indicate a high quality of the used optimisation method itself, whereas a good agreement between those from (4) and (5) would mean that the ASM provides a reasonable solution close to the fine model optimum.

We experienced that results for different initial parameter vectors $\mathbf{u}_0$ are comparable. For illustration we here present the results of two exemplary test runs, considering the same target dataset $\mathbf{y}_d$, but different initial parameters $\mathbf{u}_0$. In both cases we used $K = 20$, $M = 8{,}760 \cdot 5$ and $\beta = 20$. Hence for the coarse model we obtain $\hat{M} = M/\beta = 2{,}190$ discrete time steps per spatial grid point z_i.

In Figures 5 and 6, the output of the fine model was sampled as described above. We again point out that the qualitative behaviour of the other tracers and of the shown tracers at different times and spatial layers – for simplicity not presented in this figure – is similar. Table 1 shows the resulting values of the parameters, the cost function and the gained reduction in the computational cost.

Figure 5 Fine model output $\mathbf{y}$ for dissolved inorganic nitrogen (left) and for zooplankton (right) at depth $z \approx 2.68$ m (top) and $z \approx 108.15$ m (bottom)

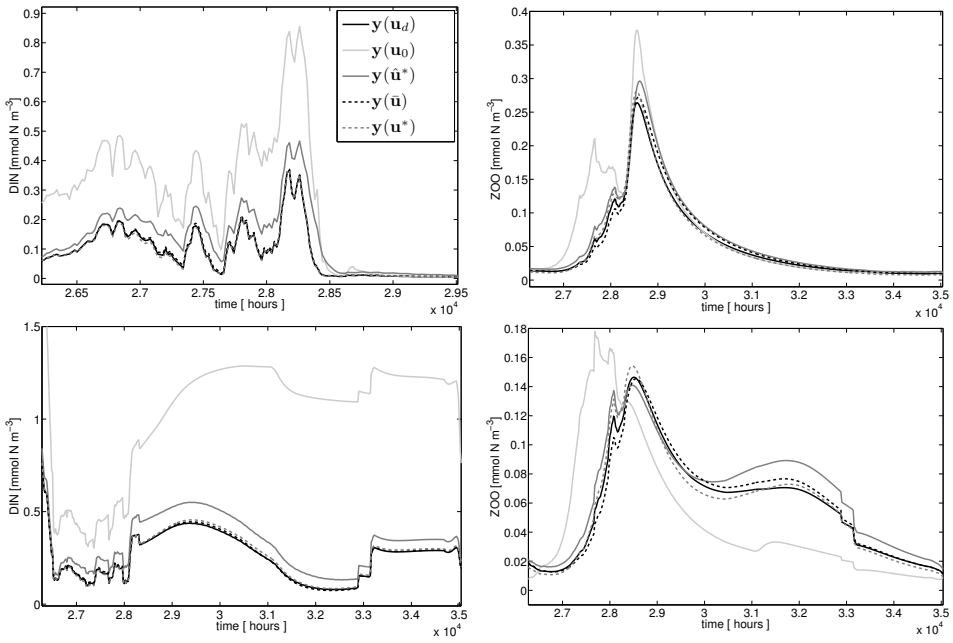


Notes: Shown are, in the legend from top to bottom:

(1) target $\mathbf{y}_d$, i.e., fine model output at randomly chosen parameters $\mathbf{u}_d$
(2) fine model output at the initial value $\mathbf{u}_0$
(3) at the coarse model optimum $\hat{\mathbf{u}}^*$
(4) at the result of the ASM algorithm $\bar{\mathbf{u}}$
(5) at the result of the direct fine model optimisation yielding $\mathbf{u}^*$.

On the top left, we only show the interesting time interval. Curves corresponding to (1), (4) and (5) are very close.

Figure 6 Same as Figure 5, but for a different initial guess $\mathbf{u}_0$ for the optimisation runs



Notes: Again the curves corresponding to (1), (4) and (5) are very close. See the text for further details.

8.4 Test results for model output and computational cost

Figure 5 illustrates the results of the ASM approach and of a typical fine and coarse model optimisation for a first choice of the initial guess $\mathbf{u}_0$. Corresponding parameters and values of the cost function J are given in the upper part of Table 1. Furthermore the table shows the total cost of the fine ($C_{opt,h}$) and the coarse model optimisation ($C_{opt,l}$) and of the Quasi-Newton iterations of the ASM algorithm (C_{QN}) in terms of the total number of equivalent fine model evaluations, which were required to reach the given value of the cost function J. Equivalent in this case means that for example β evaluations of the coarse model used here with a coarsening factor β are equivalent to (or, as expensive as) one fine model evaluation. Note that the total cost in the ASM approach consists of the cost for the coarse model optimisation $C_{opt,l}$ and those for solving the non-linear system of equations by the Quasi-Newton method, i.e., C_{QN}. For details see also the next subsection.

From Figure 5, we see that by the direct fine model optimisation we yield a very reasonable optimal fit $\mathbf{y}(\mathbf{u}^*)$ (grey dashed line) of the target data $\mathbf{y}_d$ (black line). This corresponds to a cost function value of $J(\mathbf{y}(\mathbf{u}^*)) = 1.611e-05$ obtained after 281 function evaluations (cf., Table 1). We furthermore see that by the coarse model optimisation we yield parameters $\hat{\mathbf{u}}^*$ with a fit $\mathbf{y}(\hat{\mathbf{u}}^*)$ (light grey line) which obviously provides only a rough approximation of the target data, but in $C_{opt,l} = 19.95$ equivalent fine model evaluations only. Using the ASM approach, we finally obtain a solution $\bar{\mathbf{u}}$ with an optimal fit $\mathbf{y}(\hat{\mathbf{u}}^*)$ (black dashed line) and parameter match lying very close to that obtained by the fine model optimisation.

Table 1 Results of the fine and coarse model optimisation and of the ASM algorithm from two illustrative test runs, corresponding to Figures 5 (top) and 6 (bottom), see the text for details

	$u_{k,1}$	$u_{k,2}$	...									$u_{k,12}$	J	C_i
$\mathbf{u}_0$	0.486	0.644	0.019	0.01	0.037	0.933	1.905	0.006	0.18	0.017	0.406	6.937	5.9e-03	
	Fine model optimisation:					$\mathbf{u}^* := \text{argmin}_{\mathbf{u}\in U} J(\mathbf{y}(\mathbf{u}))$								
$\mathbf{u}^*$	0.764	0.599	0.027	0.01	0.035	1.018	1.93	0.01	0.218	0.02	0.495	5.866	1.6e-05	281
	Coarse model optimisation:					$\hat{\mathbf{u}}^* := \text{argmin}_{\mathbf{u}\in U} J(\hat{\mathbf{y}}(\mathbf{u}))$								
$\hat{\mathbf{u}}^*$	0.759	0.363	0.025	0.012	0.029	1.118	0.864	0.007	0.194	0.016	0.491	5.42	1.8e-03	19.95
	ASM:					Solve $\mathbf{F}(\bar{\mathbf{u}}) := \mathbf{p}(\bar{\mathbf{u}}) - \hat{\mathbf{u}}^* = 0$								
$\bar{\mathbf{u}}$	0.759	0.587	0.027	0.011	0.034	0.944	1.524	0.01	0.179	0.02	0.49	6.073	5.0e-05	80.25
$\mathbf{u}_d$	0.75	0.6	0.025	0.01	0.03	1.0	2.0	0.01	0.205	0.02	0.5	6.0	57.54% Reduction	
	$u_{k,1}$	$u_{k,2}$	...									$u_{k,12}$	J	C_i
$\mathbf{u}_0$	0.565	0.672	0.015	0.012	0.036	1.096	2.335	0.013	0.209	0.028	0.452	5.235	7.0e-02	
	Fine model optimisation:					$\mathbf{u}^* := \text{argmin}_{\mathbf{u}\in U} J(\mathbf{y}(\mathbf{u}))$								
$\mathbf{u}^*$	0.871	0.593	0.029	0.012	0.038	1.0478	0.952	0.011	0.223	0.019	0.466	5.836	5.6e-05	418
	Coarse model optimisation:					$\hat{\mathbf{u}}^* := \text{argmin}_{\mathbf{u}\in U} J(\hat{\mathbf{y}}(\mathbf{u}))$								
$\hat{\mathbf{u}}^*$	0.759	0.356	0.029	0.012	0.037	1.138	0.848	0.007	0.188	0.016	0.502	5.475	1.8e-03	26.35
	ASM:					Solve $\mathbf{F}(\bar{\mathbf{u}}) := \mathbf{p}(\bar{\mathbf{u}}) - \hat{\mathbf{u}}^* = 0$								
$\bar{\mathbf{u}}$	0.761	0.572	0.031	0.011	0.043	0.96	1.529	0.011	0.174	0.02	0.512	5.976	5.9e-05	91.15
$\mathbf{u}_d$	0.75	0.6	0.025	0.01	0.03	1.0	2.0	0.01	0.205	0.02	0.5	6.0	71.27% Reduction	

Notes: Also shown are the corresponding values of the cost function J given in (12) and the computational cost C_i in terms of the total number of equivalent fine model evaluations required to obtain the given cost function value J, again for the three cases, i.e., $C_i \in \{C_{opt,h}, C_{opt,l}, C_{QN}\}$.

The key point now is that the ASM solution $\bar{\mathbf{u}}$ is obtained in only C_{QN} = 80.25 equivalent fine model evaluations for the Quasi-Newton steps. Summarising, using the ASM approach, we hence obtained a very reasonable solution in totally

$$C_{ASM} := C_{opt,l} + C_{QN} \approx 100$$

equivalent fine model evaluations. The same cost function value by the direct fine model optimisation was obtained after 236 model evaluations (cf., Table 1). This leads to a reduction in the total optimisation cost of about 57%.

Figure 6 and the lower part of Table 1 show the corresponding results of another run with different initial parameters $\mathbf{u}_0$. Obviously results of the fine and coarse model optimisation look similar. Also a similar convenient fit was obtained by the ASM approach. Here we required about 117 equivalent fine model evaluations to obtain

the ASM solution while about 409 evaluations were necessary to yield the same cost function value by the fine model optimisation. This leads to an even better reduction in the total optimisation cost of about 71%.

8.5 Analysis of the optimisation cost

In order to compare the total optimisation cost of the ASM approach, in the following denoted by C_{ASM}, with the one obtained from the fine and coarse model optimisation ($C_{opt,h}$, $C_{opt,l}$) we consider the cost in terms of total number of equivalent fine model evaluations. We generally yield the following:

$$\begin{aligned} C_{ASM} &:= C_{opt,l} + C_{QN}, \quad C_{QN} := N_{ASM} \cdot C_p \cdot N_{LS}^{qn}, \\ C_p \quad &:= C_{align} + 1, \quad C_{align} := N_{opt,p} \cdot (C_{grad} + N_{LS}^{opt})/\beta, \\ C_{opt,l} &:= N_{opt,l} \cdot (C_{grad} + N_{LS}^{opt})/\beta, \\ C_{opt,h} &:= N_{opt,h} \cdot (C_{grad} + N_{LS}^{opt}), \quad C_{grad} = 12. \end{aligned} \tag{14}$$

The optimisation cost for the fine and coarse model optimisation is given as the number of iterations, denoted by $N_{opt,h}, N_{opt,l}$, times the cost of the gradient C_{grad} plus the number of line search steps done per iteration, denoted by N_{LS}^{opt}. Note that for simplicity we assume the value N_{LS}^{opt} to be equal for the different optimisation runs. To obtain the cost of the coarse model optimisation $C_{opt,l}$ we furthermore have to divide by the coarsening factor β to yield the number of equivalent fine model evaluations as descibed above. For the cost for a gradient evaluation we use the number of optimisation variables (i.e., 12 here) which corresponds to the usual effort for a finite difference approximation and that for a forward mode AD (algorithmic/automatic differentiation) gradient.

As we described in Section 7, the ASM algorithm involves firstly to solve for the coarse model optimum $\hat{\mathbf{u}}^*$ resulting in $C_{opt,l}$ equivalent fine model evaluations. The second part within the ASM algorithm involves the (globalised) Quasi-Newton iteration (cf., Section 7) which results in C_{QN} equivalent fine model evaluations. C_{QN} is given as N_{ASM}, the overall number of steps in the ASM algorithm (i.e., the number of Quasi-Newton iterations), times C_p, the cost of calculating the mapping $\mathbf{p}$, times N_{LS}^{qn}, the number of line search steps. The cost of the mapping, C_p, is furthermore given by one fine model evaluation plus the cost of the minimisation (alignment) required for the mapping [cf., (5) and (13)], which we denote by C_{align}. Since the alignment is an optimisation in the coarse model space, its cost C_{align} is given similar to the cost for the coarse model optimisation ($C_{opt,l}$) now with $N_{opt,p}$ denoting the corresponding number of iterations in this optimisation process.

In the numerical tests we obtained the following: at average $\bar{C}_{opt,h} \approx 230$, $\bar{C}_{opt,l} \approx 1.4 \cdot \bar{C}_{opt,h}/\beta$, $\bar{C}_{align} \approx \bar{C}_{opt,l}$, $\bar{N}_{ASM} \approx 3.125$, $\bar{N}_{LS}^{qn} \approx 1.2$, i.e.,

$$\begin{aligned} \bar{C}_{ASM} \quad &\approx 1.4 \cdot \bar{C}_{opt,h}/\beta + 3.125 \cdot (1.4 \cdot \bar{C}_{opt,h}/\beta + 1) \cdot 1.2, \\ \text{i.e., } \bar{C}_{ASM}/\bar{C}_{opt,h} &\approx 1.4/\beta + 3.125 \cdot (1.4/\beta + 1/\bar{C}_{opt,h}) \cdot 1.2. \end{aligned} \tag{15}$$

Applying the ASM algorithm, using a coarsening factor of $\beta = 20$ for the coarse model and $\bar{C}_{opt,h} \approx 230$ as given above we hence yield an average reduction in the total optimisation cost of $(1 - \bar{C}_{ASM}/\bar{C}_{opt,h}) \cdot 100 \approx 65\%$.

The aim of the calculation above was to illustrate how the different parts of the ASM algorithm contribute to the total cost C_{ASM}. In principle, values for N_{ASM} and N_{LS}^{qn} are up to the user refinement and will be problem specific. Allowing for a greater number of ASM steps N_{ASM} or number of line search steps N_{LS}^{qn} in the globalised Quasi-Newton algorithm increases the total cost C_{ASM} while not necessarily yielding a more accurate solution in the end. To improve the results of the ASM approach in terms of cost reduction one might also decrease the cost for the coarse model optimisation ($C_{opt,l}$) by terminating the optimisation run of the coarse model after a certain number of iterations ($N_{opt,l}$) as more iterations might not necessarily improve the quality of its solution. Concluding, the setting for the ASM algorithm has to be carefully chosen to yield a reasonable solution at a sufficient reduction in the total optimisation cost.

9 Conclusions

Parameter optimisation of models that couple ocean circulation and a marine ecosystem model can be very expensive in terms of model and gradient evaluations, especially for complex three-dimensional models. Hence methods reducing the total optimisation cost, such as surrogate-based optimisation techniques, are highly desirable. In this paper we successfully applied the ASM approach to parameter optimisation of a one-dimensional coupled marine ecosystem model (fine model). We used a coarser discretisation in time to create a reasonable coarse model. We showed that using the mapping definition, the ASM approach is based on, for correction of the coarse model one can yield a reasonable surrogate for the fine model. We recalled that the ASM approach is a conditionally equivalent approach to directly use this surrogate in the optimisation run, replacing the fine model under consideration. Furthermore, we used a globalised Quasi-Newton iteration to obtain the ASM solution. We verified our approach by using synthetic target data comparing the results to those of direct fine model optimisation. All in all we have shown that with the ASM approach we could yield a very reasonable solution within a few number of fine model evaluations only, resulting in a significant reduction in the total optimisation cost of about 65% at average.

Acknowledgements

The authors would particularly like to thank Andreas Oschlies of IFM Geomar, Kiel and Johannes Rückelt of Institute of Computer Science, Kiel.

References

Bakr, M.H., Bandler, J.W., Madsen, K., Rayas-Sánchez, J.E. and Søndergaard, J. (2000) 'Space mapping optimization of microwave circuits exploiting surrogate models', *IEEE Trans. Microwave Theory Tech.*, Vol. 48, No. 12, pp.2297–2306.

Bandler, J.W., Biernacki, R.M., Chen, S.H., Grobelny, P.A. and Hemmers, R.H. (1994) 'Space mapping technique for electromagnetic optimization', *IEEE Transactions on Microwave Theory and Techniques*, Vol. 42, No. 12, pp.2536–2544.

Bandler, J.W., Biernacki, R.M., Chen, S.H., Hemmers, R.H. and Madsen, K. (1995) 'Electromagnetic optimization exploiting aggressive space mapping', *Microwave Theory and Techniques, IEEE Transactions on*, Vol. 43, No. 12, pp.2874–2882.

Bandler, J.W., Cheng, Q.S., Dakroury, S.A., Mohamed, A.S., Bakr, M.H., Madsen, K. and Søndergaard, J. (2004a) 'Space mapping: the state of the art', *IEEE T. Microw. Theory*, Vol. 52, No. 1, pp.337–361.

Bandler, J.W., Cheng, Q.S., Hailu, D.M., Mohamed, A.S., Bakr, M.H., Madsen, K. and Pedersen, F. (2004b) 'Recent trends in space mapping technology', With referee.

Broyden, C.G. (1965) 'A class of methods for solving nonlinear simultaneous equations', *Mathematics of Computation*, Vol. 19, No. 92, pp.577–593.

Conn, A.R., Gould, N.I.M. and Toint, P.L. (2000) *Trust-Region Methods*, Society for Industrial and Applied Mathematics, Philadelphia, PA.

Echeverría, D. and Hemker, P. (2005) 'Space mapping and defect correction', *Computational Methods in Applied Mathematics*, Vol. 5, No. 2, pp.107–136.

Fennel, W. and Neumann, T. (2004) *Introduction to the Modelling of Marine Ecosystems*, Elsevier, Amsterdam, The Netherlands.

Fletcher, C.A.J. (1991) *Computational Techniques for Fluid Dynamics*, 2nd ed., Vol. 1, Springer, Berlin, Heidelberg, New York.

Forrester, A.I. and Keane, A.J. (2009) 'Recent advances in surrogate-based optimization', *Prog. Aerosp. Sci.*, Vol. 45, Nos. 1–3, pp.50–79.

Hackbusch, W. (2010) *Elliptic Differential Equations: Theory and Numerical Treatment*, Springer Series in Computational Mathematics, Springer Berlin.

Kosmol, P. (1993) *Methoden zur numerischen Behandlung nichtlinearer Gleichungen und Optimierungsaufgaben*, Teubner.

Koziel, S., Bandler, J. and Cheng, Q. (2010) 'Robust trust-region space-mapping algorithms for microwave design optimization', *IEEE T. Microw. Theory*, Vol. 58, No. 8, pp.2166–2174.

Koziel, S., Bandler, J.W. and Madsen, K. (2006) 'A space – mapping framework for engineering optimization – theory and implementation', *IEEE Transactions on Microwave Theory and Techniques*, Vol. 54, No. 10, pp.3721–3730.

Marchuk, G.I. (1982) *Methods of Numerical Mathematics*, 2nd ed., Springer, New York.

Nocedal, J. and Wright, S. (2000) *Numerical Optimization*, Springer, New York.

Oschlies, A. and Garcon, V. (1999) '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.

Queipo, N.V., Haftka, R.T., Shyy, W., Goel, T., Vaidyanathan, R. and Tucker, P.K. (2005) 'Surrogate-based analysis and optimization', *Prog. Aerosp. Sci.*, Vol. 41, No. 1, pp. 1–28.

Rückelt, J., Sauerland, V., Slawig, T., Srivastav, A., Ward, B. and Patvardhan, C. (2010) 'Parameter optimization and uncertainty analysis in a model of oceanic co2-uptake using a hybrid algorithm and algorithmic differentiation', *Nonlinear Analysis B Real World Applications*, Vol. 11, No. 5, pp.3993–4009.

Søndergaard, J. (2003) 'Optimization using surrogate models – by the space mapping technique', PhD thesis, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby. Supervisor: Kaj Madsen.

Sarmiento, J. and Gruber, N. (2006) *Ocean Biogeochemical Dynamics*, Princeton University Press, Princeton, Oxford.

Tarantola, A. (2005) *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM, Philadelphia.

Notes

1 MATLAB is a registered trademark of the MathWorks, Inc., http://www.mathworks.com.

Journal of Computational Science 2 (2011) 335–344

Contents lists available at SciVerse ScienceDirect

Journal of Computational Science

journal homepage: www.elsevier.com/locate/jocs

Surrogate-based optimization of climate model parameters using response correction

M. Prieß [a,*], S. Koziel [b], T. Slawig [a]

[a] Institute for Computer Science, Cluster The Future Ocean, Christian-Albrechts Universität zu Kiel, 24098 Kiel, Germany
[b] Engineering Optimization & Modeling Center, School of Science and Engineering, Reykjavik University, Menntavegur 1, 101 Reykjavik, Iceland

ARTICLE INFO

Article history:
Received 6 March 2011
Received in revised form 12 August 2011
Accepted 21 August 2011
Available online 26 August 2011

Keywords:
Climate models
Marine ecosystem models
Surrogate-based optimization
Parameter optimization
Response correction

ABSTRACT

We present a computationally efficient methodology for the optimization of climate model parameters applied to a (one-dimensional) representative of a class of marine ecosystem models. We use a response correction technique to create a surrogate from a temporarily coarser discretized physics-based low-fidelity model. We demonstrate that replacing the direct parameter optimization of the high-fidelity ecosystem model by iteratively updating and re-optimizing the surrogate leads to a very satisfactory solution while yielding significant cost saving – about 84% when compared to the direct high-fidelity model optimization.

1. Introduction

In this paper we present the application of a *Surrogate-based Optimization* approach, based on a multiplicative response correction, on parameter identification problems in a climate model.

Surrogate-based optimization is a methodology to efficiently optimize complex, so-called *high-fidelity* or *fine* models that require substantial computational effort already for a model evaluation. As a consequence, optimization and control problems for them are often still beyond the capability of modern mathematical algorithms and computer power. The idea of surrogate-based optimization is to introduce a surrogate, a computationally cheap and yet reasonably accurate representation of the high-fidelity model. The surrogate replaces the high-fidelity model in the optimization process in the sense of providing predictions of the model optimum. Also, it is updated using the high-fidelity model data accumulated during the process. The prediction-updating scheme is normally iterated in order to refine the search and to locate the high-fidelity model optimum as precisely as possible. The surrogate can be created by approximating sampled high-fidelity model data or by employing a physically-based *low-fidelity* or *coarse* model. In this work, we use the latter approach. The low-fidelity model is normally less accurate, therefore, it has to be iteratively corrected by suitable methods. The correction (or alignment) can be realized using a limited number (in many cases, only one) evaluations of the high-fidelity model and possibly also its derivatives. Surrogate-based optimization is widely and very successfully used in engineering sciences, compare [1–4].

The development and use of low-fidelity models obtained by, e.g. coarser discretizations (in time and/or space) or by parametrizations is common in climate research [5]. We point out that surrogate-based optimization is not just a direct low-fidelity model optimization, but includes a special alignment between the high- and the low-fidelity model. In that sense, the application of surrogate-based methods on parameter optimization in climate models is new.

Climate models are typically given as time-dependent partial differential or differential algebraic equations (PDEs/DAEs) [5–7]. Since the number of processes that have to be included and the needed temporal and spatial resolution is quite high, so is the computational effort. As a result, many processes on small temporal or spatial scales are, as denoted in the climate community, *parameterized*, i.e., they are represented by simpler models that usually include a number of parameters that have to be properly chosen or adjusted. A typical example – not only used in climate models for ocean or atmosphere simulations – is turbulence modeling [8]. There are also processes in the climate system where even

* Corresponding author. Tel.: +49 0431 880 7452; fax: +49 0431 880 7618.
E-mail addresses: mpr@informatik.uni-kiel.de (M. Prieß), koziel@ru.is (S. Koziel), ts@informatik.uni-kiel.de (T. Slawig).
URLs: http://www.informatik.uni-kiel.de/co2/mitarbeiterinnen/dipl-phys-malte-priess/ (M. Prieß), http://koziel.ru.is (S. Koziel), http://www.informatik.uni-kiel.de/co2/mitarbeiterinnen/prof-dr-thomas-slawig/ (T. Slawig).

1877-7503/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jocs.2011.08.004

without much simplification several quantities or parameters are unknown or very difficult to measure. This is for example the case for growth and dying rates in marine ecosystem models [9,10], one of which is taken as a test case for the surrogate-based optimization approach we analyze in this paper. Marine ecosystem models describe photosynthesis and other biogeochemical processes in the marine ecosystem that are important, e.g., to compute and predict the oceanic uptake of carbon dioxide (CO_2) as part of the global carbon cycle [9]. We point out that the mathematical formulation of the climate models we use is quite general, such that our approach is not limited to them but remains applicable for a wide range of time-dependent models.

The aim of parameter optimization is to adjust or identify the model parameters, ideally to fit given measurement data by the corresponding model output [11]. The mathematical task thus can be classified as a least-squares type optimization or inverse problem [12,13]. The number of optimization parameters range from about 10 to 100 discrete real-valued ones in marine ecosystem models (where they are growth and dying rates, etc.) up to distributed functions (or thousands and more discrete values after discretization), for example when an initial model state or boundary condition is unknown and target of the optimization. The optimization parameters and the model state are coupled by the constraint of the time-dependent PDE, i.e., the climate model. Additionally, constraints on the parameters (e.g., non-negativity of growth-rates in ecosystem models) and on the state variables (non-negativity of concentrations of biological species as algae, etc. or of temperature) might be given.

This optimization process requires a substantial number of (typically expensive) function and optionally sensitivity or gradient or even Hessian matrix evaluations. If the latter are computed by finite difference approximations, the critical quantity determining the computational effort of the optimization is that of the cost function evaluation, which is basically a single model simulation. Hence, decreasing the effort related to the function evaluations (or, equivalently, cutting down the number of function calls necessary to find the optimum) is of primary importance to reduce the overall optimization cost. This becomes particularly significant for computationally expensive three-dimensional coupled models, as for example global climate models [7].

In this paper we analyze the application of a *multiplicative response correction technique* to create a surrogate for one specific type of a climate model, a one-dimensional marine ecosystem model that uses pre-computed ocean circulation data [14]. This model was chosen because here extensive optimization runs with different methods including local, gradient-based and so-called global, genetic algorithms have been performed, see [15]. The underlying physically-based low-fidelity model is obtained from a temporarily coarser discretization of the high-fidelity one. We verify our approach by using synthetic target data and by comparing the results of surrogate-based optimization to those obtained from the direct fine model optimization. The application on real data is performed as a next step. Furthermore, this exemplary application shall serve as a test for three-dimensional model runs, which are much more costly with respect to computing time and storage.

The structure of the paper is as follows: The general form of climate models and the parameter optimization problem considered is described in Section 2. We first recall the basic idea of surrogate-based optimization in Section 3. The ecosystem model, which is taken as an example in this paper, is introduced in Section 4, and its low-fidelity counterpart that we use as a basis for the surrogate is described in Section 5. The response correction, the construction of the surrogate model and the quality of the surrogate are described and analyzed in Section 6. The setup of the optimization which is used to compare the results is given in Section 7. Numerical results and discussion of an exemplary test run are provided in Section 8. Section 9 concludes the paper with a summary and an outlook.

2. Model equations and optimization problem

In this section we give the formulations of what we call a *model* and of the corresponding parameter optimization problem. Our formulations are quite general and appropriate for a big class of applications, for which climate models are only one example.

2.1. Continuous and discrete model formulation

We start from an initial boundary value problem (IBVP) for a system of time-dependent partial differential or differential algebraic equations (PDEs/DAEs) of the following form:

$$\left.\begin{aligned} E\frac{\partial y}{\partial t} &= f(y,u) && \text{in } \Omega\times(0,T)\\ y(x,0) &= y_{init}(x) && \text{in } \Omega\\ y(x,t) &= y_{bdr}(x,t) && \text{on } \partial\Omega\times(0,T). \end{aligned}\right\} \tag{1}$$

Here y is the vector of the *state variables*, and E is a matrix with the size of y, typically being the identity matrix for a PDE while having rank deficiency for a DAE [16]. We include DAEs in this formulation since in climate models, e.g., ocean circulation models, the Navier–Stokes equations [17] are an important part, and – after space discretization – take the form of a DAE system. Then y may for example consist of velocity field, pressure, temperature and salinity. In our example of a marine ecosystem model (which is formulated as PDE system), the matrix E can be set to the identity and thus omitted. In this case the state vector y contains all relevant biogeochemical tracers as phyto- and zooplankton, etc., see Section 4 for the details.

The spatial domain Ω is an open subset of $\mathbb{R}^d$, $d\in\{1,2,3\}$, with boundary $\partial\Omega$. We assume that the initial time is $t=0$, and define the end of the considered time interval as $T>0$. The right-hand side of the system is given by a vector-valued and usually nonlinear function f which includes spatial differential operators. In climate models, it often additionally depends explicitly on the space and time variables x and t, respectively, which is skipped in the notation. Moreover f depends on a number of model parameters which are summarized in the vector u. The vector-valued functions y_{init} and y_{bdr} are given initial and boundary values. They can also be subject to optimization, but for simplicity of notation and since in our example this is not the case, we do not formulate (1) this way. The IBVP given by (1) can be called a *continuous model*, since its solution – if it exists – is mathematically given in an appropriate function space Y.

After spatial and temporal discretization, we obtain a discretized or *discrete model*. We will consider here a time discretization which is performed by a sequential integration at time steps $0=t_0<\cdots<t_j<\cdots<t_M=T$. In many applications (among them climate models), this integration is at least partially implicit, i.e., it involves the solution of linear or non-linear systems. With respect to the spatial discretization, we are quite general, i.e., it may be either finite volumes or differences (as in most climate models) or finite elements. We assume that we have K discrete values for the discrete approximation $\mathbf{y}_j$ of the state y at time t_j, i.e., $\mathbf{y}_j\in\mathbb{R}^K$ and $\mathbf{y}_j\approx y(\cdot,t_j)$. Thus, after discretization we end up with the following discrete model:

$$\left.\begin{aligned} \mathbf{y}_0 &\approx y_{init},\\ \mathbf{y}_{j+1} &= \Phi_j(\mathbf{y}_j,\mathbf{u}), \quad j=1,\dots,M-1. \end{aligned}\right\} \tag{2}$$

Again, only for simplicity of notation, we do not include multi-step time integration schemes here, but they can be easily treated similarly. The first line in (2) indicates that already the discrete initial

value usually is only an approximation (or restriction) of the exact initial data y_{init} given in the continuous model (1). We use the boldface notation for discrete vectors, e.g., here for $\mathbf{y}_j \in \mathbb{R}^I$. The whole discrete state $\mathbf{y} := (\mathbf{y}_j)_{j=1,\ldots,M}$ then is in $\mathbb{R}^{MI}$. Note that I not necessarily equals the number of points in the spatial grid, but is often a multiple of it since we have a system of equations in (1) and y consists of several components. We moreover introduced a discrete vector $\mathbf{u} \in \mathbb{R}^n$ of parameters, either after their discretization, or (as in our model example), since they have been real numbers, from the beginning. Here we assume that the parameters do not depend on the time t_j and thus have no index j. The right-hand sides Φ_j in (2) depend on f, the actual time t_j (since f depends on it), the time discretization scheme and its step size.

If the discretization schemes in space and time are chosen properly, we can regard the discrete model as a mapping

$$\mathbf{u} \mapsto \mathbf{y}(\mathbf{u})$$

from the parameters to the discrete space state, i.e., from $\mathbb{R}^n =: U$ to $\mathbb{R}^{MI} =: Y$.

2.2. Optimization problem

In this subsection we formulate the optimization problem for the discrete model. Omitting the boldface notation, the same formulation holds for the continuous model, but naturally would require further analysis, which is beyond the scope of this paper.

The key task in parameter optimization is to minimize a least-squares type cost function measuring the misfit between the discrete model output $\mathbf{y} = \mathbf{y}(\mathbf{u})$, i.e., the solution of (2), and given observational data $\mathbf{y}_d$ [11,13]. We assume that $\mathbf{y}_d \in Y$, otherwise an appropriate observation/restriction operator has to be introduced. In most cases, the cost function is constrained by parameter bounds. Thus the parameter optimization problem can be written as

$$\min_{\mathbf{u} \in U_{ad}} J(\mathbf{y}(\mathbf{u})) \tag{3}$$

where

$$J(\mathbf{y}) := \frac{1}{2} \left\| \mathbf{y} - \mathbf{y}_d \right\|_Y^2, \qquad U_{ad} := \{\mathbf{u} \in \mathbb{R}^n : \mathbf{b}_l \le \mathbf{u} \le \mathbf{b}_u\},$$

$$\mathbf{b}_l, \mathbf{b}_u \in \mathbb{R}^n, \qquad \mathbf{b}_l < \mathbf{b}_u.$$

The inequalities in the definition of the set U_{ad} of admissible parameters are meant component-wise. The functional J may additionally include a regularization term for the parameters, which was not necessary in our case.

Additional constraints on the state variable $\mathbf{y}$ might be necessary, e.g., to ensure non-negativity of the temperature or of the concentrations of biogeochemical quantities. In our example model however, non-negativity of the state variables can be ensured by using appropriate parameter bounds $\mathbf{b}_l$ and $\mathbf{b}_u$. This was already observed and used in [15].

3. Surrogate-based optimization

For many nonlinear optimization problems, a high computational cost of evaluating the objective function and its sensitivity, and, in some cases, the lack of sensitivity information, is a major bottleneck. The need for decreasing the computational cost of the optimization process is especially important while handling complex three-dimensional models.

Surrogate-based optimization (SBO) [1–4] addresses these issues by replacing the original high-fidelity model $\mathbf{y}$ by its surrogate model $\mathbf{s}$. The surrogate should be computationally cheap and analytical tractable. It can be obtained by approximating the sampled high-fidelity model data using a suitable technique, e.g., polynomial regression [1], kriging [18] or support-vector regression [19]. These so-called *functional surrogates* are constructed without any particular knowledge of the system and will not be addressed further in this paper.

Another possibility, explored in this paper, is to construct the surrogate through correction/alignment of a physical low-fidelity or coarse model, a less accurate but computationally cheap representation of $\mathbf{y}$. These so-called *physically-based surrogates* [20] inherit more characteristics of the high-fidelity model under consideration. Possible ways to create the underlying physical low-fidelity model are by using a coarser discretization (while employing the same simulation tool as for the high-fidelity model), simplified physics or different ways of describing the same physical phenomenon or even by using analytical formulas if available. In this paper we use a coarser discretization in time to create a physical low-fidelity model (cf. Section 5).

The low-fidelity model correction aims at reducing misalignment between the low- and high-fidelity model output. The specific correction technique exploited in this work is described in detail in Section 6. The surrogate model is updated at each iteration of the optimization algorithm, typically using available high-fidelity model data. In particular, the surrogate model $\mathbf{s}_k$ at iteration k can be constructed by only using the high-fidelity model output $\mathbf{y}(\mathbf{u}_k)$ at the current optimization variable vector $\mathbf{u}_k$ and the corresponding low-fidelity model output. The next iterate, $\mathbf{u}_{k+1}$, is obtained by optimizing the surrogate $\mathbf{s}_k$, i.e.,

$$\mathbf{u}_{k+1} = \operatorname*{argmin}_{\mathbf{u} \in U_{ad}} J(\mathbf{s}_k(\mathbf{u})). \tag{4}$$

Then the updated surrogate $\mathbf{s}_{k+1}$ is determined by re-aligning the low-fidelity model at $\mathbf{u}_{k+1}$ and optimized again as in (4). The process of aligning the low-fidelity model to obtain the surrogate and subsequent optimization of this surrogate is repeated until a user-defined termination condition is satisfied, which can use certain convergence criteria, assumed level of cost function value or a specific number of iterations (particularly if the computational budget of the optimization process is limited). A discussion of the termination condition used in this work can be found in Section 8.

A well performing surrogate-based algorithm is capable of yielding a satisfactory solution at a low computational cost, typically corresponding to only a few evaluations of the high-fidelity model. The key prerequisites to ensure this are a cheap and yet reasonably accurate low-fidelity model as well as a properly selected and low-cost alignment procedure (i.e., using a limited number of high-fidelity model evaluations, preferably just one).

If the surrogate $\mathbf{s}_k$ satisfies so-called 0-order and 1st-order consistency conditions [21,22] with the high-fidelity model at $\mathbf{u}_k$, i.e.,

$$\mathbf{s}_k(\mathbf{u}_k) = \mathbf{y}(\mathbf{u}_k), \qquad \mathbf{s}_k{}'(\mathbf{u}_k) = \mathbf{y}'(\mathbf{u}_k), \tag{5}$$

the surrogate-based scheme (4) is provable convergent to at least a local optimum of (3), provided that both the low- and high-fidelity models are sufficiently smooth, and the surrogate optimization step is enhanced by the trust-region (TR) safeguard [21,22], i.e.,

$$\mathbf{u}_{k+1} = \operatorname*{argmin}_{\mathbf{u} \in U_{ad}, \|\mathbf{u} - \mathbf{u}_k\| \le \delta_k} J(\mathbf{s}_k(\mathbf{u})),$$

with δ_k being the trust-region radius updated according to the TR rules.

Note that the 1st-order consistency requires high-fidelity sensitivity data, which is not utilized here. In this work, the surrogate is defined to satisfy the 0-order consistency only which is sufficient to ensure good performance as demonstrated in Section 6.3 and Section 8.

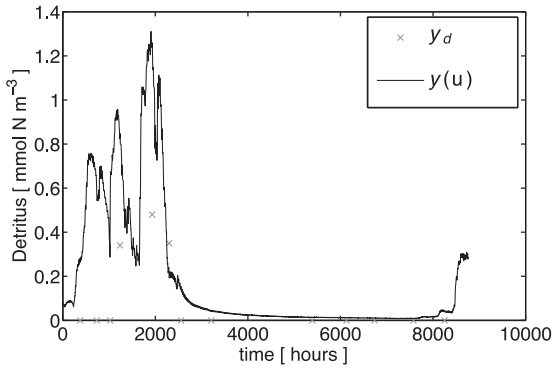


Fig. 1. Model output $\mathbf{y}^{(D)}$ (detritus) and observation data $\mathbf{y}_d^{(D)}$ for one year at depth $z \simeq -25$ m.

4. Example: a marine ecosystem model

In this section we briefly describe the model – both in continuous and discrete form – that we used to analyze the application of a surrogate-based optimization approach on a climate model. The considered example is a one-dimensional marine ecosystem model driven by pre-computed ocean circulation data [14].

4.1. The continuous model

Simulating the marine ecosystem has become a key tool for understanding the ocean carbon cycle and its variability. The marine ecosystem contains several biogeochemical quantities (called *tracers*), for example nutrients, phyto- and zooplankton which interact and are moreover transported by the ocean circulation and influenced by temperature and salinity. Thus ecosystem simulations require modeling and computation both of ocean circulation and biogeochemistry. The underlying continuous models are governed by coupled systems of nonlinear, parabolic PDEs or DAEs, for ocean circulation (ocean models, i.e., Navier–Stokes equations with additional temperature and salinity transport equations) and transport of biogeochemical tracers (marine ecosystem models, i.e., convection- or advection-diffusion-reaction type equations) [9]. Thus they fit in our general formulation (1) and its discrete counterpart (2).

In ecosystem models, the parameters to be optimized – summarized in the vector $\mathbf{u}$ in (2) – are for example growth and dying rates of the tracers and thus appear in the usually nonlinear coupling or interaction terms in the model.

Our example ecosystem model was developed by Oschlies and Garcon [14] and simulates the interaction of dissolved inorganic nitrogen, phytoplankton, zooplankton and detritus (thus also called *NPZD* model).

One aim was to reproduce observations $\mathbf{y}_d$ at different North Atlantic locations by the optimization of model parameters within credible limits. Fig. 1 shows the model output and target data, respectively, as illustration for the tracer detritus for a certain depth and a part of the time interval. The model uses pre-computed ocean circulation and temperature data from an ocean model (in a sometimes called *off-line modus*), i.e., no feedback by the biogeochemistry on the circulation and temperature is modeled [14]. Thus the continuous model (1) here just contains the biochemistry, whereas all circulation data are hidden in the right-hand side f.

As a test case and since biogeochemistry – except for sinking processes – mainly happens locally in space, we use here a one-dimensional version of the model. This version simulates one water column at a given horizontal position. This is additionally motivated by the fact that there have been special time series studies at fixed locations. Clearly the computational effort in a one-dimensional simulation is significantly smaller than in the three-dimensional case. Thus, before going to 3D, this model serves as a good test example for the applicability of surrogate-based optimization approaches, since it includes all significant features of ecosystem models.

In the *NPZD* model, the concentrations (in mmol N m^{-3}) of dissolved inorganic nitrogen N, phytoplankton P, zooplankton Z, and detritus (i.e., dead material) D are summarized in the vector $y = (y^{(l)})_{l=N,P,Z,D}$ and described by the following coupled PDE system

$$\left.\begin{aligned} \frac{\partial y^{(l)}}{\partial t} &= \frac{\partial}{\partial z}\left(\kappa \frac{\partial y^{(l)}}{\partial z}\right) + Q^{(l)}(y, u_2, \ldots, u_n), && l = N, P, Z \\ \frac{\partial y^{(D)}}{\partial t} &= \frac{\partial}{\partial z}\left(\kappa \frac{\partial y^{(D)}}{\partial z}\right) + Q^{(D)}(y, u_2, \ldots, u_n) - \frac{\partial y^{(D)}}{\partial z} u_1, && l = D \end{aligned}\right\} \quad (6)$$

$$\text{in } (-H, 0) \times (0, T) \quad (6)$$

with additional appropriate initial values. Here, z denotes the only remaining, vertical spatial coordinate, and H the depth of the water column. The terms $Q^{(l)}$ are the biogeochemical coupling (or *source-minus-sink*) terms for the four tracers and $\mathbf{u} = (u_1, \ldots, u_n)$ is the vector of unknown physical and biological parameters. The sinking term is only apparent in the equation for detritus. In the one-dimensional model no advection term is used, since a reduction to vertical advection would make no sense. Thus, the circulation data (taken from an ocean model) are the turbulent mixing coefficient $\kappa = \kappa(z, t)$ and the temperature $\Theta = \Theta(z, t)$, which goes into the nonlinear coupling terms $Q^{(l)}$ but is omitted in the notation.

4.2. Discretization scheme and discretized model

The continuous model (6) is discretized and solved using an operator splitting method, which for a given a time step τ reads

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

$$j = 1, \ldots, M. \quad (7)$$

Recall that by $\mathbf{y}_j$ we denote the discrete solution in time step j given as

$$\mathbf{y}_j = (y_{ji})_{i=1,\ldots,I}, \quad j = 1, \ldots, M. \quad (8)$$

at the discrete spatial points. Since in our case the model output consists of four tracers, I denotes the number of spatial discrete points times 4. If the discrete state $\mathbf{y}_j$ is given in such a way that the four discrete tracer vectors at the time step j are concatenated, the matrices A_j^{diff}, A^{sink} in (7) are (4×4)-block-diagonal matrices. They represent the discretization of the diffusion (with second order central differences) and the sinking (discretized by an upstream scheme), respectively.

In every time step $j \to j+1$, at first the nonlinear coupling operators Q_j (that depend on t_j directly and/or via the temperature field Θ) are computed at every spatial grid point and integrated by four explicit Euler steps, each of which is described by the nonlinear operator

$$B_j^Q(\mathbf{y}_j) := \left[\mathbf{y}_j + \frac{\tau}{4} Q_j(\mathbf{y}_j)\right].$$

Note that, for simplicity, we omitted the additional arguments of the term Q_j in the formulation above. Then, an explicit Euler step with full step size τ is performed for the sinking term. This step is represented by the matrix B^{sink}. Since the sinking velocity is temporarily constant, this matrix does not depend on the time step j. Finally, an implicit Euler step for the diffusion operator is applied.

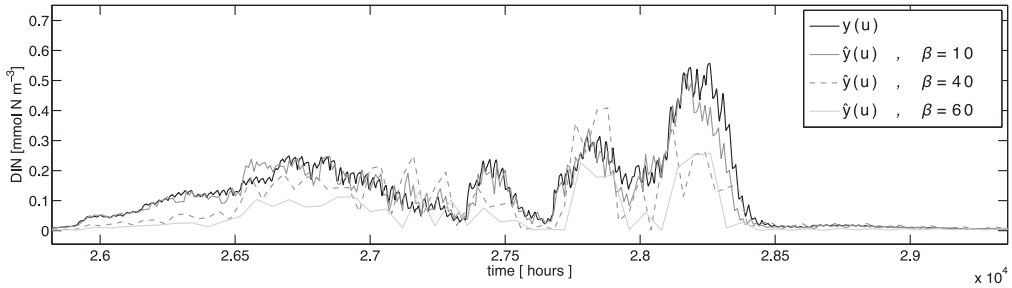


Fig. 2. Fine and coarse model output $\mathbf{y}$, $\hat{\mathbf{y}}$, respectively, for the state dissolved inorganic nitrogen at depth $z \approx -2.68$ m for different values of the coarsening factor β and the same randomly chosen parameter vector $\mathbf{u}$. For simplicity we skip super- and subscripts in the legends of all figures.

Due to $\kappa = \kappa(z, t)$ the resulting matrix B_j^{diff} depends on j and is non-symmetric [23, Section 5]. It is tridiagonal, and the system is solved directly by splitting it up into the four blocks. Writing this last step formally as a matrix inversion, formulation (7) corresponds to (2).

In the original discrete model (6) the time step τ is chosen as one hour. By chosing this time step all relevant processes are captured and further decrease of the time step does not improve the accuracy of the model. From now on, we refer to this version as the high-fidelity or fine model.

5. The low-fidelity model

In this paper we use a low-fidelity (or coarse) model which has a coarser time discretization to obtain a physically-based surrogate (cf. Section 6).

Using a coarser discretization (in space or time or both) is probably the most straightforward way to create a physical low-fidelity model and most likely quite easy to implement. It also turns out that such coarse models are quite suitable for our chosen correction approach yielding a reasonable approximation of our fine model (cf. Section 6).

5.1. Coarser time discretization

The coarse model is obtained by using a coarser time discretization with

$$\hat{\tau} = \beta\tau$$

with $\tau = 1$ h, the time step of the fine model and with a *coarsening factor* $\beta \in \mathbb{N} \setminus \{0, 1\}$, while keeping the spatial discretization fixed. The state variable for this coarser discretized model will be denoted by $\hat{\mathbf{y}}$, the corresponding number of discrete time steps by $\hat{M} = M/\beta$. Note that the parameters $\mathbf{u}$ for this coarse model are the same as for the fine model. Fig. 2 shows the fine and coarse model output $\mathbf{y}$, $\hat{\mathbf{y}}$ for the state dissolved inorganic nitrogen, for different values of β and at the same randomly chosen parameter vector $\mathbf{u}$.

It follows from Fig. 2 that the choice of the coarsening factor determines the quality of the coarse model. This will in turn affect the quality of a surrogate based on this model and hence will also affect the solution obtained by a SBO (cf. Section 3).

On the one hand, if choosing β too large, we might not capture all relevant characteristics of the fine model output (cf. Fig. 2) while, in turn, the cost for one coarse model evaluation and hence of a SBO run will be low. The solution of this optimization might not be accurate enough or the algorithm may not converge. Moreover, it is important to keep in mind that choosing β too large could lead to numerical instabilities [24] resulting in oscillations of the model output.

On the other hand, choosing a smaller value for β will improve the accuracy of the coarse model and the surrogate whereas leading to an increased cost for one coarse model evaluation and a SBO run.

Overall, we seek for a reasonable trade-off between the accuracy and speed of the coarse model. By using a coarsening factor of $\beta = 40$, the most relevant characteristics of the fine model output can be captured while the cost for one coarse model evaluation is reasonably low. Furthermore, the resulting coarse model output does not show any numerical instabilities. This was investigated by visual inspection of the model output (cf. Fig. 2) and motivates the use of this value in the optimization.

6. The surrogate

The surrogate model is constructed here in a simple way using a multiplicative response correction of the coarse model. The correction term is calculated at the beginning of each iteration of the algorithm (4) using a single fine model evaluation. It turns out that this way of correcting the coarse model is quite suitable for the considered problem because the relation between the coarse and the fine model response values is rather well preserved for various sets of parameters $\mathbf{u}$, at least locally.

6.1. Smoothing

As the coarse model output is very noisy (cf. Fig. 2), it is necessary to smoothen the coarse and, consequently, also the fine model output before calculating the multiplicative correction factors. Initial experiments indicated (details omitted for the sake of brevity) that the surrogate-based optimization exploiting the unsmoothed model outputs is not able to yield a reasonable solution.

For the smoothening of the fine and coarse model output $\hat{\mathbf{y}}$, $\mathbf{y}$, respectively, we use a walking average with span $\pm n$ given as:

$$\begin{aligned} \tilde{\hat{y}}_{ji} &:= \frac{1}{2n+1}\sum_{m=j-n}^{j+n}\left[\frac{1}{2n+1}\sum_{p=m-n}^{m+n}\hat{y}_{pi}\right] \\ \tilde{y}^{\beta}_{ji} &:= \frac{1}{2n+1}\sum_{m=j-n}^{j+n}\left[\frac{1}{2n+1}\sum_{p=m-n}^{m+n} y^{\beta}_{pi}\right] \\ j &= 1,\ldots,\hat{M}, \quad i = 1,\ldots,I, \end{aligned} \tag{9}$$

where j, i are the temporal and spatial indices, respectively (cf. (8)) and where we used the *down-sampled* fine model output, denoted by $\mathbf{y}^{\beta} \in \mathbb{R}^{\hat{M}I}$, which is given by

$$y^{\beta}_{ji} := y_{\beta j,i}, \quad j = 1,\ldots,\hat{M}, \quad i = 1,\ldots,I, \tag{10}$$

to be commensurable with the coarse model output. In this paper, we use $n = 3$. Also, the smoothing is performed twice. 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. It turns out, also by visual inspection, that the chosen value of $n = 3$ and "double" smoothing are suitable for the considered problem.

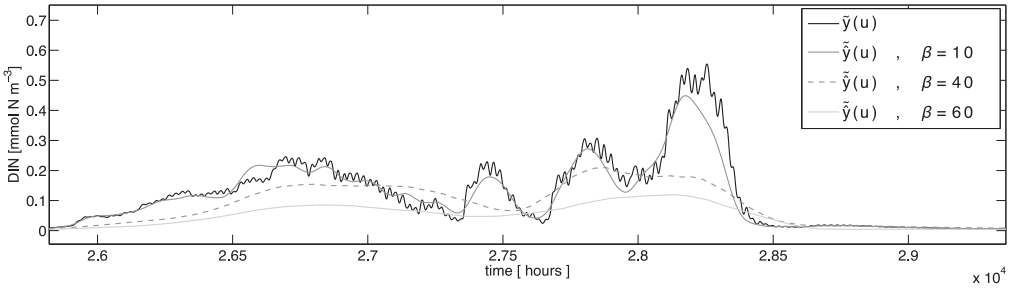


Fig. 3. Same as in Fig. 2 but now using smoothing (cf. (9)) for both the coarse and the fine model. Smoothing helps removing the numerical noise in the model outputs so that the optimization process is able to identify and track relevant changes of the traces of interest.

Fig. 3 shows the fine and coarse model outputs as in Fig. 2 while here using smoothing and, for comparison with the curves in Fig. 2, again for different values of the coarsening factor β. As was motivated in Section 5.1, we use a coarse model with a coarsening factor of $\beta = 40$ in the optimization.

6.2. Response correction

In this work, the surrogate model output is generated, at iteration k of the optimization process, by multiplicative correction of the coarse model output (cf. Section 3). The *correction factor*, denoted as A_{kji}, is defined by pointwise division of the smoothed fine by the smoothed coarse model output at the point $\mathbf{u}_k$, i.e.,

$$\left.\begin{array}{rcl} s_{kji}(\mathbf{u}) & := & A_{kji}\ \tilde{\tilde{y}}_{ji}(\mathbf{u}), \\ A_{kji} & := & \dfrac{\tilde{y}^{\beta}_{ji}(\mathbf{u}_k)}{\tilde{\tilde{y}}_{ji}(\mathbf{u}_k)} \end{array}\right\} \begin{array}{l} k = 1, 2, \ldots, \\ j = 1, \ldots, \hat{M}, \quad i = 1, \ldots, I, \end{array} \tag{11}$$

where $\tilde{\mathbf{y}}^{\beta}$ is given by (9). We call $A_k := (A_{kji})_{j,i} \in \mathbb{R}^{\hat{M}\times I}$ the *correction matrix* in step k. We use it to write the correction step in iteration k on the whole discrete state vector as

$$\mathbf{s}_k(\mathbf{u}) := A_k \circ \tilde{\tilde{\mathbf{y}}}(\mathbf{u}), \quad \mathbf{s}_k \in \mathbb{R}^{\hat{M}I}$$

where the operation "$\circ$" is defined by (11).

Note that the surrogate model is constructed using just one evaluation of the fine model. This simple correction scheme is justified by the fact that the overall "shape" of the coarse model output resembles that of the fine one. In particular, the high-value outputs for both models are corresponding to each other on the time scale, which is the consequence of the coarse model being physically based. Also, the relative changes of the outputs while changing the model parameters are similar for both coarse and fine models so that the multiplicative correction seems to be a natural choice.

It should be emphasized that our surrogate model does not use high-fidelity model sensitivity data. Still, as demonstrated in Section 8, it is able to yield remarkably good results, not only with respect to the quality of the final solution, but, most importantly, in terms of the low computational cost of the optimization process.

6.3. Consistency conditions and generalization capability

By definition, the surrogate model (11) satisfies the 0-order consistency condition (cf. (5)) in the point of alignment $\mathbf{u}_k$, i.e.,

$$\mathbf{s}_k(\mathbf{u}_k) = \mathbf{y}^{\beta}(\mathbf{u}_k).$$

As we do not use sensitivity information, the 1st-order consistency condition cannot be satisfied exactly. Nevertheless, our surrogate model exhibits quite good generalization capability, which means that the surrogate provides a reasonable approximation of the fine one in the neighborhood of $\mathbf{u}_k$. This is of primary importance from the point of view of the robustness of the optimization process.

In the following, in order to demonstrate the generalization property of the surrogate model, we analyze the quality of the surrogate during the optimization run.

For this purpose we consider an iterate, say $\mathbf{u}_k$, and $\overline{\mathbf{u}}_k \in B_\epsilon(\mathbf{u}_k)$, i.e., in the ball around $\mathbf{u}_k$ with the radius ϵ which is an estimate for the step size or trust-region radius in this iteration step (cf. Section 3). The surrogate model $\mathbf{s}_k$ is established at $\mathbf{u}_k$ so that we have $\mathbf{s}_k(\mathbf{u}_k) = \mathbf{y}^{\beta}(\mathbf{u}_k)$. We would like to have at least approximate satisfaction of the 0-order consistency condition at $\overline{\mathbf{u}}_k$, i.e., $\mathbf{s}(\overline{\mathbf{u}}_k) \approx \mathbf{y}^{\beta}(\overline{\mathbf{u}}_k)$.

We performed this test of the generalization property in every iteration of our surrogate optimization and show here the results for one iteration at the beginning and one towards the end. Since the step size within the optimization usually decreases at the end we chose a smaller ϵ in the second case. Fig. 4 shows the smoothed surrogate's, fine and coarse model output at those two iterations.

Obviously, since the surrogate's model output at $\mathbf{u}_k$ exactly reproduces the fine model output at $\mathbf{u}_k$, the surrogate exactly satisfies the 0-order consistency condition (6.3) in the point of alignment. Furthermore, in the neighborhood, the surrogate's model output still provides a very reasonable approximation of the fine one. This is due to the fact that the relation between the coarse and the fine model output is rather well preserved for various sets of parameters $\mathbf{u}$ (here $\mathbf{u}_k$ and $\overline{\mathbf{u}}_k$), which, in turn, is a consequence of both the coarse and fine model enjoying the same underlying physics. This is also verified by Fig. 4 (upper) comparing the coarse and the fine model at these two parameter vectors.

Specifically, this setup demonstrates that towards the end of the optimization run, our surrogate has even better generalization properties, which improves the robustness of the optimization process.

Occasionally, there might occur a situation where the coarse model output is close to zero (and maybe even negative due to approximation errors) and a few magnitudes smaller than the fine one, which leads to large (possibly negative) entries in the corresponding correction tensor A_k. If this is only true in one iterate $\mathbf{u}_k$ but not anymore in the neighborhood, the resulting surrogate might provide a poor approximation there. These issues can be dealt with using some simple means [25], e.g., by introducing upper and lower bounds for the multiplication factors as well as non-negative bounds for the model outputs (the negative output is non-physical and is a result of numerical errors due to using large time steps in the coarse model).

However, even without addressing these issues, our approach yields good results both in terms of the quality of the final solution and, most importantly, in terms of the relative reduction in the total optimization cost.

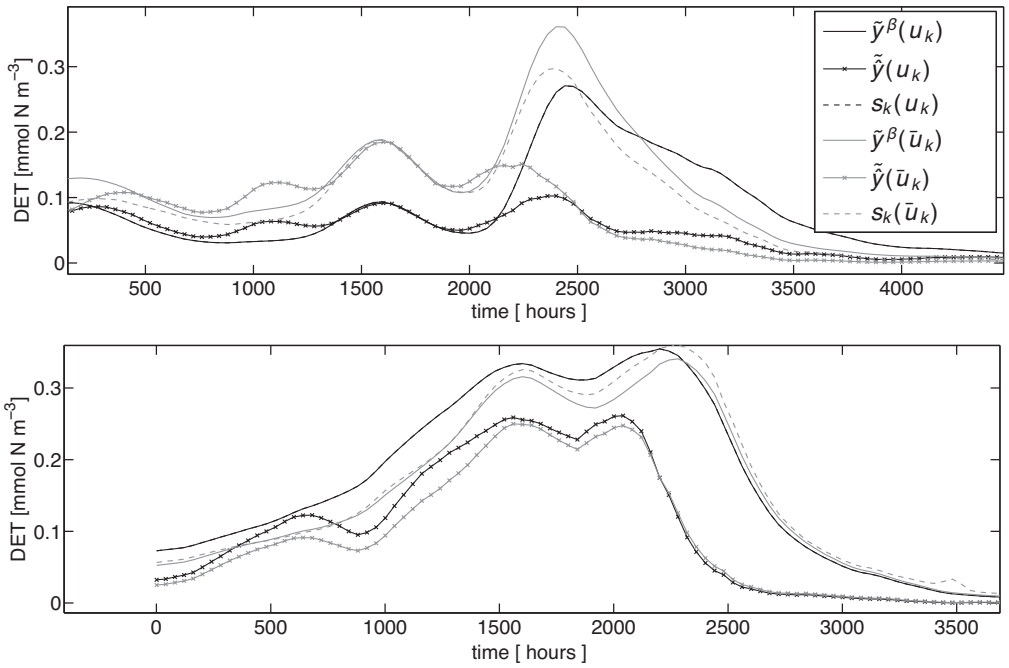


Fig. 4. Surrogate's, fine (down-sampled) and coarse model output $\tilde{\mathbf{y}}^{\beta}$, $\tilde{\tilde{\mathbf{y}}}$, $\mathbf{s}_k$ for the state detritus at depth $z \approx -2.68$ m and at two iterates $\mathbf{u}_k$ and with different neighborhood radii ϵ, see the text for details. The surrogate 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.

7. Optimization setup

The optimization approach proposed in this work has been tested using synthetic target data. We compare the quality of the solution and the computational cost of the surrogate-based optimization to those obtained by direct fine and coarse model optimization. For all optimizations we used the MATLAB[1] function `fmincon`, exploiting the active-set algorithm.

At a randomly chosen parameter vector $\mathbf{u}_d \in U_{ad}$ we computed the fine model output $\mathbf{y}(\mathbf{u}_d)$ and down-sampled it to be commensurable with the coarse and surrogate model outputs. The resulting data set is used as our synthetic target data $\mathbf{y}_d$ and given as:

$$(y_d)_{ji} := y^{\beta}_{ji}(\mathbf{u}_d), \quad j = 1, \ldots, \hat{M}, \quad i = 1, \ldots, I.$$

where $\mathbf{y}^{\beta}$ was defined in (10).

7.1. Cost functions

Now we define the following cost functions:

$$J(\mathbf{z}) := ||\,\mathbf{z} - \mathbf{y}_d||^2 = \sum_{i=1}^{I} \sum_{j=1}^{\hat{M}} (z_{ji} - (y_d)_{ji})^2, \tag{12}$$

$$\tilde{J}(\mathbf{z}) := ||\,\mathbf{z} - \tilde{\mathbf{y}}_d\,||^2 = \sum_{i=1}^{I} \sum_{j=1}^{\hat{M}} (z_{ji} - (\tilde{y}_d)_{ji})^2, \quad \mathbf{z} \in \mathbb{R}^{\hat{M}I}. \tag{13}$$

Note that, since for the optimization of the coarse model and the surrogate we have to consider the smoothed model output (see Section 6.1), we also have to consider the smoothed target data, yielding $\tilde{\mathbf{y}}_d$.

[1] MATLAB is a registered trademark of The MathWorks, Inc., http://www.mathworks.com.

In order to yield a fair comparison between the results obtained from direct fine model optimization and those obtained from coarse model and surrogate optimization we also consider the sampled fine model output $\mathbf{y}^{\beta}$ given by (10) for the fine model optimization. We thus have the following three optimization problems:

- Fine model optimization:

$$\mathbf{u}^* := \operatorname*{argmin}_{\mathbf{u} \in U_{ad}} J(\mathbf{y}^{\beta}(\mathbf{u}))$$

- Coarse model optimization:

$$\hat{\mathbf{u}}^* := \operatorname*{argmin}_{\mathbf{u} \in U_{ad}} \tilde{J}(\tilde{\tilde{\mathbf{y}}}(\mathbf{u}))$$

- Surrogate optimization:

$$\mathbf{u}_{k+1} = \operatorname*{argmin}_{\mathbf{u} \in U_{ad}} \tilde{J}(\mathbf{s}_k(\mathbf{u})), \quad k = 0, 1, \ldots$$

We furthermore constitute that using smoothing of the fine model output is not essential. However it actually turned out that even when using smoothing, results of the fine model optimization are not significantly affected.

7.2. Optimization cost

We will denote the total optimization cost of the surrogate, the fine and of the coarse model optimization by C_s, C_f and C_c, respectively.

This cost is given in terms of the total number of *equivalent* fine model evaluations. For example, β evaluations of the coarse model used here with a coarsening factor β are equivalent to (or, as expensive as) one fine model evaluation. On the other hand, the cost of one iteration of the surrogate-based optimization procedure (in

terms of equivalent fine model evaluations) equals to the number of coarse model evaluations necessary to optimize the surrogate model divided by β, and increased by one (since there is only one fine mode evaluation necessary per iteration).

8. Results and discussion

The operation and performance of the proposed algorithm is illustrated through the results of an exemplary test run with $I = 33 \cdot 4$, $M = 8760 \cdot 5$ and $\beta = 40$, which means that we obtain $\tilde{M} = M/\beta = 1095$ discrete time steps for the coarse model.

Below, we consider the following quantities:

(i) The target $\mathbf{y}_d$, i.e., the sampled fine model output at a randomly chosen parameter vector $\mathbf{u}_d$.
(ii) The sampled fine model output $\mathbf{y}^\beta$ at another randomly chosen parameter vector $\mathbf{u}_0$, serving as initial value of the optimization runs.
(iii) At the result $\hat{\mathbf{u}}^*$ of a coarse model optimization.
(iv) At the result $\mathbf{u}_s^*$ of the surrogate-based optimization.
(v) At the output of a fine model optimization yielding $\mathbf{u}^*$.

For the comparison of the results of the three optimization approaches we use the value of the cost function J given in (12), i.e., the one using the sampled fine model output and unsmoothed target data.

For the results provided in the following paragraph, 13 re-alignments (cf. Section 3) of the surrogate were required to satisfy the termination condition $J(\mathbf{y}^\beta(\mathbf{u}_k)) \leq 50$. This particular value was selected as it ensures good visual agreement between the fine model output and the target. Furthermore, we used a specific number of iterations within each surrogate optimization run (4), here 7. The reason is that it is not necessary to run the surrogate model optimization until convergence: an approximate solution is sufficient as the surrogate model is not perfectly accurate so that using a fixed (and rather limited) number of function calls allows us to reduce the computational cost of the optimization process. We should also point out that choosing the above termination condition is up to the user and it is generally problem dependent. We refer the reader to Section 8.2 for a more thorough discussion.

Altogether, a good agreement between (i) and (v) would indicate a high quality of the algorithm exploited to optimize the model (fine/coarse/surrogate), whereas a good agreement between (iv) and (v) would mean that the SBO works well.

8.1. Numerical results

Fig. 5 illustrates the results of the fine, coarse model and surrogate optimization for the randomly chosen initial parameter vector $\mathbf{u}_0$. Corresponding parameters and values of the cost function J are given in Table 1. Furthermore, the table shows the total optimization cost of the fine model (C_f), the coarse model (C_f) and of the surrogate optimization (C_s) as were described in Section 7.2.

Note that Fig. 5 shows selected (representative) tracers for a part of the whole time interval at some distinct depth layers only. The total number of depth layers considered in the optimization process is 33 and the entire time scale is 43,800 so that it is impossible to present a full model output here. We emphasize that the qualitative behavior of the other tracers and at different times and spatial layers is similar.

Fig. 5 indicates that the direct fine model optimization yields a very reasonable final solution $\mathbf{y}^\beta(\mathbf{u}^*)$ of the target data $\mathbf{y}_d$. This corresponds to a cost function value of $J(\mathbf{y}^\beta(\mathbf{u}^*)) = 1.267\text{e}-02$ obtained after 983 function evaluations (cf. Table 1), hence leading to the optimization cost of $C_f = 983$. Furthermore, in can be observed that the parameters $\hat{\mathbf{u}}^*$ obtained by coarse model optimization provide only a rough approximation $\mathbf{y}^\beta(\hat{\mathbf{u}}^*)$ of the target data corresponding to $J(\mathbf{y}^\beta(\hat{\mathbf{u}}^*)) = 2.96\text{e}+03$. The optimization cost is only $C_c = 11.275$ equivalent fine model evaluations. Optimization of the surrogate finally provides a solution $\mathbf{u}_s^*$ with a remarkably good optimal fit $\mathbf{y}^\beta(\mathbf{u}_s^*)$ and parameter match corresponding to a cost function of $J(\mathbf{y}^\beta(\mathbf{u}_s^*)) = 48.527$.

The key point is that the computational cost of the SBO is low: only $C_s = 59.575$ equivalent fine model evaluations were required to yield $\mathbf{u}_s^*$. Roughly the same cost function value $J \approx 48$ was obtained by direct fine model optimization after $C_f = 375$ model evaluations. Altogether, a reduction in the total optimization cost of about 84% could be obtained by using this SBO approach.

We point out that the performance looks similar for other initial conditions $\mathbf{u}_0$ as well as for other target data. It is also worth noticing that although using different routines for fine/surrogate model optimization might yield different results, the relative reduction in the total optimization cost using the surrogate in the optimization run would probably be maintained. For example, in [15] better cost function values were obtained by direct fine model optimization using a different optimization method (other than MATLAB's `fmincon`) for the same problem and the same model.

8.2. Appropriate choice of number of alignment steps

It should be emphasized again that the SBO method presented in this paper does not use sensitivity information and that the surrogate model satisfies exactly only the 0-order consistency condition with the fine model (cf. Section 6.3). Because of the specific choice of the model alignment method that is tailored to the relationship between the coarse and the fine model, our algorithm is able to yield a rapid improvement of the cost function. On the other hand, the algorithm convergence can be quite slow in the vicinity of the optimal solution. Both points are illustrated in the following paragraphs.

Results are presented in Fig. 6 showing the value of the cost function J (cf. (12)) calculated at the single iterates of the fine and coarse model optimization runs (Fig. 5 and Table 1) and at those of this extended surrogate optimization run. The x-axis represents the number of equivalent fine model evaluations which were required to reach the given value of the cost function. The same figure indicates several points corresponding to the specific values of the reduction in the total optimization cost.

The point showing 84% reduction marks the result $\mathbf{u}_s^*$ which we presented in the previous paragraph corresponding to a value of the cost function $J(\mathbf{y}^\beta(\mathbf{u}_s^*)) \approx 48$ (cf. Fig. 5, Table 1).

The figure also shows that approximately 95% reduction could be achieved after only 4 equivalent fine model evaluations corresponding to a termination condition of $J(\mathbf{y}^\beta(\mathbf{u}_k)) \leq 2780$). Of course the quality of the final solution at this point is not as good as the quality of the solution given above in Fig. 5 and Table 1, i.e., the one obtained after approximately equivalent 60 fine model evaluations. It is worth noticing that with even more than those 60 model evaluations, no significant reduction in the cost function value J can be further achieved by the surrogate optimization process.

Decreasing the threshold value in the termination condition to $J(\mathbf{y}^\beta(\mathbf{u}_k)) \leq 0.1$ leads to a significant increase of the number of equivalent fine model evaluations in the SBO of approximately 400.

On the other hand, optimization of the coarse model yields a solution $\hat{\mathbf{u}}^*$, which was obtained after approximately 11 equivalent fine model evaluations (cf. Table 1) corresponding to $J(\mathbf{y}^\beta(\hat{\mathbf{u}}^*)) \approx 2960$. This result is much worse than that obtained using the surrogate.

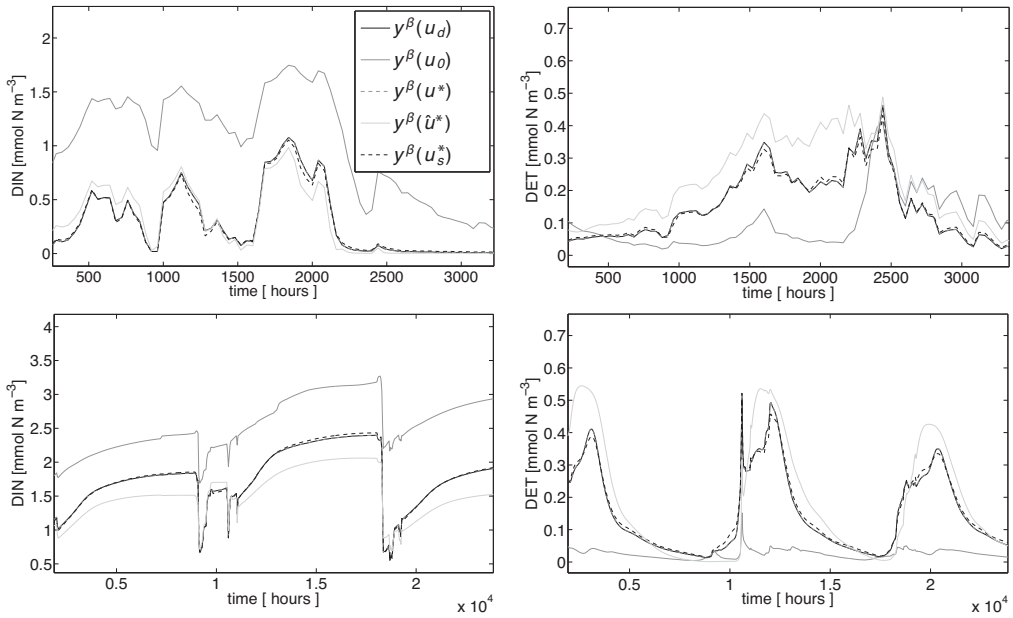


Fig. 5. Fine model output $\mathbf{y}^\beta$ (down-sampled) for state dissolved inorganic nitrogen (left) and the state detritus (right) at depth $z \approx -2.68$ m (top) and $z \approx -184.32$ m (bottom). Shown are, in the legend from top to bottom: (i) Target $\mathbf{y}_d$, i.e., the sampled fine model output at a randomly chosen parameter vector $\mathbf{u}_d$, (ii) fine model output at the initial value $\mathbf{u}_0$, (iii) at the result of the direct fine model optimization yielding $\mathbf{u}^*$, (iv) at the coarse model optimum $\hat{\mathbf{u}}^*$ and (v) at the optimum $\mathbf{u}_s^*$ obtained by surrogate optimization. Curves corresponding to (i), (iii) and (v) are very close. For clarity, the sampled fine model output is only shown at the selected (representative) time intervals. In the lower figures, a greater section can be shown since the model output at this deeper depth layer is not as noisy as in upper layers.

Table 1
Initial and optimal parameters $\mathbf{u}_0$, $\mathbf{u}^*$, $\hat{\mathbf{u}}^*$, $\mathbf{u}_s^*$, the corresponding values of the cost function J (which we use for comparison, cf. (12)) as well as the computational cost $C_i \in \{C_f, C_c, C_s\}$ (cf. Section 7.2) for an illustrative fine, coarse model and surrogate optimization run. The cost is given in terms of the total number of equivalent fine model evaluations required to obtain the given cost function value (see the text for details). Corresponding results in terms of the model output are given in Fig. 5.

Iterate	$u_{k,1}$	$u_{k,2}$	...									$u_{k,12}$	$J\ (\mathbf{y}^\beta(\mathbf{u}))$	C_i
$\mathbf{u}_0$	0.718	0.314	0.018	0.06	0.026	1.992	0.839	0.001	0.152	0.079	0.661	3.823	6.609e+04	
	Fine model optimization:			$\mathbf{u}^* := \text{argmin}_{\mathbf{u}\in U_{ad}}\ J(\mathbf{y}^\beta(\mathbf{u}))$										
$\mathbf{u}^*$	0.747	0.596	0.025	0.01	0.03	0.999	2.046	0.01	0.203	0.02	0.493	4.31	1.267e−02	983
	Coarse model optimization:			$\hat{\mathbf{u}}^* := \text{argmin}_{\mathbf{u}\in U_{ad}}\ \tilde{J}(\tilde{\mathbf{y}}(\mathbf{u}))$										
$\hat{\mathbf{u}}^*$	0.3	1.066	0.036	0.065	0.064	0.025	0.04	0.065	0.01	0.012	0.73	3.448	2.96e+03	11.275
	Surrogate optimization:			$\mathbf{u}_s^* := \text{argmin}_{\mathbf{u}\in U_{ad}}\ \tilde{J}(\mathbf{s}_k(\mathbf{u}))$										
$\mathbf{u}_s^*$	0.705	0.626	0.044	0.015	0.06	0.937	1.908	0.016	0.147	0.02	0.629	4.237	48.527	59.575
$\mathbf{u}_d$	0.75	0.6	0.025	0.01	0.03	1.0	2.0	0.01	0.205	0.02	0.5	4.32	∼ 84% reduction	

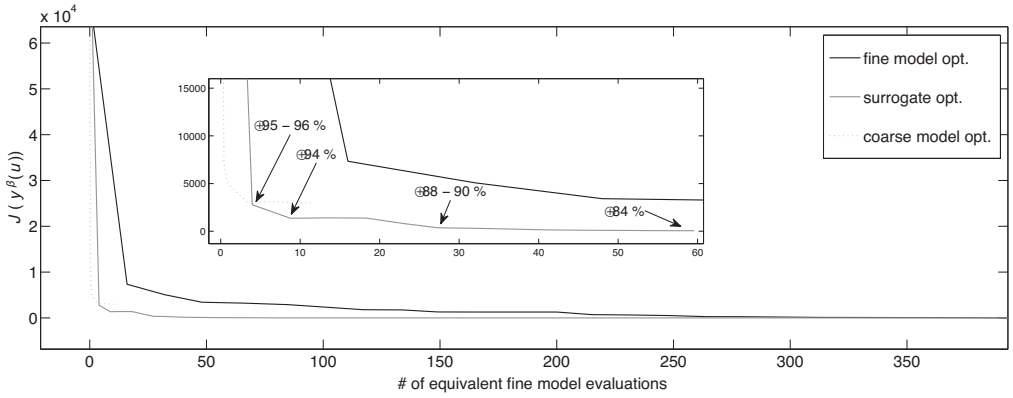


Fig. 6. The values of the cost function J (cf. (12)) versus the equivalent number of fine model evaluations for the fine, coarse and the SBO run. Several points corresponding to various values of the relative reduction in the total optimization cost (SBO versus straightforward fine model optimization) are also indicated. Results of fine model and surrogate optimization given in Fig. 5 and Table 1 correspond to the point marked as ∼84%.

9. Conclusions

Parameter optimization in climate models can be very expensive in terms of the cost function and gradient evaluations, especially for three-dimensional cases. Therefore, methods that aim at reducing the optimization cost, including surrogate-based optimization techniques, are highly desirable.

In this paper, we successfully applied a surrogate optimization technique to the optimization of a one-dimensional coupled marine ecosystem model. We use a physically-based surrogate constructed from a low-fidelity (or coarse) model that is the same as the original, high-fidelity (or fine) one, but utilizes a coarser time discretization. The surrogate is constructed through a simple multiplicative response correction of the coarse model. We demonstrated that the relation between the coarse and the fine model response values is rather well preserved for various sets of parameters, which shows that our correction method is quite suitable for the considered problem.

The optimization approach proposed in this work has been verified using synthetic target data. We furthermore compared the results, both in terms of the quality of the solution and the computational cost, to those obtained by direct fine and coarse model optimization. Although the direct fine model optimization yields an almost exact fit of the target data, its computational cost is high. On the other hand, the surrogate-based optimization produces a remarkably good results (both in terms of the quality of the final solution and the corresponding parameters) within a very few number of fine model evaluations only, resulting in a significant reduction of the total optimization cost of over 84%.

Acknowledgements

The authors would like to thank Andreas Oschlies, IFM Geomar, Kiel and Johannes Rückelt, Institute of Computer Science, Christian-Albrechts Universität zu Kiel. This research was supported by the DFG Cluster of Excellence Future Ocean.

References

[1] N.V. Queipo, R.T. Haftka, W. Shyy, T. Goel, R. Vaidyanathan, P.K. Tucker, Surrogate-based analysis and optimization, Prog. Aerosp. Sci. 41 (2005) 1–28.
[2] A.I. Forrester, A.J. Keane, Recent advances in surrogate-based optimization, Prog. Aerosp. Sci. 45 (2009) 50–79.
[3] J.W. Bandler, Q.S. Cheng, S.A. Dakroury, A.S. Mohamed, M.H. Bakr, K. Madsen, J. Søndergaard, Space mapping: the state of the art, IEEE Trans. Microw. Theory 52 (2004).
[4] L. Leifsson, S. Koziel, Multi-fidelity design optimization of transonic airfoils using physics-based surrogate modeling and shape-preserving response prediction, J. Comput. Sci. 1 (2010) 98–106.
[5] K. McGuffie, A. Henderson-Sellers, in: A Climate Modelling Primer, 3rd ed., Wiley, 2005.
[6] A. Majda, in: Introduction to PDE's and Waves for the Atmosphere and Ocean, AMS, 2003.
[7] A.E. Gill, in: Atmosphere–Ocean Dynamics, Volume 30 of International Geophysics Series, Academic Press, 1982.
[8] D.C. Wilcox, in: Turbulence Modeling for CFD, 2nd ed., DCW Industries, 1998.
[9] J. Sarmiento, N. Gruber, in: Ocean Biogeochemical Dynamics, Princeton University Press, 2006.
[10] W. Fennel, T. Neumann, in: Introduction to the Modelling of Marine Ecosystems, Elsevier, 2004.
[11] H.T. Banks, K. Kunisch, in: Estimation Techniques for Distributed Parameter Systems, Birkhäuser, 1989.
[12] H.M. Bucker, O. Fortmeier, M. Petera, Solving a parameter estimation problem in a three-dimensional conical tube on a parallel and distributed software infrastructure, Journal of Computational Science 2 (2011) 95–104, Simulation Software for Supercomputers.
[13] A. Tarantola, in: Inverse Problem Theory and Methods for Model Parameter Estimation, SIAM, 2005.
[14] A. Oschlies, V. Garcon, An eddy-permitting coupled physical-biological model of the north Atlantic. 1. sensitivity to advection numerics and mixed layer physics, Global Biogeochem. Cytochem. 13 (1999) 135–160.
[15] J. Rückelt, V. Sauerland, T. Slawig, A. Srivastav, B. Ward, C. Patvardhan, Parameter optimization and uncertainty analysis in a model of oceanic CO_2-uptake using a hybrid algorithm and algorithmic differentiation, Nonlinear Anal.: Real World Appl. Online (2010).
[16] P. Kunkel, V. Mehrmann, in: Differential-algebraic Equations: Analysis and Numerical Solution, EMS, 2006.
[17] R. Temam, Navier–Stokes Equations, North-Holland, Amsterdam, 1979.
[18] T. Simpson, J. Poplinski, P.N. Koch, J. Allen, Metamodels for computer-based engineering design: survey and recommendations, Eng. Comput. 17 (2001) 129–150, 10.1007/PL00007198.
[19] A.J. Smola, B. Schölkopf, A tutorial on support vector regression, Stat. Comput. 14 (2004) 199–222, doi:10.1023/B:STCO.0000035301.49549.88.
[20] J. Søndergaard, Optimization using surrogate models – by the space mapping technique, Ph.D. thesis, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, 2003, Supervisor: Kaj Madsen.
[21] A.R. Conn, N.I.M. Gould, P.L. Toint, in: Trust-region Methods, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2000.
[22] S. Koziel, J.W. Bandler, Q.S. Cheng, Robust trust-region space-mapping algorithms for microwave design optimization, IEEE Trans. Microw. Theory 58 (2010) 2166–2174.
[23] W. Hackbusch, in: Elliptic Differential Equations: Theory and Numerical Treatment, Springer Series in Computational Mathematics, Springer, Berlin, 2010.
[24] C.A.J. Fletcher, in: Computational Techniques for Fluid Dynamics, vol. 1, 2nd ed., Springer, 1991.
[25] M. Priess, S. Koziel, T. Slawig, Improved surrogate-based optimization of climate model parameters using response correction, in: International Conference on Simulation and Modeling Methodologies, Technologies and Application, SIMULTECH 2011, Noordwijkerhout, The Netherlands, 2011, pp. 449–457.

M. Prieß received his diploma degree in physics with foucs on theoretical physics, mathematics and computer science from the Leipniz Universität Hannover, Germany, in 2007. He is currently a Ph.D. student in the research group Algorithmic Optimal Control – CO_2 Uptake of the Ocean (headed by Prof. Thomas Slawig) at the Christian-Albrechts-Universität zu Kiel in the Institute for Computer Science. His research interest lie in the field of parameter optimization (e.g. for climate models), surrogate-based optimization, space mapping, algorithms for nonlinear optimization and numerical analysis.

S. Koziel received the M.Sc. and Ph.D. degrees in electronic engineering from Gdansk University of Technology, Poland, in 1995 and 2000, respectively. He also received the M.Sc. degrees in theoretical physics and in mathematics, in 2000 and 2002, respectively, as well as the Ph.D. in mathematics in 2003, from the University of Gdansk, Poland. He is currently an Associate Professor with the School of Science and Engineering, Reykjavik University, Iceland. His research interests include CAD and modeling of microwave circuits, surrogate-based optimization, space mapping, circuit theory, analog signal processing, evolutionary computation and numerical analysis.

T. Slawig is professor at the Christian-Albrechts-Universität zu Kiel in the Institute for Computer Science. He is leader of the research group Algorithmic Optimal Control – CO_2 Uptake of the Ocean in the Cluster Future Ocean. His research interest lies in the field of optimal control for partial differential equations, optimization and algorithmic differentiation.

Marine Ecosystem Model Calibration through Enhanced Surrogate-Based Optimization

Malte Prieß[1], Slawomir Koziel[2], and Thomas Slawig[1]

[1] Institute for Computer Science, Cluster The Future Ocean,
Christian-Albrechts Universität zu Kiel, 24098 Kiel, Germany,
{mpr,ts}@informatik.uni-kiel.de
[2] Engineering Optimization & Modeling Center, School of Science and Engineering,
Reykjavik University, Menntavegur 1, 101 Reykjavik, Iceland,
koziel@ru.is

Abstract. Mathematical optimization of models based on simulations usually requires a substantial number of computationally expensive model evaluations and it is therefore often impractical. An improved surrogate-based optimization methodology, which addresses these issues, is developed for the optimization of a representative of the class of one-dimensional marine ecosystem models. Our technique is based upon a multiplicative response correction technique to create a computationally cheap but yet reasonably accurate surrogate from a temporarily coarser discretized physics-based coarse model. The original version of this methodology was capable of yielding about 84% computational cost savings when compared to the fine ecosystem model optimization. Here, we demonstrate that by employing relatively simple modifications, the surrogate model accuracy and the efficiency of the optimization process can be further improved. More specifically, for the considered test case, the optimization cost is reduced three times, i.e., from about 15% to only 5% of the cost of the direct fine model optimization.

Keywords: Marine Ecosystem Models, Surrogate-Based Optimization, Parameter Optimization, Response Correction, Data Assimilation

1 Introduction

Numerical simulations nowadays play an important role to simulate the earth's climate system and to forecast its future behavior. The processes to be modeled and simulated are ranging from fluid mechanics (in atmosphere and oceans) to bio- and biochemical interactions, e.g., in marine or other type of ecosystems. The underlying models are typically given as time-dependent partial differential or differential algebraic equations [7, 10, 12].

Among them, marine ecosystem models describe photosynthesis and other biogeochemical processes in the marine ecosystem that are important, e.g., to compute and predict the oceanic uptake of carbon dioxide (CO_2) as part of the global carbon cycle [17]. They are typically coupled to ocean circulation models.

Since many important processes are non-linear, the numerical effort to simulate the whole or parts of such a coupled system with a satisfying accuracy and resolution is quite high.

There are processes in the climate system where where even without much simplification (through e.g. "parametrizations" to reduce the system size, see for example [12]) several quantities or parameters are unknown or very difficult to measure. This is for example the case for growth and dying rates in marine ecosystem models [5, 17], one of which our work in this paper is based on. Before a transient simulation of a model (e.g., used for predictions) is possible, the latter has to be calibrated, i.e., relevant parameters have to be identified using measurement data (sometimes also known as data assimilation). For this purpose, large-scale optimization methods become crucial for a climate system forecast.

The aim of parameter optimization is to adjust or identify the model parameters such that the model response fits given measurement data. The mathematical task thus can be classified as a least-squares type optimization or inverse problem [2, 3, 21]. This optimization (or calibration) process requires a substantial number of function and optionally sensitivity or even Hessian matrix evaluations. Evaluation times for the high-fidelity model of several hours, days or even weeks are not uncommon. As a consequence, optimization and control problems are often still beyond the capability of modern numerical algorithms and computer power. For such problems, where the optimization of coupled marine ecosystem models is a representative example, development of faster methods that would reduce the number of expensive simulations necessary to yield a satisfactory solution becomes critical.

Computationally efficient optimization of expensive simulation models (*high-fidelity* or fine models) can be realized using surrogate-based optimization (SBO), see for example [1, 6, 9, 15]. The idea of SBO is to exploit a surrogate, a computationally cheap and yet reasonably accurate representation of the high-fidelity model. The surrogate replaces the original high-fidelity model in the optimization process in the sense of providing predictions of the model optimum. Also, it is updated using the high-fidelity model data accumulated during the process. The prediction-updating scheme is normally iterated in order to refine the search and to locate the high-fidelity model optimum as precisely as possible. One of possible ways of creating the surrogate, our work in this paper is based on, is to utilize a physics-based *low-fidelity* (or coarse) model. The development and use of low-fidelity models obtained by, e.g., coarser discretizations (in time and/or space) or by parametrizations is common in climate research [12], whereas their applications for surrogate-based parameter optimization in this area is new.

In [14], a surrogate-based methodology has been developed for the optimization of climate model parameters. As a case study, a selected representative of the class of one-dimensional marine ecosystem models was considered. Since biochemistry mainly happens locally in space and since the complexity of the biogeochemical processes included in this specific model is high, this model serves as a good test example for the applicability of surrogate-based optimization ap-

proaches. The technique described in [14] is based on a multiplicative response correction of a temporally coarser discretized physics-based low-fidelity model. It has been successfully applied and demonstrated to yield substantial computational cost savings of the optimization process when compared to a direct optimization of the high-fidelity model.

In this paper, we demonstrate that by employing simple modifications of the original response correction scheme, one can improve the surrogate's accuracy, as well as further reduce the computational cost of the optimization process. We verify our approach by using synthetic target data and by comparing the results of SBO with the improved surrogate to those obtained with the original one. The optimization cost is reduced three times when compared to previous results, i.e., from about 15% to only 5% of the cost of the direct high-fidelity ecosystem model optimization (used as a benchmark method). The corresponding time savings are increased to from 84% to 95%.

It should be emphasized that the proposed approach does not rely on high-fidelity model sensitivity data. As a consequence, the first-order consistency condition between the surrogate and the high-fidelity model (i.e., agreement of their derivatives) is not fully satisfied. Nevertheless, the combination of the knowledge about the marine system under consideration embedded in the low-fidelity model and the response correction is sufficient to obtain a quality solution in terms of good model calibration, i.e., its match with the target output.

The paper is organized as follows. The high-fidelity ecosystem model, considered here as a test problem, as well as the low-fidelity counterpart that we use as a basis to construct the surrogate model, are described in Section 2. The optimization problem under consideration is formulated in Section 3. The original and improved response correction schemes and the comparison of the corresponding surrogate model qualities are discussed in Section 4. Numerical results for an illustrative SBO run are provided in Section 5. Section 6 concludes the paper.

2 Model Description

The considered example for the class of one-dimensional marine ecosystem models simulates the interaction of dissolved inorganic nitrogen, phytoplankton, zooplankton and detritus (dead material), thus is of so-called $NPZD$ type [13]. The model uses pre-computed ocean circulation and temperature data from an ocean model (in a sometimes called *off-line mode*), i.e., no feedback by the biogeochemistry on the circulation and temperature is modeled, see again [13]. The original high-fidelity (fine) model and its low-fidelity (coarse) counterpart which we use as a basis to construct a surrogate for further use in the optimization process are briefly described below.

2.1 The High-Fidelity Model

The $NPZD$ model simulates one water column at a given horizontal position. This is motivated by the fact that there have been special time series studies at

fixed locations. Clearly, the computational effort in a one-dimensional simulation is significantly smaller than in the three-dimensional case. However, as pointed out in the introduction, the model – from point of view of the complexity of the included processes – serves as a good test example for the applicability of SBO approaches.

In the $NPZD$ model, the concentrations (in mmol N m^{-3}) of dissolved inorganic nitrogen N, phytoplankton P, zooplankton Z, and detritus (i.e., dead material) D are summarized in the vector $y = (y^{(l)})_{l=N,P,Z,D}$ and described by the following coupled PDE system

$$\begin{aligned} \frac{\partial y^{(l)}}{\partial t} &= \frac{\partial}{\partial z}\left(\kappa \frac{\partial y^{(l)}}{\partial z}\right) + Q^{(l)}(y, u_2, \ldots, u_n), \quad l = N, P, Z, \\ \frac{\partial y^{(D)}}{\partial t} &= \frac{\partial}{\partial z}\left(\kappa \frac{\partial y^{(D)}}{\partial z}\right) + Q^{(D)}(y, u_2, \ldots, u_n) - \frac{\partial y^{(D)}}{\partial z} u_1, \quad l = D, \end{aligned} \tag{1}$$

in $(-H, 0) \times (0, T)$, with additional appropriate initial values. Here, z denotes the only remaining, vertical spatial coordinate, and H the depth of the water column. The terms $Q^{(l)}$ are the biogeochemical coupling (or *source-minus-sink*) terms for the four tracers and $\mathbf{u} = (u_1, \ldots, u_n)$ is the vector of unknown physical and biological parameters, with $n = 12$ for this specific model. The sinking term (with the sinking velocity u_1) is only apparent in the equation for detritus. In the one-dimensional model no advection term is used, since a reduction to vertical advection would make no sense. Thus, the circulation data (taken from an ocean model) are the turbulent mixing coefficient $\kappa = \kappa(z, t)$ and the temperature $\Theta = \Theta(z, t)$, which goes into the nonlinear coupling terms $Q^{(l)}$ but is omitted in the notation.

The parameters $\mathbf{u}$ to be optimized are, for example, growth and dying rates of the tracers and thus appear in the nonlinear coupling terms $Q^{(l)}_{l=N,P,Z,D}$ in (1). For the sake of brevity and for the purpose of this paper we omit the explicit formulation of the coupling terms as well as the explicit physical meaning of the involved parameter. For details we refer the reader to [13, 16].

2.2 Numerical Solution

The continuous model (1) is discretized and solved using an operator splitting method [11], an explicit Euler time stepping scheme for the nonlinear coupling terms Q and the sinking term while using an implicit scheme for the diffusion term. For further details we refer the reader to [13, 14].

More explicitly, in every discrete time step, at first the nonlinear coupling operators Q_j (that depend on t_j directly and/or via the temperature field Θ) are computed at every spatial grid point and integrated by four explicit Euler steps with step size $\tau/4$. Then, an explicit Euler step with full step size τ is performed for the sinking term. Finally, an implicit Euler step for the diffusion operator, again with full step size τ, is applied.

In the original model, the time step τ is chosen as one hour. By choosing this time step, all relevant processes are captured and further decrease of the time

step does not improve the accuracy of the model. The model with this particular time step will be referred to as the high-fidelity or fine one in the following.

We furthermore denote by $\mathbf{y}_j \approx y(\cdot, t_j)$ the discrete fine model solution of the continuos model (1) in time step j (containing all tracers N, P, Z, D) given as

$$\mathbf{y}_j = (y_{ji})_{i=1,\dots,I}, \quad j = 1, \dots, M_f, \quad \mathbf{y} \in \mathbb{R}^{M_f I}, \quad I = n_z n_t, \tag{2}$$

where I denotes the number of spatial discrete points n_z times the number of tracers n_t, which is four for the considered model, and where M_f denotes the total number of discrete time steps, given the discrete time step τ_f. More specifically, the model consists of $n_z = 66$ vertical layers and is integrated over totally $M_f = 8760\,\text{time steps/year} \times 5\,\text{years} = 43800$ discrete time step. We will furthermore use the subscript f to distinguish the relevant fine model variables, which read $\mathbf{y}_f, \tau_f$ and M_f, from those we will introduce for the coarse model, respectively.

2.3 The Low-Fidelity Model

Marine ecosystem model, are typically given as coupled time-dependent partial differential equations, compare [5, 17]. One straightforward way to introduce a low-fidelity (or coarse) model for these models is to reduce the spatial and/or temporal resolution, whereas, in this paper, we exploit the latter one.

The coarse model, which is a less accurate but computationally cheap representation of $\mathbf{y}_f$ is obtained by using a coarser time discretization with a discrete time step τ_c given as

$$\tau_c = \beta \tau_f, \tag{3}$$

with a *coarsening factor* $\beta \in \mathbb{N} \setminus \{0, 1\}$, while keeping the spatial discretization fixed. The state variable for this coarser discretized model will be denoted by $\mathbf{y}_c$, the corresponding number of discrete time steps by $M_c = M_f/\beta$, i.e., we have

$$(\mathbf{y}_c)_j = ((y_c)_{ji})_{i=1,\dots,I}, \quad j = 1, \dots, M_c, \quad \mathbf{y}_c \in \mathbb{R}^{M_c I}, \quad I = n_z n_t\,. \tag{4}$$

Note that the parameters $\mathbf{u}$ for this model are the same as for the fine one.

Clearly, the choice of the temporal discretization, or equivalently, the coarsening factor β, determines the quality of the coarse model and of a surrogate if based upon the latter one. Moreover, both the computational cost, the performance and quality of the solution obtained by a SBO process might be affected.

Altogether, we seek for a reasonable trade-off between the accuracy and speed of the coarse model. From numerical experiments, a value of $\beta = 40$ turned out be a reasonable choice, as was shown in [14]. Numerical results presented in Section 4 demonstrate that such a coarse model leads to a reliable approximation of the original fine ecosystem model when a response correction technique as described in this paper is utilized. Furthermore, it was observed that, for this specific choice of β, while additionally restricting the parameter u_1, i.e., the sinking velocity, using an appropriate upper bound, the resulting model response does not show any numerical instabilities.

3 Optimization Problem

The task of parameter optimization in climate science typically is to minimize a least-squares type cost function measuring the misfit between the discrete model output $\mathbf{y} = \mathbf{y}(\mathbf{u})$ and given observational data $\mathbf{y}_d$ [2, 21]. In most cases, the problem is constrained by parameter bounds. The optimization problem can generally be written as

$$\min_{\mathbf{u} \in U_{ad}} J(\,\mathbf{y}(\mathbf{u})\,), \tag{5}$$

where

$$\begin{aligned} &J(\,\mathbf{y}\,) := ||\,\mathbf{y} - \mathbf{y}_d\,||^2, \\ &U_{ad} := \{\mathbf{u} \in \mathbb{R}^n : \mathbf{b}_l \leq \mathbf{u} \leq \mathbf{b}_u\}, \mathbf{b}_l, \mathbf{b}_u \in \mathbb{R}^n,\ \mathbf{b}_l < \mathbf{b}_u\,. \end{aligned} \tag{6}$$

The inequalities in the definition of the set U_{ad} of admissible parameters are meant component-wise. The functional J may additionally include a regularization term for the parameters. However, from numerical experiments, it turned out that such a term is not necessary to ensure a well performing optimization process.

Additional constraints on the state variable $\mathbf{y}$ might be necessary, e.g., to ensure non-negativity of the temperature or of the concentrations of biogeochemical quantities. In our example model, however, by using appropriate parameter bounds $\mathbf{b}_l$ and $\mathbf{b}_u$, non-negativity of the state variables can be ensured. This was already observed and used in [16].

4 Surrogate-Based Optimization

For many nonlinear optimization problems, a high computational cost of evaluating the objective function and its sensitivity, and, in some cases, the lack of sensitivity information, is a major bottleneck. The need for decreasing the computational cost of the optimization process is especially important while handling complex three-dimensional models.

Surrogate-based optimization [1, 6, 9, 15] is a methodology that addresses these issues by replacing the original high-fidelity or fine model $\mathbf{y}$ by a surrogate, in the following denoted by $\mathbf{s}$, a computationally cheap and yet reasonably accurate representation of $\mathbf{y}$.

Surrogates can be created by approximating sampled fine model data (*functional* surrogates). Popular techniques include polynomial regression, kriging, artificial neural networks and support vector regression [15, 18, 19]. Another possibility, exploited in this work, is to construct the surrogate model through appropriate correction/alignment of a low-fidelity or coarse model (*physics-based* surrogates) [20].

Physics-based surrogates inherit physical characteristics of the original fine model so that only a few fine model data is necessary to ensure their good

alignment with the fine model. Moreover, generalization capability of the physics-based models is typically much better than for functional ones. As a results, SBO schemes working with this type of surrogates normally require small number of fine model evaluations to yield a satisfactory solution. On the other hand, their transfer to other applications is less straightforward since the underlying coarse model and chosen correction approach is rather problem specific. The specific correction technique exploited in this work is recalled in Section 4.1 (see also [14]).

The surrogate model is updated at each iteration k of the optimization algorithm, typically using available fine model data from the current and/or also from previous iterates. The next iterate, $\mathbf{u}_{k+1}$, is obtained by optimizing the surrogate $\mathbf{s}_k$, i.e.,

$$\mathbf{u}_{k+1} = \underset{\mathbf{u}\in U_{ad}}{\operatorname{argmin}}\ J\left(\mathbf{s}_k(\mathbf{u})\right), \tag{7}$$

where, again U_{ad} denotes the set of admissible parameters. The updated surrogate $\mathbf{s}_{k+1}$ is determined by re-aligning the coarse model at $\mathbf{u}_{k+1}$ and optimized again as in (7). The process of aligning the coarse model to obtain the surrogate and subsequent optimization of this surrogate is repeated until a user-defined termination condition is satisfied, which can be based on certain convergence criteria, assumed level of cost function value or a specific number of iterations (particularly if the computational budget of the optimization process is limited).

If the surrogate $\mathbf{s}_k$ satisfies so-called zero-order and first-order consistency conditions with the fine model at $\mathbf{u}_k$, i.e.,

$$\mathbf{s}_k(\mathbf{u}_k) = \mathbf{y}_f(\mathbf{u}_k), \quad \mathbf{s}_k'(\mathbf{u}_k) = \mathbf{y}_f'(\mathbf{u}_k), \tag{8}$$

with $\mathbf{y}'$ and $\mathbf{s}_k'(\mathbf{u}_k)$ denote the derivatives of the responses, the surrogate-based scheme (7) is provable convergent to at least a local optimum of (5) under mild conditions regarding the coarse and fine model smoothness, and provided that the surrogate optimization scheme (7) is enhanced by the trust-region (TR) safeguard, i.e.,

$$\mathbf{u}_{k+1} = \underset{\substack{\mathbf{u}\in U_{ad},\\ \|\,\mathbf{u}-\mathbf{u}_k\,\|\le\delta_k}}{\operatorname{argmin}}\ J\left(\mathbf{s}_k(\mathbf{u})\right), \tag{9}$$

with δ_k being the trust-region radius updated according to the TR rules. We refer the reader to e.g. [4, 8] for more details.

4.1 Surrogate Model Using Basic Multiplicative Response Correction

It has been found in [14] that a natural way of constructing the surrogate would be *multiplicative response correction*. This approach is motivated by the fact that the qualitative relation of the fine and coarse model response is rather well preserved (at least locally) while moving from one parameter vector to another. As a result, a multiplicative correction allows constructing a surrogate model with a good generalization capability. The technique is briefly recalled below.

The surrogate response $\mathbf{s}_k(\mathbf{u})$, at iteration k of the optimization process, is generated by multiplicative correction of the *smoothed* coarse model response, denoted by $\tilde{\mathbf{y}}_c$, which we briefly formulate as

$$\left.\begin{aligned} \bar{\mathbf{s}}_k(\mathbf{u}) &:= \mathbf{a}_k\, \tilde{\mathbf{y}}_c(\mathbf{u}), \\ \mathbf{a}_k &:= \frac{\tilde{\mathbf{y}}_f^{\beta}(\mathbf{u}_k)}{\tilde{\mathbf{y}}_c(\mathbf{u}_k)} \end{aligned}\right\} \quad \begin{aligned} &k = 1, 2, \dots \\ &\beta = M_f/M_c \end{aligned} \tag{10}$$

where the operations in (10) are meant point-wise and where a_k denote the *correction factors* which are included in the vector $\mathbf{a}_k$. They are defined as the point-wise division of the smoothed and *down-sampled* fine model response, denoted by $\tilde{\mathbf{y}}_f^{\beta}$, by the smoothed coarse model response at the point $\mathbf{u}_k$.

It was observed that smoothing allows us to remove the numerical noise from the coarse model response and identify the main characteristics of the traces of interest (see [14] for details). The fine model response is smoothed accordingly in the formulation (10).

Down-sampling was necessary to make the fine model response commensurable with the corresponding response of the coarse model. The down-sampled fine model response $\mathbf{y}_f^{\beta}$ is simply given as

$$y_{ji}^{\beta} := y_{\beta j,i}, \quad j = 1, \dots, M_c, \quad i = 1, \dots, I\,. \tag{11}$$

By definition, the surrogate model is zero-order consistent with the (down-sampled and smoothed) fine model in the point $\mathbf{u}_k$, i.e.,

$$\mathbf{s}_k(\mathbf{u}_k) = \tilde{\mathbf{y}}_f^{\beta}(\mathbf{u}_k)\,. \tag{12}$$

As we do not use sensitivity information from the fine model, the first-order consistency condition cannot be satisfied exactly. Nevertheless, as was shown in [14], this surrogate model exhibits quite good generalization capability, which means that the surrogate provides a reasonable approximation of the fine one in the neighborhood of $\mathbf{u}_k$.

Figure 1 shows the surrogate's, fine (down-sampled) and coarse model responses $\mathbf{s}_k, \tilde{\mathbf{y}}_f^{\beta}$ and $\tilde{\mathbf{y}_c}$ at two different points, $\mathbf{u}_k$ and $\bar{\mathbf{u}}_k$. The surrogate model is established at $\mathbf{u}_k$ and, therefore, its response is perfectly aligned with the one of the fine model at $\mathbf{u}_k$, whereas its prediction is still reasonably accurate at $\bar{\mathbf{u}}_k$.

Note that only the selected tracers for a chosen section in the whole time interval and at one selected depth layer are shown. The total dimension of the model response is too large to present a full response here. We emphasize that shown responses are representative for the overall qualitative behavior the other tracers, time sections and depth layers.

4.2 Difficulties of Basic Surrogate Formulation

Occasionally, when using the surrogate given in (10), there might occur a situation where the coarse model response is close to zero (and maybe even negative

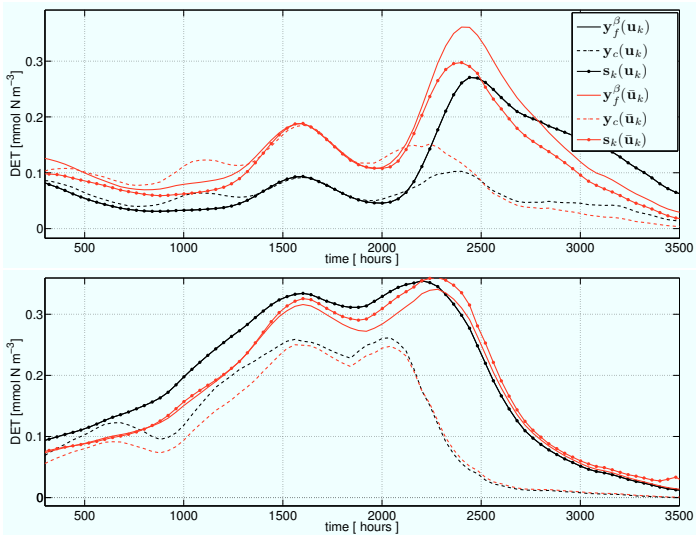


Fig. 1: Surrogate's, fine (down-sampled, smoothed) and coarse (smoothed) model responses $\mathbf{s}_k, \tilde{\mathbf{y}}_f^\beta$ and $\tilde{\mathbf{y}}_c$ for the tracer detritus at the uppermost depth layer at two points $\mathbf{u}_k$ and corresponding perturbation $\bar{\mathbf{u}}_k$, illustrating the generalization capability of the surrogate.

due to approximation errors) and a few magnitudes smaller than the fine one, which leads to large (possibly negative) correction factors a_k. While such a correction ensures zero-order consistency at the point where it was established (i.e., $\mathbf{u}_k$), it may lead to (locally) poor approximation in the vicinity of $\mathbf{u}_k$.

Figures 2 and 3 (top) illustrate these issues by showing the smoothed surrogate's, fine (down-sampled) and coarse model responses $\mathbf{s}_k, \tilde{\mathbf{y}}_f^\beta$ and $\tilde{\mathbf{y}}_c$ for the state detritus at one illustrative time interval and depth layer. Shown are the model responses at the same points $\mathbf{u}_k$ and its neighborhood $\bar{\mathbf{u}}_k \in B_\delta(\mathbf{u}_k)$ as in Figure 1.

It should be pointed out that the overall shape of the surrogate's response still provides a reasonable approximation of the fine model one (and more accurate than the corresponding coarse model response) despite of the distortion illustrated in Figures 2 and 3. This is supported by the fact that even without addressing these issues, the SBO was able to yield satisfactory results, not only with respect to the quality of the final solution, but, most importantly, in terms of the low computational cost of the optimization process. This was already demonstrated in [14].

4.3 Improved Response Correction Scheme

The response distortion described in the previous section is problematic towards the end of the surrogate-based optimization run when a higher accuracy of the

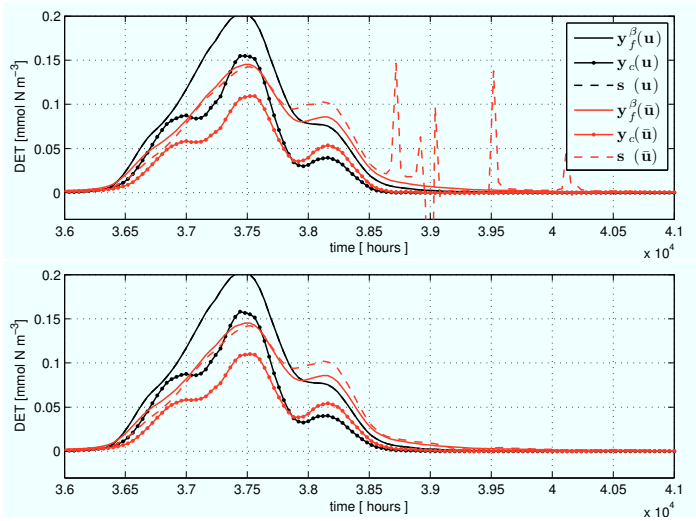


Fig. 2: Responses as in Figure 1 for a different time interval using the basic surrogate formulation (10) (top) and exploiting the modifications (13) of the response correction scheme (bottom).

surrogate is required to locate the fine model optimum more accurately. The "'spikes"' appearing in the response due to large values of the correction term can be viewed, in a way, as a numerical noise that slows down the algorithm convergence and makes the optimum more difficult to locate.

A few simple means described below can address these issues and further improve the accuracy of the surrogate's response as well as the performance of the optimization algorithm. We introduce non-negative bounds for the coarse model response (the negative response is non-physical and is a result of numerical errors due to using large time steps in the numerical solution of the coarse model) and an upper bound a_{ub} for the correction factors. We furthermore restrict the correction factors to one in case the fine and coarse model responses are below a certain threshold ϵ which should be of the order of the discretization error below which the responses can be treated as zero.

More specifically, the following modifications of the model outputs and the scaling factors are performed for each iteration k

$$(i)\ \mathbf{y}_c = \begin{cases} 0; & \text{if } \mathbf{y}_c \le 0 \\ \mathbf{y}_c; & \text{else} \end{cases}, \quad (ii)\ \mathbf{a}_k = \begin{cases} a_{\text{ub}}; & \text{if } \mathbf{a}_k \ge a_{\text{ub}} \\ \mathbf{a}_k; & \text{else} \end{cases}, \tag{13}$$

$$(iii)\ \mathbf{a}_k = 1 \text{ if } (\tilde{\mathbf{y}}_f^\beta \le \epsilon \text{ and } \tilde{\mathbf{y}}_c \le \epsilon),$$

where the operations are again meant point-wise and where (i) is applied before smoothing. From numerical experiments, $a_{\text{ub}} = 10$ turned out to be a reasonable choice and we furthermore consider $\epsilon = 10^{-4}$.

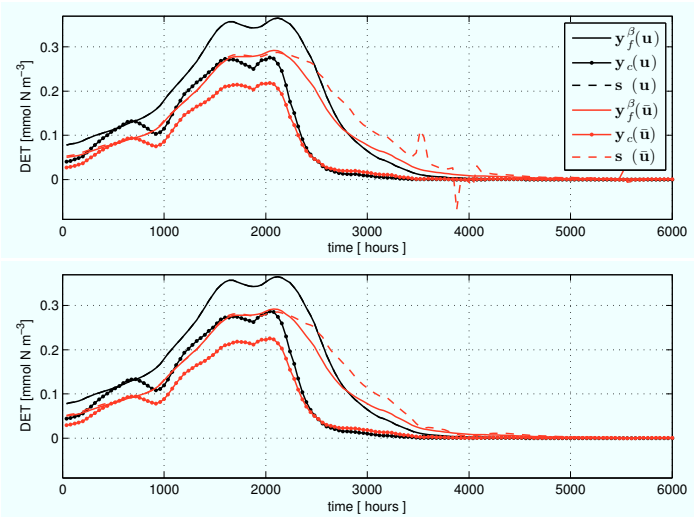


Fig. 3: Responses as in Figure 2, but for yet another section within the whole time interval. Again, after employing the improvements in (13), the positive and negative peaks are removed (bottom).

Figure 2 (bottom) shows the surrogate's, fine (down-sampled) and coarse model response for the same illustrative tracer, time interval and depth layer, however, while employing the improvements given in (13). It can be observed that the positive and negative peaks present in the surrogate responses shown in Figure 2 (top) are removed after applying (13). As additional evidence, Figure 3 (bottom) shows the same model responses but for a different section within the whole time interval.

The numerical results presented in Section 5 demonstrate that this enhanced response correction scheme allows us to further improve the computational efficiency of the SBO.

5 Numerical Results

For all optimization runs, we use the MATLAB[1] function `fmincon`, exploiting the active-set algorithm. The following cost functions

$$J(\mathbf{z}) := \| \mathbf{z} - \mathbf{y}_d \|^2 = \sum_{i=1}^{I} \sum_{j=1}^{M_c} \left(z_{ji} - (y_d)_{ji} \right)^2 , \tag{14}$$

$$\tilde{J}(\mathbf{z}) := \| \mathbf{z} - \tilde{\mathbf{y}}_d \|^2 = \sum_{i=1}^{I} \sum_{j=1}^{M_c} \left(z_{ji} - (\tilde{y}_d)_{ji} \right)^2 , \tag{15}$$

[1] MATLAB is a registered trademark of The MathWorks, Inc., http://www.mathworks.com

were the target data – as a test case – is given by model generated, attainable data as

$$\mathbf{y}_d := \mathbf{y}_f^\beta(\mathbf{u}_d)\,.$$

For the optimization runs presented in this paper we employ the following cost functions: for the fine model optimization, we use (14) with $\mathbf{z} = \mathbf{y}_f^\beta$, for the coarse model optimization, (15) with $\mathbf{z} = \tilde{\mathbf{y}}_c$ and for the SBO, (15) with $\mathbf{z} = \mathbf{s}_k$, whereas (14) was used in the termination condition and to compare the results and where the down-sampled fine model response $\mathbf{y}_f^\beta$ is defined by (11). Sampling was necessary to yield a comparable fine model optimization run while in (15) the smoothed target data is considered accordingly, since the coarse model and thus also the surrogate's response are smoothed. Note that the cost functions we employ are not normalized by the total number of discrete model points. The dimension of the responses is of the order of 10^5. Clearly, this has to be taken into account for presented cost function values in the following.

We perform an exemplary direct fine and coarse model optimization as well as a SBO based on the surrogate in (10) exploiting the original and improved response correction scheme (cf. Sections 4.1, 4.3). In the following, the solutions of the four optimization runs are compared through visual inspection of the (down-sampled) fine model response $\mathbf{y}_f^\beta$ and the corresponding cost function value $J(\mathbf{y}_f^\beta)$ (cf. (14)) at the respective optima.

The optimization cost is measured in *equivalent* fine model evaluations which are determined taking into account the coarsening factor β. More specifically, one evaluation of the coarse model with a coarsening factor β is equivalent to $1/\beta$ evaluations of the fine model. On the other hand, the cost of one iteration of the SBO (in terms of equivalent fine model evaluations) equals to the number of coarse model evaluations necessary to optimize the surrogate model divided by this factor β, and increased by the cost for the response correction. Recall that the specific correction (10) we use in this paper requires one fine model evaluation only.

Figure 4 shows the value of the cost function $J(\mathbf{y}_f^\beta)$ versus the equivalent number of fine model evaluations for the SBO algorithm using the surrogate model exploiting the original and the improved correction scheme, as well as for the fine and coarse model optimization. Points 1 and 3 in Figure 4 indicate those solutions obtained in the SBO runs that correspond to a termination condition of $J \leq 50$. This particular value was selected as it ensures good visual agreement between the fine model output and the target. Point 2 denotes the solution in the improved SBO run which could be obtained at the same optimization cost as the one at point 1 in the original SBO run.

Figure 5 shows the fine model response at the solutions $\mathbf{u}_{s1}^*$ and $\mathbf{u}_{s2}^*$ (corresponding to points 1 and 2 in Figure 4) obtained using the SBO algorithm with the original and improved response correction scheme (cf. Sections 4.1 and 4.3) as well the responses at the solutions $\mathbf{u}_f^*, \mathbf{u}_c^*$ of a direct fine and coarse model optimization. For illustration, responses for two representative tracers and for

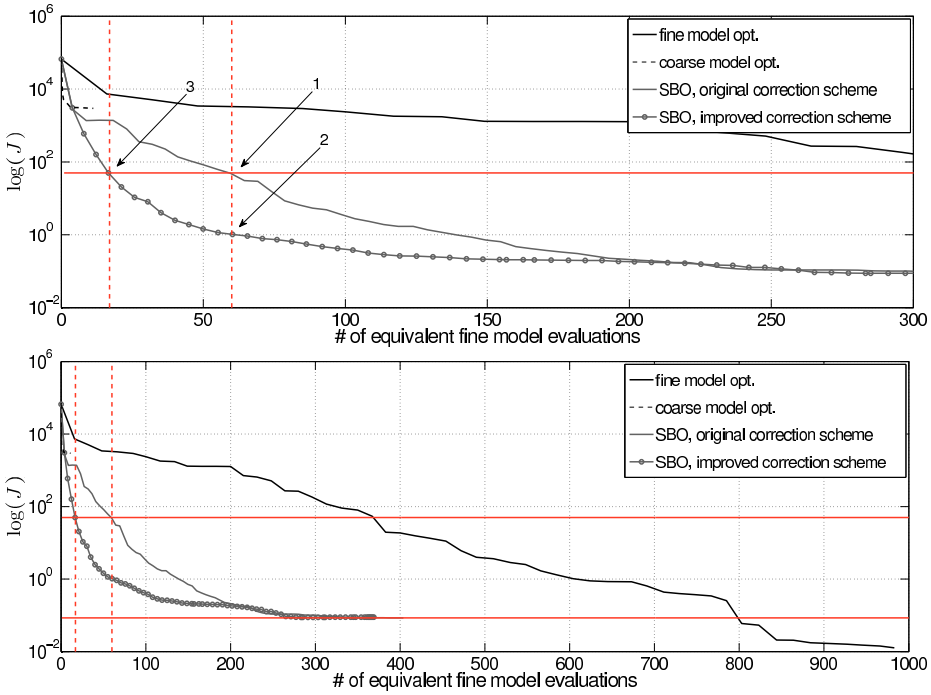


Fig. 4: Values of the cost function J versus the optimization cost measured in *equivalent number* of fine model evaluations for an exemplary SBO run exploiting the original and the improved correction scheme, as well as for a fine and coarse model optimization run. Points 1 and 3 correspond to a termination condition of $J \leq 50$ (upper horizontal line), ensuring good visual agreement between the fine model output and the target. Solution at point 2 in the improved SBO is significantly more accurate and obtained at the same cost as the one at point 1. Overall, SBO converges to a cost function value of to $J \approx 10^{-1}$ (lower horizontal line).

a selected depth level and time interval are shown. Corresponding parameter values are provided in Table 1.

It can be observed that coarse model optimization yields a solution far away from the target and a rather inaccurate parameter match (cf. Table 1), whereas the optimization cost of only 11 equivalent fine model evaluations is very low. However, results indicate that the accuracy of the coarse model is not sufficient to use this very model directly in an optimization.

On the other hand, direct fine model optimization yields a solution $\mathbf{u}_f^*$ with an almost perfect fit of the target data (cf. Figure 5) and of the optimal parameters $\mathbf{u}_d$ (cf. Table 1), corresponding to a very low cost function of $J \approx \cdot 10^{-2}$. However, the optimization cost is substantially higher: about 980 fine model evaluations.

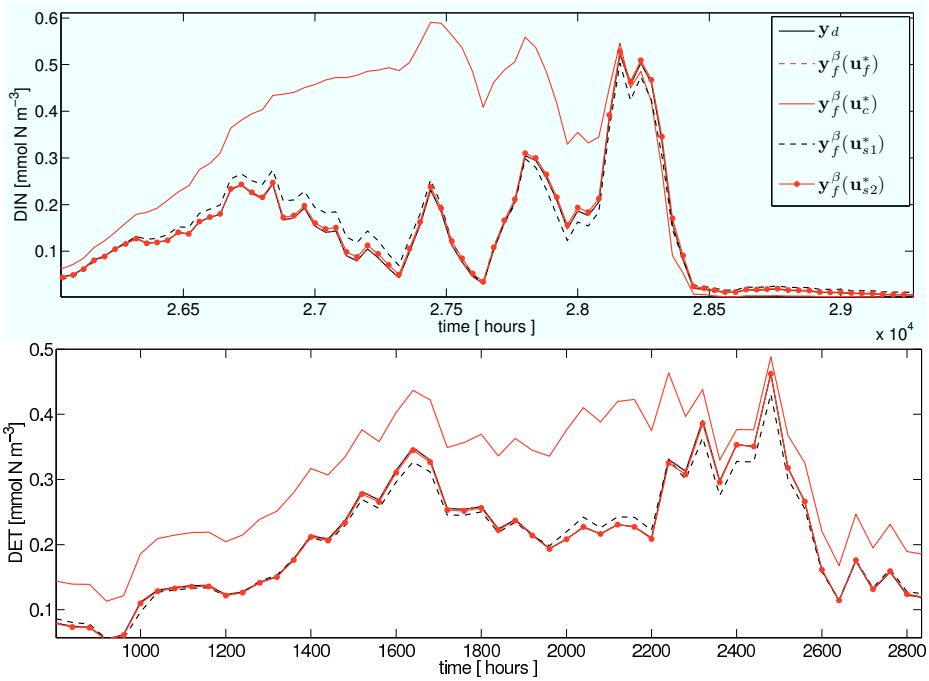


Fig. 5: Synthetic target data $\mathbf{y}_d$ at optimal parameters $\mathbf{u}_d$ and fine model response $\mathbf{y}_f^\beta$ (down-sampled) for two illustrative tracers and at the uppermost depth layer for the solutions $\mathbf{u}_f^*, \mathbf{u}_c^*, \mathbf{u}_{s1}^*$ and $\mathbf{u}_{s2}^*$ of a direct fine and coarse model optimization as well as of a SBO run exploiting the original and the improved correction scheme. Solutions $\mathbf{u}_{s1}^*$ and $\mathbf{u}_{s2}^*$ correspond to points 1 and 2 in Figure 4.

In [14], we demonstrated that in a exemplary SBO run based on the original response correction scheme, a reasonably accurate solution $\mathbf{u}_{s1}^*$ could be obtained at the cost of approximately 60 equivalent fine model evaluations only (point 1 in Figure 4). This resulted in a significant reduction of the total optimization cost of about 84% when compared to the direct fine model optimization (correspondingly, 375 evaluations were required in the fine model optimization to reach this cost function value, cf. Figure 4).

Exploiting the improved scheme, a similarly accurate solution – both in terms of parameter match and optimal fit of the target data – can be obtained at a remarkably lower cost of only 17 equivalent fine model evaluations (point 3 in Figure 4). This is over three times less than for the original response correction scheme corresponding to a reduction of the total optimization cost of about 96%. Specific parameter values and model responses of this solution are omitted here, since they are similar to those of the original solution $\mathbf{u}_{s1}^*$.

On the other hand, when exploiting the improved correction scheme, a solution $\mathbf{u}_{s2}^*$ (point 2 in Figure 4) with a significantly higher accuracy – again both

Table 1: Solutions $\mathbf{u}_c^*, \mathbf{u}_f^*, \mathbf{u}_{s1}^*$ and $\mathbf{u}_{s2}^*$ of an illustrative coarse, fine model optimization and of a SBO run, exploiting the original and the improved correction scheme. Solutions $\mathbf{u}_{s1}^*$ and $\mathbf{u}_{s2}^*$ correspond to points 1 and 2 in Figure 4.

iterate	u_1	u_2	...									u_{12}
	SBO (original and improved scheme)											
$\mathbf{u}_{s1}^*$	0.705	0.626	0.044	0.015	0.060	0.937	1.908	0.016	0.147	0.020	0.629	4.237
$\mathbf{u}_{s2}^*$	0.738	0.604	0.028	0.010	0.036	1.024	1.678	0.010	0.206	0.020	0.541	4.318
	Coarse model optimization											
$\mathbf{u}_c^*$	0.300	1.066	0.036	0.065	0.064	0.025	0.040	0.065	0.010	0.012	0.730	3.448
	Fine model optimization											
$\mathbf{u}_f^*$	0.747	0.596	0.025	0.010	0.030	0.999	2.046	0.010	0.203	0.020	0.493	4.310
$\mathbf{u}_d$	0.750	0.600	0.025	0.010	0.030	1.000	2.000	0.010	0.205	0.020	0.500	4.320

in terms of parameter match and optimal fit of the target data – can be obtained (cf. Figure 5 and Table 1) at the same cost as were required for the original one $\mathbf{u}_{s1}^*$, i.e., 60 equivalent fine model evaluations.

It should be emphasized that the surrogate model utilized in this work only satisfies zero-order consistency with the fine model. Still, as demonstrated in this section, the performance of our surrogate-based optimization process is satisfactory, particularly in terms of obtaining a good match between the model response and a given target output. Improved matching between the optimized model parameters and those corresponding to the target output could be obtained by executing larger number of SBO iterations (cf. Figure 4), which is mostly because of low sensitivity of the model with respect to some of the parameters. Also, the use of derivative information together with the trust-region convergence safeguards [4, 8] would bring further improvement in terms of matching accuracy. Clearly, the trade-offs between the accuracy of the solution and the extra computational overhead related to sensitivity calculation has to be investigated. The aforementioned issues will be the subject of future research.

6 Conclusions

Parameter identification in climate models can be computationally very expensive or even beyond the capabilities of modern computer power. Before a transient simulation of a model (e.g., used for predictions) is possible, the latter has to be calibrated, i.e., relevant parameters have to be identified using measurement data. This is the point where large-scale optimization methods become crucial for a climate system forecast.

Using the high-fidelity (or fine) model under consideration in conventional optimization algorithms that require large number of model evaluations is often

infeasible. Therefore, the development of faster methods that aim at reducing the optimization cost, such as surrogate-based optimization (SBO) techniques, are highly desirable. The idea of SBO is to replace the high-fidelity model in the optimization run by a surrogate, its computationally cheap and yet reasonably accurate representation.

As a case study, we have investigated parameter optimization of a representative of the class of one-dimensional marine ecosystem models. As demonstrated in our previous work, a simple multiplicative response correction applied to a temporally coarser discretized physics-based low-fidelity (coarse) model of the system of interest is sufficient to create a reliable surrogate of the original, high-fidelity ecosystem model, which can be used as a prediction tool to calibrate the latter. This approach allowed us to yield remarkably good results, both in terms of the quality of the final solution and, most importantly, in terms of the relative reduction in the total optimization cost, about 84% when compared to the direct fine model optimization.

In this paper, we demonstrated that the correction scheme can be enhanced to alleviate the difficulties of its original version, which results in further improvement of the surrogate model accuracy and overall performance of the optimization algorithm utilizing this surrogate. The optimization cost was reduced by a factor of three (from 16% to 5% of the direct high-fidelity model optimization optimization cost), which corresponds to the cost savings of 95%.

Improvements of the present approach by utilizing additionally sensitivity information of the low- and the high-fidelity model in the alignment of the low-fidelity model as well as trust-region convergence safeguards applied to enhance the optimization process are expected to further improve the robustness of the algorithm and the accuracy of the solution. The trade-offs between the accuracy and extra costs due too sensitivity evaluation will have to be inspected.

Acknowledgments The authors would like to thank Andreas Oschlies, IFM Geomar, Kiel. This research was supported by the DFG Cluster of Excellence Future Ocean.

Bibliography

[1] J.W. Bandler, 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.*, 52(1), 2004.

[2] H.T. Banks and K. Kunisch. *Estimation Techniques for Distributed Parameter Systems*. Birkhäuser, 1989.

[3] H.M. Bucker, O. Fortmeier, and M. Petera. Solving a parameter estimation problem in a three-dimensional conical tube on a parallel and distributed software infrastructure. *Journal of Computational Science*, 2(2):95–104, May 2011. Simulation Software for Supercomputers.

[4] A.R. Conn, N.I.M. Gould, and P.L. Toint. *Trust-region methods*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2000.

[5] W. Fennel and T. Neumann. *Introduction to the Modelling of Marine Ecosystems*. Elsevier, 2004.

[6] A.I.J. Forrester and A.J. Keane. Recent advances in surrogate-based optimization. *Prog. Aerosp. Sci.*, 45(1-3):50–79, 2009.

[7] A.E. Gill. *Atmosphere - Ocean Dynamics*, volume 30 of *International Geophysics Series*. Academic Press, 1982.

[8] S. Koziel, J.W. Bandler, and Q.S. Cheng. Robust trust-region space-mapping algorithms for microwave design optimization. *IEEE T. Microw. Theory.*, 58(8):2166–2174, August 2010.

[9] 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*, 1(2):98–106, 2010.

[10] A. Majda. *Introduction to PDE's and Waves for the Atmosphere and Ocean*. AMS, 2003.

[11] G.I. Marchuk. *Methods of Numerical Mathematics*. Springer, 2nd edition, 1982.

[12] K. McGuffie and A. Henderson-Sellers. *A Climate Modelling Primer*. Wiley, 3nd edition, 2005.

[13] 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.*, 13:135–160, 1999.

[14] M. Prieß, S. Koziel, and T. Slawig. Surrogate-based optimization of climate model parameters using response correction. *Journal of Computational Science*, 2011 (in press).

[15] N.V. Queipo, R.T. Haftka, W. Shyy, T. Goel, R. Vaidyanathan, and P.K. Tucker. Surrogate-based analysis and optimization. *Prog. Aerosp. Sci.*, 41(1):1–28, 2005.

[16] J. Rückelt, V. Sauerland, T. Slawig, A. Srivastav, B. Ward, and C. Patvardhan. Parameter optimization and uncertainty analysis in a model of oceanic CO_2-uptake using a hybrid algorithm and algorithmic differentia-

tion. *Nonlinear Analysis B Real World Applications*, 10(1016):3993–4009, 2010.

[17] J. L. Sarmiento and N. Gruber. *Ocean Biogeochemical Dynamics*. Princeton University Press, 2006.

[18] T.W. Simpson, J.D. Poplinski, P.N. Koch, and J.K. Allen. Metamodels for computer-based engineering design: Survey and recommendations. *Eng. Comput.*, 17:129–150, 2001. 10.1007/PL00007198.

[19] A.J. Smola and B. Schölkopf. A tutorial on support vector regression. *Stat. Comput.*, 14:199–222, 2004. 10.1023/B:STCO.0000035301.49549.88.

[20] J. Søndergaard. *Optimization using surrogate models - by the space mapping technique*. PhD thesis, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, 2003. Supervisor: Kaj Madsen.

[21] A. Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM, 2005.

Low-Cost Marine Ecosystem Model Calibration Using Surrogate-Based Optimization and Trust Regions

M. Prieß[a,1,*], S. Koziel[c], T. Slawig[a]

[a] *Institute for Computer Science, Cluster The Future Ocean, Christian-Albrechts Universität zu Kiel, 24098 Kiel, Germany*
[b] *Leibniz Institute of Marine Science (IFM-GEOMAR), Marine Biogeochemistry, Biological Oceanography, Düsternbrooker Weg 20, 24105 Kiel, Germany*
[c] *Engineering Optimization & Modeling Center, School of Science and Engineering, Reykjavik University, Menntavegur 1, 101 Reykjavik, Iceland*

Abstract

Model calibration plays a key role in simulations and predictions of the earth's climate system. Calibration is normally formulated as an inverse problem where a set of control parameters are to be found so that the model fits given measurement data. Straightforward calibration attempts by direct adjustment of the high-fidelity (or fine) model parameters using conventional optimization algorithms are often tedious or even infeasible as they normally require a large number of simulations. The development of faster methods becomes critical, particularly for the models that are computationally expensive. The optimization of coupled marine ecosystem models simulating biogeochemical processes in the ocean is here a representative example. In this paper, we introduce a surrogate-based optimization (SBO) methodology where the calibration of the expensive fine model is carried out by means of a surrogate: its fast and yet reasonably accurate representation. As a case study, we consider a representative of the class of one-dimensional marine ecosystem models. The surrogate is obtained from a temporarily coarser discretized physics-based low-fidelity (or coarse) model and a multiplicative response correction technique. In our previous work, a basic formulation of this surrogate was sufficient to create a reliable approximation, yielding a remarkably accurate solution at low computational costs. This was verified by model-generated attainable data. The application on real (measurement) data is covered in this paper. Enhancements of the basic formulation by utilizing additionally fine and coarse model sensitivity information as well as trust-region convergence safeguards allow us to further improve the robustness of the algorithm and the accuracy of the solution. The trade-offs between the solution accuracy and the extra computational overhead related to sensitivity calculation is also addressed. We demonstrate that SBO is able to yield a very accurate solution at nevertheless low computational cost. The time savings are up to 85 percent when compared to the direct fine model optimization.

Keywords: Climate models, marine ecosystem models, surrogate-based optimization, parameter optimization, parameter identification, response correction, computationally efficient optimization

1. Introduction

Numerical simulations play a key role to simulate and predict processes in the earth's climate system, ranging from fluid dynamics (in atmosphere and oceans), thermodynamics, radiative transfer to bio- and biochemical interactions, e.g., in marine or other type of ecosystems. The underlying models are typically formulated as time-dependent partial differential or differential algebraic equations (PDEs/DAEs) [1, 2, 3].

Since many important processes are non-linear, the numerical effort to simulate the whole or parts of the climate system with a satisfying accuracy and resolution is quite high. This motivates the development and use of reduced order models by e.g. coarser discretizations (in time and/or space) or by parametrizations to reduce the system size and thus the computational effort [1]. Through those parametrizations, several additional parameters enter the system. Many of them are not known beforehand and not directly measurable.

Growth and dying rates in marine ecosystem models [4, 5], one of which is taken as a test case for the proposed optimization methodology, are examples for such unknown parameters. Marine ecosystem models describe the transport, interactions and biogeochemistry among ocean biota. The modeled processes comprise the marine biogeochemical cycles among carbon and the major nutrients – therein the marine carbon cycle [see, e.g., 4, 5, 6]. Marine ecosystem models are of great importance for understanding the oceanic uptake of carbon dioxide and for projections of the marine ecosystem's responses to climate change.

Generally, before a transient simulation is possible, a marine ecosystem model has to be calibrated, i.e., the relevant parameters have to be identified such that the simulated tracer concentrations ideally resembles the actual physical and biogeochemical processes. Moreover, the ability to forecast future dynamics within the marine ecosystem crucially depends upon parameterizations of the desired biogeochemical processes. Thus, since

*Corresponding author (*phone:* +49-(0)431 880 7452, *fax:* +49-(0)431 880 7618)
Email addresses: mpr@informatik.uni-kiel.de (M. Prieß), koziel@ru.is (S. Koziel), ts@informatik.uni-kiel.de (T. Slawig)
[1] Research supported by DFG Cluster The Future Ocean

there is no general consensus on what is *the* "correct" ecosystem model or model structure to represent the observed quantities under consideration, an assessment of the different models/parametrizations highly depends on their validation against the given observed quantities. Mathematically, this parameter identification can be classified as a least-squares type optimization or inverse problem (see, e.g., [7, 8, 9, 10]). This optimization (or calibration) process requires a substantial number of (typically expensive) function and optionally sensitivity or even Hessian matrix evaluations.

Straightforward attempts by employing the *high-fidelity* or *fine* model under consideration directly in an optimization loop using conventional optimization techniques are therefore tedious or even beyond the capability of modern computer power, especially when using traditional, gradient-based techniques. The need for an accelerated optimization process, which especially becomes important while handling complex three-dimensional models, becomes critical.

Surrogate-based optimization (SBO) addresses this issue by shifting the computational burden from the accurate and expensive high-fidelity model to its fast but yet reasonably accurate surrogate. More specifically, the idea of SBO is to replace the fine model in the optimization process in the sense of providing predictions of the model optimum. The surrogate can be created by approximating sampled fine model data (so-called *function-approximation surrogates*, see [11, 12, 13]) or by employing a so-called physics-based *low-fidelity* or *coarse* model, a computationally cheap representation of the fine one. The latter approach is used in this paper. Since the accuracy of the coarse model is usually not sufficient to directly use the latter in lieu of the fine model in an optimization loop, it is often necessary to use suitable alignment/correction techniques to reduce the misalignment between the coarse and fine model responses. Popular correction/alignment techniques include response correction [14] and space mapping [15]. Surrogate-based optimization is widely and very successfully used in engineering sciences (see, e.g., [11, 15, 16, 17].

As a case study, in order to investigate the applicability of a SBO methodology to the optimization of marine ecosystem models, we consider a representative of the class of one-dimensional models. Clearly, the computational effort in a one-dimensional simulation is significantly smaller than in the three-dimensional case. However, biochemistry mainly happens locally in space and, moreover, the complexity of the response of this specific model is comparably high. Thus, although one-dimensional, this model serves as a suitable and computational affordable test example, before considering computationally more expensive three-dimensional models.

Exhaustive optimization runs by using both local, gradient-based and global, genetic algorithms have been previously performed for this specific model (see, e.g., [18, 19, 20]). However, it is not the focus of this paper to further assess the quality of the optimal solution obtained there. For the purpose of this paper we tentatively accept the previously found minima. Also, we don't seek a quantitative interpretation of the solutions obtained by SBO in the biogeochemical context. Our aim clearly is to demonstrate the applicability of the proposed methodology to the parameter optimization of the considered model. More specifically, the focus is to demonstrate that, by exemplary optimization runs, SBO is able to yield a solution close to the one obtained by a direct fine model optimization at low optimization costs.

One straightforward way to introduce a physics-based coarse model is to reduce the spatial and/or temporal resolution, whereas the latter is used for the selected model in this paper. Moreover, we use a multiplicative response correction technique for the alignment of the coarse and fine model response.

In our previous work [21], a basic formulation of this surrogate was sufficient to create a reliable approximation, yielding a remarkably accurate solution at low computational costs. This was verified by model generated, attainable data.

In this paper, the application on real data is covered. Utilizing additionally fine and coarse model sensitivity information ensures the zero- and first-order consistency conditions between the fine model and the surrogate, i.e., agreement in function values and first-order derivatives. Embedding the algorithm in a trust-region framework [22, 23], this allows us to further improve the robustness of the SBO and accuracy of its solution. The trade-offs between the solution accuracy and the extra computational overhead related to sensitivity calculation will be addressed. We show the results of an exemplary SBO run and compare the solution to those obtained by a direct fine and coarse model optimization. We demonstrate that a direct optimization of the fine model requires a substantial number of comparably expensive fine model evaluations, whereas a direct coarse model optimization is computationally cheap but yields a rather inaccurate solution only. We subsequently show that SBO yields a solution close to the one obtained by a direct fine model optimization while greatly reducing the optimization costs – down to 15% of those of a direct fine model optimization.

The structure of the paper is as follows: We briefly recall the general form of numerical models common in climate science and highlight the special properties of marine ecosystem models in Section 2. We introduce the basic idea of surrogate-based optimization in Section 3. The ecosystem model and corresponding optimization problem, which is taken as an example in this paper, is introduced in Section 4. The coarse model that we use as a basis to create a surrogate, is recalled in Section 5. The response correction approach used to obtain the surrogate is motivated and described in Section 6. The optimization setup, numerical results and discussion of exemplary optimization runs are provided in Sections 7 and 8. Section 9 concludes the paper with a summary and an outlook.

2. Climate Models – A General Formulation

Numerical models that are used to simulate processes in the climate system (what we denote by *climate models*) can be quite generally written as coupled systems of time-dependent partial differential or differential algebraic equations (PDEs/DAEs) [1,

2, 3], for example in the following form:

$$\begin{aligned} E\,\frac{\partial y}{\partial t} &= f\left(y, \frac{\partial y}{\partial x_i}, \frac{\partial^2 y}{\partial x_i \partial x_j}, u\right) && \text{in } I \times \Omega \\ y(t_0, x) &= y_{init} && \text{in } \Omega \\ By &= 0 && \text{on } I \times \Gamma, \end{aligned} \tag{1}$$

where $y(t, x) : I \times \Omega \to \mathbb{R}$ is the vector of the *state variable*, with a definite time interval $I = [t_0, t_0 + T]$, $t_0 \in \mathbb{R}$ an initial point in time, $T \in \mathbb{R}$ a duration, $\Omega \in \mathbb{R}^3$ a domain and where $\Gamma = \partial\Omega$ denotes its boundary. The time variable is denoted by $t \in I$ and the spacial variable by $x = (x_1, x_2, x_3)^\top \in \overline{\Omega}$. We use a boldfaced notation to distinguish a vector from a continuous or scalar variable in the following.

The right-hand side f includes all spatial differential operators as well as the coupling between the components of the state variable y. In climate models, it often additionally depends explicitly on the space and time variables x and t, respectively, which, for simplicity, is omitted in the notation. Moreover, f depends on a number of model parameters which are summarized in the vector u. The vector-valued function $y_{init} : \Omega \to \mathbb{R}$ includes the initial model data and B denotes the boundary operator which – when representing for example a Neumann boundary condition – is nonlinear and includes the first normal derivative.

E is a matrix with the size of y, typically being the identity matrix for a PDE while having rank deficiency for a PDAE [24]. We include PDAEs in this formulation since for example in ocean circulation models [3], the underlying Navier-Stokes equations are – when written in the above form – a PDAE system. Then y may for example consist of the velocity, pressure, temperature or salinity field. In the case of marine ecosystem models, which are formulated as a PDE system, the matrix E can be set to the identity and thus omitted. In this case, the state vector y contains so-called *biogeochemical tracers* such as phytoplankton, see Section 2.1 below and 4 for details.

2.1. Marine Ecosystem Models

Marine ecosystem models mainly consist of two parts, namely the ocean circulation and the biogeochemical model [see, e.g., 4, 5, 25]. The coupling between ocean circulation and the biogeochemical interactions such as photosynthesis is mostly regarded as a one-way coupling. This means that the influence of the biota on the circulation (including temperature and maybe salinity distribution) is assumed to be negligible and thus is often omitted (so-called *off-line* mode). See for example [26] where such an off-line computation has been thoroughly described and investigated for an atmospheric model. Velocity and temperature fields are computed beforehand by an ocean circulation model and only used as *forcing data* for the biogeochemical simulations which significantly reduces the computational effort (see, e.g., [27]). Our example model (cf. Section 4) is simulated in such an off-line mode.

The model equations consist of a system of coupled advection-diffusion-reaction equations, where the reaction terms (also called *source minus sink*, or *sms* terms) are given by the biogeochemical interactions between the biogeochemical tracers. As a special form of (1), a system of these *transport equations* for n_t tracers then generally reads

$$\frac{\partial y_i}{\partial t} = \operatorname{div}(\kappa \nabla y_i) - \operatorname{div}(v y_i) + q_i(y, \mathbf{u}), \quad i = 1, \ldots, n_t \tag{2}$$

where $y_i(t, x) : I \times \Omega \to \mathbb{R}$ denotes the concentration of tracer i at time t and the spatial location x. If no interactions with the atmosphere is taken into account, homogeneous Neumann conditions on the boundary Γ for all concentrations are employed, i.e.,

$$\frac{\partial y_i}{\partial n} = n \cdot \nabla y = 0 \quad \text{on } I \times \Gamma, \quad i = 1, \ldots, n_t, \tag{3}$$

where n denotes the normal vector. The time dependent turbulent mixing/diffusion coefficient $\kappa(t, x) : I \times \Omega \to \mathbb{R}$ as well as the velocity vector field $v(t, x) : I \times \Omega \to \mathbb{R}^3$ with $v = (v_i)_{i=1,2,3}$, both satisfy the Navier-Stokes equations. Since, here, the parameters $\mathbf{u} \in \mathbb{R}^{n_p}$, which are subject to the parameter optimization, are scalar coefficients in the nonlinear biogeochemical coupling terms q_i, we use a boldfaced notation.

3. Surrogate-Based Optimization

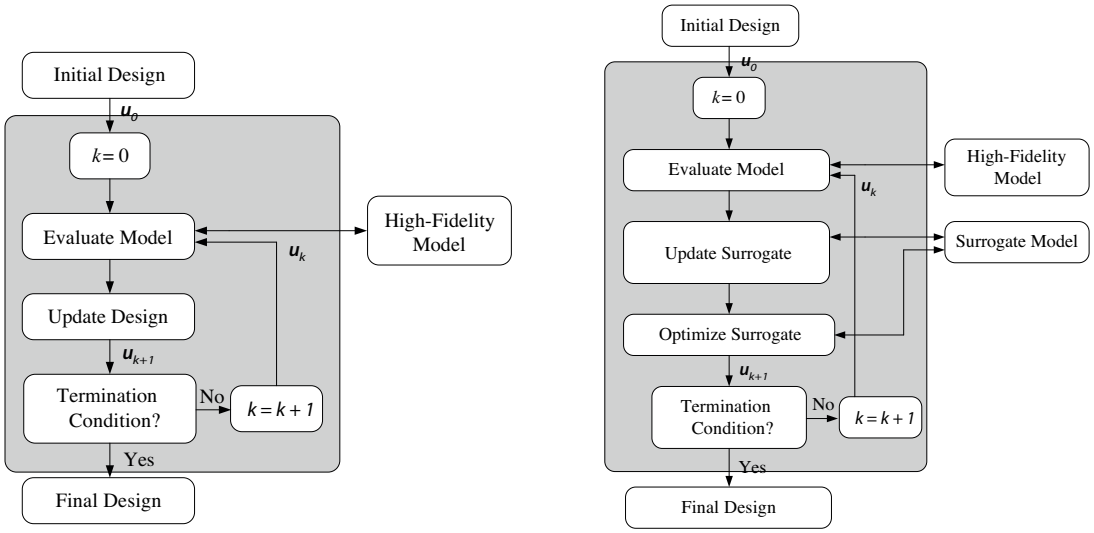
The optimization task is typically formulated as the minimization problem of the form

$$\min_{\mathbf{u}} J(\mathbf{y}(\mathbf{u})) \quad \text{s.t. constraints}, \tag{4}$$

where J denotes a cost function measuring the misfit between relevant quantities (which are obtained from the discrete model response $\mathbf{y}$ at the parameters/design $\mathbf{u}$) and some desired specifications. For the considered optimization problem in this paper, these quantities are tracer concentrations, whereas the desired specifications are corresponding observed quantities (cf. Section 4.4). However, for the purpose of this section, to sketch the basic ideas of SBO, we omit a more detailed formulation of J here. The state $\mathbf{y}$ is normally evaluated through a computationally expensive numerical simulation and will be referred to as the high-fidelity or fine model in the following.

For many optimization problems, a high computational cost and/or even the lack of sensitivity information of the model under consideration is a major bottleneck. As a result, straightforward attempts of solving (4) by employing the fine model under consideration directly in an optimization loop (cf. Figure 1a) using conventional optimization algorithms are often tedious or even infeasible, since typically a large number of the expensive fine model evaluations are required. The need for an accelerated optimization process becomes critical, for which the optimization of complex marine ecosystem models is a representative example.

Surrogate-based optimization (SBO) [11, 15, 16, 17] addresses these issues by replacing the original fine model in the optimization loop by its computationally cheaper but yet reasonably accurate surrogate (cf. Figure 1b). In particular, the surrogate at the iterate $\mathbf{u}_k$, in the following denoted by $\mathbf{s}_k(\mathbf{u})$, is



(a) Direct Optimization. (b) Surrogate-Based Optimization.

Figure 1: In a direct optimization (a), the complex high-fidelity or fine model under consideration is directly employed in an optimization loop using conventional optimization approaches. In a surrogate-based approach (b), the fine model is replaced in the optimization loop in iteration k by its computationally cheaper but yet reasonably accurate surrogate. Here, $\mathbf{u}_k$ denotes the parameter vector at iteration k.

constructed typically using available fine model data from the current and possibly also from previous iterates $(\mathbf{u}_i)_{i=0,\ldots,k-1}$.

Possible ways to create a surrogate are through approximations of sampled fine model data (cf. Section 3.1) or by correction/alignment of a less accurate but computationally cheaper low-fidelity (or coarse) model (cf. Section 3.2).

The next iterate, $\mathbf{u}_{k+1}$, in a typical SBO scheme is obtained by optimizing the surrogate $\mathbf{s}_k$, i.e.,

$$\mathbf{u}_{k+1} = \underset{\mathbf{u}}{\operatorname{argmin}}\, J(\,\mathbf{s}_k(\mathbf{u})\,) \quad \text{s.t. constraints,} \tag{5}$$

where J is the cost function, typically the same as in (4). The process of updating the surrogate and subsequent optimization is repeated until the user-defined termination condition is satisfied, which can be based on suitable convergence criteria, assumed level of cost function value or a specific number of iterations (particularly if the computational budget of the optimization process is limited).

A well performing surrogate-based algorithm is capable of yielding a reasonably accurate solution at a low computational cost, typically corresponding to only a few evaluations of the fine model. Key prerequisites to ensure this, are a cheap and yet reasonably accurate coarse model as well as a properly selected and low-cost alignment procedure (i.e., exploiting a limited number of fine model evaluations, preferably just one).

3.1. Functional Surrogates

One possibility, which will not be addressed further in this paper, is to create the surrogate by approximating sampled fine model data using suitable techniques, e.g., polynomial regression [11], kriging [12] or support-vector regression [13]. Since these so-called *function-approximation surrogates* are constructed without any particular knowledge of the system they are easily transferable to other application areas. On the other hand, such surrogates do not inherit any physical information about the fine model under consideration and normally require substantial amount of fine model data samples to ensure good accuracy so that their use to ad-hoc optimization may be questionable.

3.2. Physics-Based Surrogates

Another possibility, explored in this paper, is to construct the surrogate from a physics-based *low-fidelity* or *coarse* model, a usually computationally much cheaper but – on the other hand – less accurate representation of the fine one. This type of model are referred to as *physics-based surrogates*. Since the accuracy of the coarse model is typically not sufficient to directly directly replace the fine model in an optimization loop, it is then necessary to use suitable alignment/correction techniques to reduce the misalignment between the coarse and fine model responses and to ensure that the corrected model (the surrogate) provides a reliable prediction of the fine model optimum.

There are several methods of constructing the surrogate from a physics-based low-fidelity model. They include, among others, space mapping (SM) [15], various response correction techniques [14], manifold mapping [28], and shape-preserving response prediction (SPRP) [29]. The selection of the appropriate response correction technique is usually problem-specific.

Physics-based surrogates inherit relevant physical characteristics of the original fine model so that only a limited amount of fine model data is necessary to ensure sufficient accuracy. Also, generalization capability of the physics-based models is typically much better than that of the functional ones. As a results, SBO schemes working with physics-based surrogates normally require small number of fine model evaluations to yield a satisfactory solution.

The low-fidelity (or coarse) model can be created in various ways. The simple and straightforward approaches include the use of a coarser discretization in time and/or space (while employing the same simulation tool as for the fine model), simplified physics or different ways of describing the same physical phenomenon or even by using analytical formulas if available.

The surrogate, we use in this paper is physics-based. The specific coarse model is obtained by a coarser time discretization (cf. Section 5) which is further aligned by using a multiplicative response correction (cf. Section 6).

3.3. Consistency Conditions and Convergence of SBO

Provided that the surrogate $\mathbf{s}_k$ satisfies so-called zero- and first-order consistency conditions with the original fine model $\mathbf{y}_f(\mathbf{u}_k)$ at the iterate $\mathbf{u}_k$, i.e., agreement between the function values and first-order derivatives at the current iteration point, mathematically written as

$$\mathbf{s}_k(\mathbf{u}_k) = \mathbf{y}(\mathbf{u}_k), \quad \mathbf{s}'_k(\mathbf{u}_k) = \mathbf{y}'(\mathbf{u}_k), \tag{6}$$

the surrogate-based scheme (5) is provably convergent to at least a local optimum of (4) under mild conditions regarding the coarse and fine model smoothness (see, e.g., [30]), and provided that the surrogate optimization scheme is enhanced by the trust-region (TR) safeguard, i.e.,

$$\mathbf{u}_{k+1} = \underset{\|\mathbf{u}-\mathbf{u}_k\| \le \delta_k}{\operatorname{argmin}} J(\mathbf{s}_k(\mathbf{u})) \quad \text{s.t. constraints}, \tag{7}$$

with δ_k being the trust-region radius updated according to the TR rules. We refer the reader to e.g. [22, 23] for more details.

In (6), $\mathbf{y}'$ (and $\mathbf{s}'_k$) denotes the derivatives of the fine model and surrogate's response w.r.t. the parameter vector $\mathbf{u}$ and at the point $\mathbf{u}_k$, i.e., generally given as

$$\mathbf{y}'(\mathbf{u}_k) := \left.\frac{\mathrm{d}\,\mathbf{y}}{\mathrm{d}\,\mathbf{u}}\right|_{\mathbf{u}=\mathbf{u}_k}. \tag{8}$$

In practice, exact sensitivity information may not be obtainable, e.g., if the derivatives are calculated using finite differentiation. In such cases, the consistency conditions (6) only hold approximately. While this may not be sufficient for "theoretical" convergence, the use of trust-region and even approximate sensitivity substantially improve the SBO algorithm performance (see, e.g., [23])

The surrogate in this paper uses both fine model sensitivity information as well as trust-region convergence safeguards to increase the robustness of the optimization procedure and the accuracy of the solution obtained by SBO.

4. Example: A Marine Ecosystem Model

The model developed by Oschlies and Garçon [31] is a coupled system of four tracers with dissolved inorganic nitrogen (N), phytoplankton (P), zooplankton (Z), and detritus (D), thus also called $NPZD$ model, in the following summarized in the tracer or state vector $y = (y_i)_{i=1,\dots,n_t}$ with $n_t = 4$.

The $NPZD$ model simulates the tracer concentrations in one water column at a given horizontal position. This is motivated by the fact that there have been special time series studies at fixed locations [18]. Clearly, the computational effort in a one-dimensional simulation is significantly smaller than in the three-dimensional case. However, since biochemistry mainly happens locally in space and since the complexity of response of this specific model is high, this model serves as a good test example for the applicability of SBO approaches.

The model basically fits into our general framework (2). In the specific $NPZD$ model considered here, no advection term "div(vy_i)" as in (2) is used, since a reduction to vertical advection would make no sense. Starting from a general *continuous* formulation, the model is governed by the equations

$$\frac{\partial y_i}{\partial t} = \partial_z(\kappa\,\partial_z y_i) + q_i(y, \mathbf{u}), \quad i = 1, \dots, 4, \tag{9}$$

where, as in (2), $y_i(t, z) : I \times \Omega \to \mathbb{R}$ (with the domain $\Omega \in \mathbb{R}$) denotes the concentration of tracer i at time t and the vertical spatial location z. The coupling terms $q_i(y, \mathbf{u})$ are explicitly given as

$$\begin{aligned}
q_1(y,u) &= \Phi^z_m\, y_3 + \gamma_m\, y_4 - J(y_1, y_2, t, z)\, y_2,\\
q_2(y,u) &= J(y_1, y_2, t, z) y_2 - G(y_2, \epsilon, g)\, y_3 - \Phi^p_m\, y_2,\\
q_3(y,u) &= \beta\, G(y_2, \epsilon, g)\, y_3 - \Phi^z_m\, y_3 - \Phi^*_z\, (y_3)^2,\\
q_4(y,u) &= (1-\beta)\, G(y_2, \epsilon, g)\, y_3 + \Phi^p_m\, y_2 + \Phi^*_z\, (y_3)^2\\
&\qquad - \gamma_m\, y_4 - w_s\, \partial_z y_4.
\end{aligned} \tag{10}$$

The system involves an explicit *sinking velocity* w_s for the tracer detritus, and a non-differentiability, namely in the *growth rate of phytoplankton*, which is modeled after the minimum principle of von Liebig [32] as

$$J(y_1, y_2, t, z) = \min\left\{\bar{\mu}(y_2, t, z), V_p \cdot u(y_1, t, z)\right\}, \tag{11}$$

where the analytical solution for the *light-limited growth rate*, denoted as $\bar{\mu}(y_2, t, z)$, is given according to Evans and Parslow [33], integrated down to the given depth z [31, 19]. Here, additional parameters α, k_w and κ are involved (cf. Table 1).

The *factor for nutrient limited growth of phytoplankton* u and the *maximal phytoplankton growth rate* V_p are given as

$$u(y_1, t, z) = \frac{y_1}{k_N + y_1}, \quad V_p = \mu_m \cdot (C_{ref})^{c\,\Theta(t,z)}, \tag{12}$$

where the parameters k_N, C_{ref} and c are briefly described in Table 1 and where V_p further depends on the water temperature Θ, which has to be provided by an ocean circulation model. Due to the minimum in the growth rate of phytoplankton in (11), the

Table 1: Model parameters (cf. Section 4). Those included in the parameter vector $\mathbf{u} = (\mathbf{u})_{i=1,\ldots,12}$ are subject to the optimization.

u_i	symbol	value/range	unit (d=86400 s)	parameter meaning
	C_{ref}	1.066	1	growth coefficient
	c	1	$^\circ C^{-1}$	growth coefficient
	R	6.625	1	molar carbon to nitrogen ratio (*Redfield ratio*)
	k_w	25	m^{-1}	PAR extinction length
u_1	β	[0, 1]	1	assimilation efficiency of zooplankton
u_2	μ_m	$\mathbb{R}_0^+$	d^{-1}	phytoplankton growth rate parameter
u_3	α	$\mathbb{R}_0^+$	$m^2W^{-1}d^{-1}$	slope of photosynthesis versus light intensity
u_4	Φ_m^z	$\mathbb{R}_0^+$	d^{-1}	zooplankton loss rate
u_5	κ	$\mathbb{R}_0^+$	$m^2(mmol\ N)^{-1}$	light attenuation by phytoplankton
u_6	ϵ	$\mathbb{R}_0^+$	$m^6(mmol\ N)^{-2}d^{-1}$	grazing encounter rate
u_7	g	$\mathbb{R}_0^+$	d^{-1}	maximum grazing rate
u_8	Φ_m^p	$\mathbb{R}_0^+$	d^{-1}	phytoplankton linear mortality
u_9	Φ_z^*	$\mathbb{R}_0^+$	$m^3(mmol\ N)^{-1}d^{-1}$	zooplankton quadratic mortality
u_{10}	γ_m	$\mathbb{R}_0^+$	d^{-1}	detritus remineralization rate
u_{11}	k_N	$\mathbb{R}_0^+$	$mmol\ Nm^{-3}$	half saturation for NO_3 uptake
u_{12}	w_s	$\mathbb{R}_0^+$	$m\ d^{-1}$	detritus sinking velocity

model becomes non-differentiable. Another non-linear term in the equations is the *zooplankton grazing function* G given as

$$G(y_2, \epsilon, g) = \frac{g\,\epsilon\,(y_2)^2}{g + \epsilon\,(y_2)^2}, \tag{13}$$

which describes the transfer from phytoplankton to zooplankton and detritus with the parameters ϵ and g again briefly described in Table 1. There are totally twelve model parameters subject to the optimization, which are all summarized in Table 1. For the purpose of this paper, to demonstrate the applicability of the proposed SBO approach, we don't omit more details on the model and the involved parameters and refer the reader to [31, 18] for a more thorough description.

4.1. Carbon Primary Production

In addition to the tracers N, P, Z and D, the so-called *carbon fixation* or *carbon primary production* measured as carbon uptake (denotes as CUP in the following) is additionally taken into account in the optimization process for this model [18, 19] (see also Section 4.4). For a given depth z and time t, it can be briefly formulated as

$$y_5 := J(y_1, y_2, t, z) \cdot y_2(t, z) \cdot R$$

where R denotes the *Redfield ratio*, see, e.g., [34] and [4, Section 4.2]. It depends non-linearly on the states y_1 and y_2, i.e., the tracers dissolved inorganic nitrogen (N) and phytoplankton (P). It states that the relation between carbon (C), nitrogen (N) and phosphorus (P) in marine phytoplankton is given as $C : N : P = 106 : 16 : 1$. Thus, N can be used as a model state from which the potential uptake of CO_2 can be estimated (assuming that there is no limit on phosphorus P and carbon dioxide CO_2 in the water).

The carbon primary production obeys a daily cycle (cf. Figure 3a), since the growth of phytoplankton, $J(y_1, y_2, t, z)$, is light limited due to the term $\bar{\mu}(y_2, t, z)$ in (11) (see, e.g., [18] for details). The state CUP is calculated "internally" in the model simulation and provided as an additional state y_5 of the full model response y.

4.2. Numerical Solution

In an off-line coupled marine ecosystem model (as for example the *NPZD* model considered here), there are two ways to make use of the precomputed ocean circulation data. One way is to employ the ocean model that precomputes the data to generate so-called *transport matrices*, see [27]. These matrices usually represent a mean ocean circulation field for one month. Another approach, which the *NPZD* model is based on, is that the ocean model data is stored directly and afterwards used for assembling the system matrices for the differential operators for adjective and diffusive tracer transport in the marine ecosystem model itself.

For the numerical simulation, one may consider a spin-up into a steady quasi-periodic or periodic seasonal cycle, thus applying some kind of fixed point iteration. Another way, which is employed in the *NPZD* model considered in this paper, is to perform a complete transient run with time-dependent forcing data (as for example the temperature) to obtain a solution of (9).

More specifically, the time discretization is performed by a sequential integration at the discrete time steps $0 = t_0 < \ldots < t_j < \ldots < t_{n_\tau - 1} = T$ using a time step $\tau := t_j - t_{j-1}$ and with totally n_τ steps. A typical integration time is 5 years (see below, Section 4.3, for the details). This integration is partially implicit. An explicit Euler time-stepping scheme for the non-linear coupling terms q_i and the sinking term for

the tracer detritus is used while using an implicit Euler time-stepping scheme for the diffusion term. Furthermore, an operator splitting method is used. For details we refer the reader to [21, 31, 35].

Assuming n_z and n_τ discrete spatial and temporal grid points, with a time step $\tau = T/n_\tau$ and using again a boldfaced notation for discrete vectors, we denote by

$$\mathbf{y}_i = (\mathbf{y}_{ijk})_{\substack{j=1,\dots,n_\tau \\ k=1,\dots,n_z}} \tag{14}$$

the approximate solution of (9), i.e., $\mathbf{y}_{ijk} \approx y_i(t_j, z_k)$, denoting the concentration of tracer i at the discrete time step j and vertical depth layer k. The four state vectors for the tracers dissolved inorganic nitrogen (N), phytoplankton (P), zooplankton (Z) and detritus (D) as well as the state for the additional carbon primary production (CUP) will be summarized in the discrete vector $\mathbf{y} = (\mathbf{y}_i)_{i=1,\dots,5}$ in the following.

4.3. High-Fidelity Model

In the *original* discrete model, the time step τ is chosen as one hour. By choosing this time step all relevant processes are captured and further decrease of the time step does not improve the accuracy of the model. The number of vertical depth layers n_z is 66 and the number of discrete time steps n_τ is 43800, corresponding to 5 years with 8760 steps per year.

From now on, we will refer to this model and corresponding discrete solution as the *original* high-fidelity or fine model and will denote its state variable, time step and number of overall discrete time steps, to be distinguishable from the coarse model, by $\mathbf{y}_f$, τ_f and $n_{\tau,f}$, respectively.

4.4. Fine Model Optimization Problem

Our work is based on the following fine model optimization problem, where extensive optimization runs with different methods including local, gradient-based and also global, genetic algorithms have already been performed (see, e.g., [18, 19, 20]). It consists of finding optimal parameters yielding a minimal misfit of the discrete model response $\mathbf{y}_f$ to measurement data $\mathbf{y}_d$ as defined by the least-squares type cost function

$$\underset{\mathbf{u}\in U_{ad}}{\operatorname{argmin}}\ J_1(\mathbf{y}_f(\mathbf{u})) \tag{15}$$

where

$$\begin{aligned} J_1(\mathbf{y}_f) &:= \| C_1\,\mathbf{y}_f - \mathbf{y}_d \|_\sigma^2, \\ U_{ad} &:= \{\mathbf{u} \in \mathbb{R}^{n_p} : \mathbf{b}_l \le \mathbf{u} \le \mathbf{b}_u\}, \mathbf{b}_l, \mathbf{b}_u \in \mathbb{R}^{n_p}, \mathbf{b}_l < \mathbf{b}_u. \end{aligned} \tag{16}$$

More specifically, the measurement data $\mathbf{y}_d$ is considered for the years 1991-1995 and is taken from the *Bermuda Atlantic Time-Series Study*, called BATS, located at $31°N$, $64°W$ [18]. The inequalities in (16) in the definition of the set U_{ad} of admissible parameters are meant component-wise. The parameters $\mathbf{u}$ are the unknown scalar coefficients in the non-linear biogeochemical coupling terms q_i in (9). The specific parameter bounds $\mathbf{b}_u, \mathbf{b}_l$ that we employ in the optimization runs in this paper are provided in Table 3.

Furthermore, we have $n_p = 12$ model parameters subject to optimization (cf. Table 1) and the norm is weighted by assumed standard deviations of the measurements, $\sigma = (\sigma_j)_{j=1\dots,5}$ (see [18, 19] for details).

The functional J_1 may additionally include a regularization term for the parameters. However, regularization turns out not to be necessary to yield sufficient performance of the methods employed in this paper, and, it is therefore not used. Additional constraints on the state variable $\mathbf{y}_f$ might be necessary, e.g., to ensure non-negativity of the tracer concentrations. In our example model, this is ensured by using appropriate parameter bounds $\mathbf{b}_l$ and $\mathbf{b}_u$. This was already observed and used in [19].

C_1 is an operator describing transformations of the high-fidelity model response $\mathbf{y}_f$, to make it commensurable with the given measurement data $\mathbf{y}_d$. In brief, this operator includes the following transformation:

- A linear transformation to *chlorophyll a* (denoted as CHL) as a function of phytoplankton P, using a constant conversion factor.
- A linear transformation to *particulate organic nitrogen* (denoted as PON), calculated as the sum of phytoplankton P, zooplankton Z and detritus D.
- A spatial average of model response if the considered measurement data point lies in between two adjacent spatial grid cells.
- For zooplankton, a vertically averaged concentration in the water column down to the given depth of the measurement point (which is approximately 200 meters) is calculated.
- The observed zooplankton (with state $(\mathbf{y}_d)_3$) is furthermore transformed to $(\mathbf{y}_d)_3 = 1.23 \cdot (\mathbf{y}_d)_3 + 0.097$ in order to attempt an estimate of the total zooplankton from the measured mesozooplankton biomass (for the sake of simplicity, this is omitted in our cost function formulation (16))
- A constant temporal alignment of the model response is employed to make it commensurable with the measurement data point in time.
- A 24-hourly temporal mean of the modeled carbon primary production CUP is calculated to make it commensurable with observations from 24-hourly incubation measurements.

Except for zooplankton, only the data in the so-called *euphotic zone* – equivalent to the upper 20 discrete vertical depth layers in the model – is considered. For the sake of simplicity, we omit a more detailed description and mathematical formulation of these transformations and refer the reader to [18, 19, 20].

The measurement data in (16) is consequently given as $\mathbf{y}_d = (\mathbf{y}_d)_{i=1,\dots,5}$, with $(\mathbf{y}_d)_i$ denoting the measurement state corresponding to the transformed model response $C_1\,\mathbf{y}_f$, i.e., corresponding to the concentration of dissolved inorganic nitrogen (N), of chlorophyll a (CHL), of the total zooplankton biomass (ZOO), of particulate organic nitrogen (PON) and of

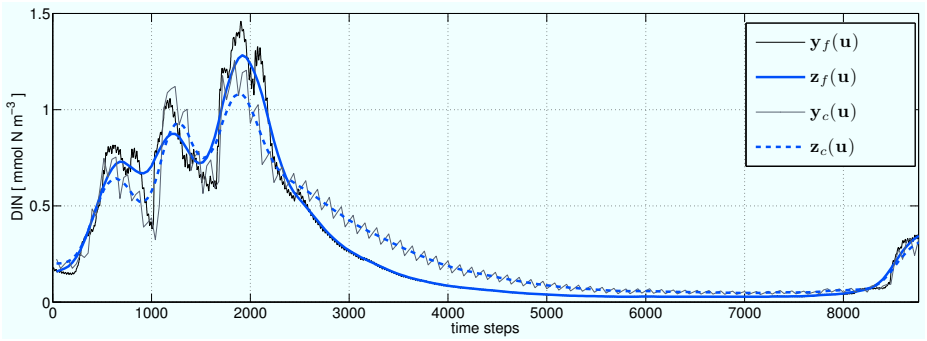


Figure 2: Fine and coarse model responses, both "raw" ($\mathbf{y}_f, \mathbf{y}_c$) and "smoothed" ($\mathbf{z}_f, \mathbf{z}_c$), employing a discrete time step of $\tau_f = 1$ h for the fine and $\tau_c = 40$ h for the coarse model, respectively (cf. Section 5). Shown is the response for one illustrative tracer (here, N), at some depth layer, part of the whole time interval and some illustrative parameter vector $\mathbf{u}$.

24-hourly incubation measurements of the carbon primary production (CUP), respectively.

From now on, we will refer to this optimization problem and corresponding cost function formulation as the *original* fine model one.

5. The Low-Fidelity Model

The way we follow here to obtain a physics-based low-fidelity (or coarse) model for the time-dependent marine ecosystem model introduced in Section 4, is to employ a coarser temporal discretization (see also Section 3.2). This has already been investigated in [21] and is briefly recalled below.

The coarse model is based upon the same model equations (9) and (10), whereas for its numerical solution (cf. Section 4.2), a larger time step, in the following denoted as τ_c, is employed with

$$\tau_c = \beta \cdot \tau_f \,. \tag{17}$$

We call $\beta \in \mathbb{N} \setminus \{0, 1\}$ the *coarsening factor* and $\tau_f = 1$ h denotes the time step employed in the original fine model solution. The spatial discretization of the coarse model is the same as for the original fine one (cf. Section 4.3). The sequential integration for the coarse model (cf. Section 4.2) is thus performed over $n_{\tau,c} = n_{\tau,f}/\beta$ discrete time steps, with $n_{\tau,f}$ denoting the total number of discrete time steps employed in the original fine model solution. In the following, the state variable of the coarse model will be denoted by $\mathbf{y}_c$, respectively.

Clearly, the choice of the temporal discretization, or equivalently, the coarsening factor β, determines the quality of the coarse model and, accordingly, of a surrogate if created from this model. Moreover, both the computational cost, the performance and quality of the solution obtained by a SBO process might be affected. Overall, we seek for a reasonable trade-off between the accuracy and speed of the coarse model. This has already been investigated in [21], where a value of $\beta = 40$ turned out to be a reasonable choice.

Furthermore, it could be demonstrated that, for the given sequential integration approach used to solve for a discrete solution of the model equations (cf. Section 4.2), a numerically stable solution (see, e.g., [36]) can be obtained if additionally restricting the parameter u_{12}, i.e., the sinking velocity (cf. parameter w_s in (10)), by using an appropriate upper bound. More specifically, from visual inspection of the model responses and from various optimization experiments, it turned out that, for the chosen coarsening factor of $\beta = 40$, $(b_u)_{12} = 5$ (cf. Table 3) ensures that the resulting coarse model response does not contain any numerical instabilities which would influence the optimization performance.

Given the temporal discretization with $\beta = 40$, we obtain for the discrete coarse model response $n_{\tau,c} = n_{\tau,f}/\beta = 43800/40 = 1095$ discrete time steps whereas the spatial vertical discretization with $n_z = 66$ is kept fixed, as noted before.

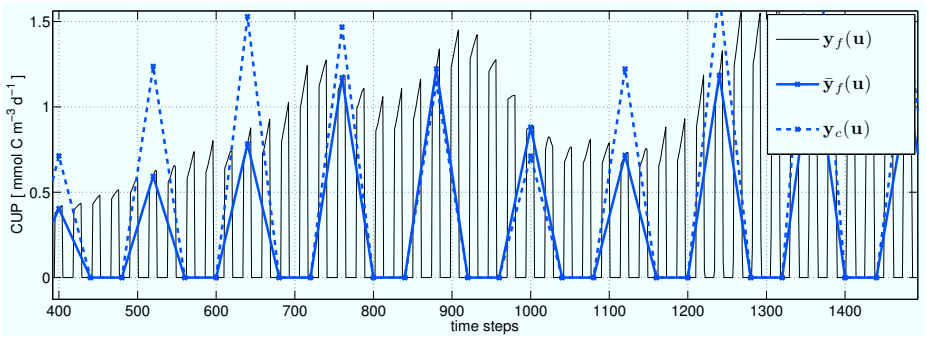
6. The Surrogate

The surrogate is obtained by a multiplicative response correction approach. It turned out that this multiplicative way of correcting the coarse model response is quite suitable for the considered problem because the overall "shape" of the coarse model response resembles that of the fine one and the relation between the coarse and the fine model responses is rather well preserved while moving from one parameter vector to another. This technique has already been investigated in [21]. We briefly recall the key ideas and describe modifications employed for the considered optimization problem with real measurement data below.

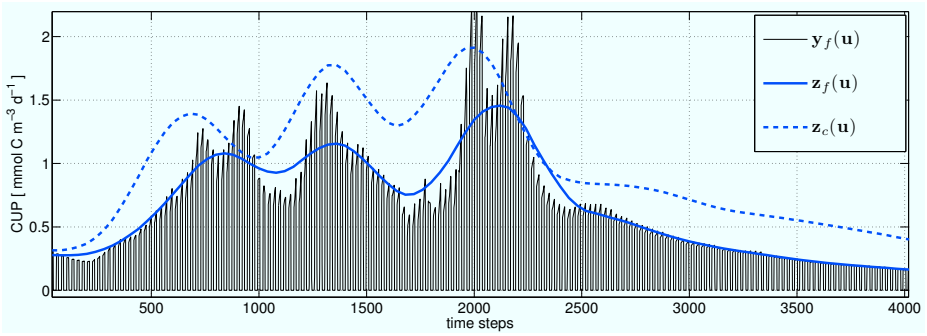
6.1. Smoothing

Due to the larger time step employed in the numerical solution of the coarse model (cf. Section 5), its response is rather inaccurate. Also both, the fine and coarse model responses, contain numerical noise which is misleading while performing model alignment. It has been demonstrated in [21] that adequate alignment of the coarse and fine model should be based on the main characteristics of the responses, which can be extracted by "smoothing".

The smoothed fine and coarse model response for the four tracers N, P, Z and D is obtained by applying a "'smoothing"'



(a) Original fine, down-sampled fine and coarse model response for the state CUP.



(b) Integration and smoothing of the coarse and of the down-sampled fine model response.

Figure 3: Fine (original and down-sampled) and coarse model response, $\mathbf{y}_f$, $\bar{\mathbf{y}}_f$ and $\mathbf{y}_c$ (Figure (a)) for the carbon primary production (CUP), at some illustrative depth layer (here, the uppermost) and some parameter vector $\mathbf{u}$. Figure (b) shows the original fine model response $\mathbf{y}_f$ as well as the (down-sampled) fine and coarse model responses after a rectangular integration and smoothing have been applied, yielding $\mathbf{z}_f$ and $\mathbf{z}_c$.

operator S as follows

$$(\mathbf{z}_f)_i := S(\bar{\mathbf{y}}_f)_i, \quad (\mathbf{z}_c)_i := S(\mathbf{y}_c)_i, \quad i = 1, \ldots, 4 \tag{18}$$

where we consider the *down-sampled* fine model response $\bar{\mathbf{y}}_f$ given by

$$\begin{aligned} &\bar{\mathbf{y}}_f := G\mathbf{y}_f, \quad (\bar{y}_f)_{i,j,k} := (y_f)_{i,\beta j,i}, \\ &i = 1, \ldots 5, \quad j = 1, \ldots, n_{\tau,c}, \quad i = 1, \ldots, n_z \end{aligned} \tag{19}$$

to be commensurable with the corresponding coarse model response. The down-sampling in (19) is employed for all 4 tracers as well as for the additional state CUP (cf. Section 4.1).

For the smoothing, we use a *walking average* with span $\pm n$, where a value of $n = 3$ and "double" smoothing turned out to be suitable for the considered coarse model (see again [21] for details). Figure 2 shows the fine and coarse model responses, both "raw" and smoothed, for the chosen temporal discretization with a coarsening factor of $\beta = 40$, for one illustrative tracer (here, N), some depth layer, part of the whole time interval and for some parameter vector $\mathbf{u}$.

Note that Figure 2 shows one selected tracer for some illustrative section in the whole time interval and at one selected depth layer. The total number of depth layers is 66 and the entire discrete time scale is 43800 so that it is impossible to present a full model response here. We emphasize that shown responses are representative for the overall qualitative behavior of the other tracers, time sections and depth layers which also holds for all subsequent plots shown in this paper.

6.1.1. *Treatment of the Carbon Primary Production*

As explained in Section 4.1, the original fine model response for the carbon primary production obeys a daily cycle. When employing a coarser temporal discretization using a coarsening factor of $\beta = 40$ (corresponding to a discrete time step in the coarse model of 40 hours, cf. (16)), it is clearly not possible to resemble the full, high-frequent fine model response of CUP (cf. Figure 3a). This would instead require a time step smaller than the period of the main features, i.e., a time step smaller than 24 hours (or, equivalently, $\beta < 24$).

As was described in Section 4.4, the original cost function (16) uses a 24-hourly temporal mean of the fine model response for the state CUP. Thus, for the surrogate-based optimization, an approximation of this mean is desired.

Within each period of 24 hours, approximately half of the points are zero whereas the other half are non-zero as shown in Figure 3a. Thus, the 24-hourly temporal mean can be approximated by applying the rectangle rule for integration to each 12-hourly period incorporating the non-zero points and by subsequent division by the length of the period, i.e., by the factor 24. The resulting value of this approximate mean is hence equal

to a simple division of the corresponding midpoint of the non-zero interval by the factor two.

We apply this integration analogously to the coarse model response. According to the original problem formulation given in Section 4.3, the subsequent division by the factor two (as an approximation of the 24-hourly mean of the original fine model response) will be shifted to the cost function formulation for the coarse model optimization which we introduce in Section 7.2.

To yield the commensurable fine model response which is neccessary for the multiplicative response correction, the same integration and subsequent division are applied to the down-sampled fine model response for the state CUP.

We furthermore smoothen the response for CUP, since it is internally calculated in the model simulation and provided as an additional state vector of the model response (cf. Section 4.1). The response for CUP is thus based on unsmoothed, i.e., possibly numerically noisy, responses of the states N and P. Note, that our aim is to treat the model as a black-box tool and thus, we try to avoid any modifications of the original implementation.

The smoothing is applied in the same way as for the tracers N, P, Z and D as described in the last Section (cf. (18)). Again, the same smoothing is applied to the down-sampled fine model response for CUP. Altogether, we briefly write for the fine and coarse model state for CUP

$$(\mathbf{z}_f)_5 := S\,I(\bar{\mathbf{y}}_f)_5, \quad (\mathbf{z}_c)_5 := S\,I(\mathbf{y}_c)_5, \tag{20}$$

with S, again, denoting the smoothing operator, I describes the rectangular integration and where $\bar{\mathbf{y}}_f$, again, denotes the down-sampled fine model response as defined in (19).

Figure 3b shows the integrated and smoothed coarse model response for the state CUP and the corresponding curve for the down-sampled fine model response, respectively.

6.2. Surrogate Construction

The surrogate in iteration k, denoted as $\mathbf{s}_k$, is obtained by a multiplicative correction of the coarse model response at the iterate $\mathbf{u}_k$. Primarily, the response for the state CUP is integrated and smoothing to all five states is applied as motivated above.

The *correction factor*, denoted as a_k, is given by the point-wise division of the (down-sampled) and smoothed fine by the smoothed coarse model response at the iterate $\mathbf{u}_k$, i.e.,

$$\mathbf{a}_k := \frac{\mathbf{z}_f(\mathbf{u}_k)}{\mathbf{z}_c(\mathbf{u}_k)}, \quad k = 1, 2, \ldots \tag{21}$$

where the smoothed responses $\mathbf{z}_f$ and $\mathbf{z}_c$ are defined through (18)-(20) and where the correction factors are summarized in the vector $\mathbf{a}_k$.

6.2.1. Zero-order Consistent Surrogate

A *zero-order consistent* surrogate $\bar{\mathbf{s}}_k$ (cf. (6)) can be simply obtained as

$$\bar{\mathbf{s}}_k(\mathbf{u}) := \mathbf{a}_k\,\mathbf{z}_c(\mathbf{u}) \tag{22}$$

where the multiplication is again meant point-wise.

The surrogate defined in (22) does not satisfy the first-order consistency condition in (6) – i.e., agreement between first-order derivatives at the current iteration point – exactly. However, since the physics-based surrogate inherits substantial knowledge about the marine model under consideration, its derivatives are expected to be at least similar to those of the fine model. The surrogate in (22) has been used in [21] and demonstrated to already yield a remarkably accurate solution at the cost of a few fine model evaluations only.

Initial experiments carried out for the problem considered in this work reveal that implementation of the first-order consistency is important to ensure sufficient performance of the algorithm. In conjunction with TR convergence safeguards (cf. Section 3.3), this allows us to locate the fine model optimum more precisely.

6.2.2. Zero- and First-Order Consistent Surrogate

To ensure first-order consistency with the fine model response, we furthermore include an additive correction term E_k in the formulation (22) as follows

$$\begin{aligned} \mathbf{s}_k(\mathbf{u}) &:= \bar{\mathbf{s}}_k(\mathbf{u}) + E_k\,(\mathbf{u} - \mathbf{u}_k), \\ E_k &:= \mathbf{z}'_f(\mathbf{u}_k) - \bar{\mathbf{s}}'_k(\mathbf{u}_k), \end{aligned} \tag{23}$$

where $\mathbf{z}'_f$ and $\bar{\mathbf{s}}'$ denote the derivatives of the smoothed coarse and (down-sampled) fine model response, defined by (8), and where the term $\bar{\mathbf{s}}_k$ is defined by (22).

Obviously, the surrogate in (23) satisfies the zero- as well as first-order consistency condition with the fine model response in the point $\mathbf{u}_k$, more specifically with the down-sampled and smoothed response (cf. Section 6.1) as

$$\mathbf{s}_k(\mathbf{u}_k) = \mathbf{z}_f(\mathbf{u}_k), \quad \mathbf{s}'_k(\mathbf{u}_k) = \mathbf{z}'_f(\mathbf{u}_k). \tag{24}$$

Since we employ finite-differences to approximate the derivatives, clearly, this equivalence in (24) is not exact. Another source of inaccuracy is the presence of numerical noise. However, as demonstrated in the literature (see, e.g., [23]) and observed through initial experiments, even the use of approximate derivatives improves the algorithm performance. In the following analysis, for the sake of simplicity, we will accordingly treat agreement in first-order derivative information as nevertheless exact and won't explicitly mention inherent approximation errors due to finite differences and/or the noise.

6.2.3. Improvements of the Basic Surrogate Formulation

Occasionally, there might occur a situation where the coarse model response is close to zero (and maybe even negative due to approximation errors) and a few magnitudes smaller than the fine one, which leads to large (possibly negative) correction factors a_k. While such a correction ensures zero-order consistency at the point where it was established (i.e., $\mathbf{u}_k$), it may lead to (locally) poor approximation in the vicinity of this point. Resulting "spikes" appearing in the response due to large values of the correction term can be viewed, in a way, as a numerical noise that slows down the algorithm convergence and makes

the optimum more difficult to locate. This has already been investigated in [37], where an upper bound for a_k as well as a non-negative bound for the coarse model response (the negative response is non-physical and is a result of large time steps employed in the coarse model solution) has been suggested to address these issues.

More specifically, we apply the following modifications at each iteration k of an SBO run (cf. (5)):

$$(i)\ \mathbf{y}_c = \begin{cases} 0; & \text{if } \mathbf{y}_c < 0 \\ \mathbf{y}_c; & \text{else} \end{cases},\ (ii)\ \mathbf{a}_k = \begin{cases} a_{\text{ub}}; & \text{if } \mathbf{a}_k > a_{\text{ub}} \\ \mathbf{a}_k; & \text{else} \end{cases}, \tag{25}$$

(iii) $\mathbf{a}_k = 1$ if ($\mathbf{z}_c \leq \delta$ and $\mathbf{z}_f \leq \delta$),

where the operations are again meant point-wise, where (i) is applied before smoothing and where δ should be of the order of the discretization error below which the responses can be treated as zero. For the considered problem, $a_{\text{ub}} = 5$ turned out to be a reasonable choice and we furthermore consider $\delta = 10^{-4}$.

As a consequence of restricting the correction factors a_k as in (25), the zero-order consistency condition in (24) can only be satisfied approximately, i.e.,

$$\mathbf{s}_k(\mathbf{u}_k) = \mathbf{z}_f(\mathbf{u}_k) + \epsilon, \tag{26}$$

with ϵ thus denoting the difference between the corrected coarse model and the fine model response in the point $\mathbf{u}_k$. Although consistency in the first-order derivative of the surrogate model (23) and the fine model response as in (24) is nevertheless satisfied, agreement in the first-order derivative of the corresponding cost function values is not, as will be explained below.

Assuming a general Euclidean least-squares norm J, measuring the misfit between the model response $\mathbf{y}$ and some specification $\mathbf{y}_d$, the gradient J' is given by

$$\begin{aligned} J'(\mathbf{y}) &= \left.\frac{\mathrm{d}J(\mathbf{y})}{\mathrm{d}\mathbf{u}}\right|_{\mathbf{u}=\mathbf{u}_k} = \left(\mathbf{y}(\mathbf{u}_k) - \mathbf{y}_d\right)^T \mathbf{y}'(\mathbf{u}_k), \\ J(\mathbf{y}) &:= \frac{1}{2} \| \mathbf{y}(\mathbf{u}) - \mathbf{y}_d \|_2^2. \end{aligned} \tag{27}$$

The surrogate's cost function gradient is then given by

$$\begin{aligned} J'(\mathbf{s}_k(\mathbf{u}_k)) &= \left(\mathbf{s}_k(\mathbf{u}_k) - \mathbf{y}_d\right)^T \mathbf{s}_k'(\mathbf{u}_k) \\ &= \left(\mathbf{z}_f(\mathbf{u}_k) + \epsilon - \mathbf{y}_d\right)^T \mathbf{z}_f'(\mathbf{u}_k). \end{aligned} \tag{28}$$

The exact agreement in the cost function gradients, i.e., $J'(\mathbf{s}_k(\mathbf{u}_k)) = J'(\mathbf{z}_f(\mathbf{u}_k))$, can only be obtained if, besides the first-order consistency, also exact zero-order consistency as is in (24) is ensured, i.e., if $\epsilon = 0$.

It is not clear, whether an exact agreement in the first-order derivative (apart from approximation errors) is actually necessary in practice. Nevertheless, we use another additive term, denoted as D_k, in the definition of the surrogate (23), that allows us to eliminate any possible influence of the problem described above. The surrogate, which we finally employ in the optimization, is formulated as

$$\begin{aligned} \mathbf{s}_k(\mathbf{u}) &:= \bar{\mathbf{s}}_k(\mathbf{u}) + D_k + E_k\,(\mathbf{u} - \mathbf{u}_k), \\ D_k &:= \mathbf{z}_f(\mathbf{u}_k) - \bar{\mathbf{s}}_k(\mathbf{u}_k) = \epsilon, \end{aligned} \tag{29}$$

where the terms $\bar{\mathbf{s}}_k$ and E_k are defined in (22) and (23) and where the underlying correction factors a_k are restricted as suggested in (25). Note, that the term D_k can be obtained at no additional costs, since the neccessary quantities at that iterate are already incorporated in the underlying surrogate formulation (cf. (23)) and are thus available.

The surrogate in (29) satisfies exact zero- and first-order consistency, both with respect to the fine model response $\mathbf{z}_f$ as well as with respect to its cost function $J(\mathbf{z}_f)$ in the current point $\mathbf{u}_k$.

7. Optimization Setup

The operation and performance of the proposed surrogate-based algorithm is illustrated through the results of exemplary optimization runs with $n_z = 33, n_{\tau,f} = 8760 \cdot 5$ and $\beta = 40$, which means that we obtain $n_{\tau,c} = n_{\tau,f}/\beta = 1095$ discrete time steps for the coarse model. The specific choice of β has been motivated in Section 5 (see also [21]). In [38] it has already been demonstrated that, at least from point of view of the optimization results, the vertical model grid can be reduced to $n_z = 33$ depth layers, instead of the originally employed 66. It has been demonstrated that optimization of both models yield practically identical results w.r.t. parameter match and quality of the optimal solution.

In the following, the solutions' quality and the computational cost of the surrogate-based optimization is compared to the results of a direct fine and coarse model optimization. The quality of the solutions is assessed by visual inspection of the model response and inspection of the corresponding cost function and parameter values.

The computational costs of the distinct optimization processes is measured in so-called *equivalent* fine model evaluations. This means, that for the considered coarse model, β evaluations (with $\beta = 40$ in this paper) are equivalent to (or, as expensive as) one fine model evaluation, which is a result of the chosen coarser discretization employing the factor β (cf. Section 5). On the other hand, the cost of one iteration of the surrogate-based optimization (in terms of equivalent fine model evaluations) equals to the number of coarse model evaluations necessary to optimize the surrogate model divided by this factor β, and increased by the cost for the response correction.

For the proposed surrogate (29), the cost of the correction is approximately 13 equivalent fine model evaluations. This mainly results from the cost of the actual fine model evaluations: one for the multiplicative correction, and 12 for the finite differentiation (the model has 12 parameters). The further cost of the coarse model evaluation plus its jacobian, correspondingly 13 divided by the factor $\beta = 40$, is negligible here.

For all optimization runs we used the MATLAB[1] routine `fmincon`, exploiting the active-set algorithm.

As has been verified for example in [19], the solution obtained by using both local, gradient-based and global, genetic algorithms, provided no suitable fit of the target. Obtaining a

[1] MATLAB is a registered trademark of The MathWorks, Inc., http://www.mathworks.com

Table 2: Optimization problems considered in this work with $\mathbf{u}$ denoting the optimization variable. $U_{ad} := \{\mathbf{u} \in \mathbb{R}^{12} : \mathbf{b}_l \leq \mathbf{u} \leq \mathbf{b}_u\}$, denotes the space of admissible parameters with component-wise upper and lower bounds $\mathbf{b}_u$ and $\mathbf{b}_l$, respectively, as more specifically given in Table 3 (see also Section 4.4). Cost functions J_1 and J_2 are formulated in (16) and (30). Prior to (O.1) - (O.3), we further perform a random search with an initial guess $\mathbf{u}_0$. For the sake of brevity, corresponding results are omitted here.

$\mathbf{u}_i$	Description	Optimization Problem
$\mathbf{u}_0$	Randomly chosen initial parameter vector	
$\mathbf{u}^*_{f1}$	Result of an *original* fine model optimization	$\mathbf{u}^*_{f1} := \underset{\mathbf{u}\in U_{ad}}{\operatorname{argmin}}\ J_1(\mathbf{y}_f(\mathbf{u}))$ (O.1)
$\mathbf{u}^*_{f2}$	Result of a *reference* fine model optimization	$\mathbf{u}^*_{f2} := \underset{\mathbf{u}\in U_{ad}}{\operatorname{argmin}}\ J_2(\mathbf{z}_f(\mathbf{u}))$ (O.2)
$\mathbf{u}^*_c$	Result of a coarse model optimization	$\mathbf{u}^*_c := \underset{\mathbf{u}\in U_{ad}}{\operatorname{argmin}}\ J_2(\mathbf{z}_c(\mathbf{u}))$ (O.3)
$\mathbf{u}^*_s$	Result of a SBO run using $\mathbf{u}^*_c$ as initial parameter vector	$\mathbf{u}_{k+1} = \underset{\mathbf{u}\in U_{ad}, \|\mathbf{u}-\mathbf{u}_k\|^2 \leq \delta_k}{\operatorname{argmin}}\ J_2(\mathbf{s}_k(\mathbf{u}))$, $k = 0, 1, \ldots$, $\mathbf{u}_0 := \mathbf{u}^*_c$ (O.4)

better result with other optimization methods seems not very likely. Thus we tentatively accept the found minima in [19] and argue that the *NPZD* model in the current formulation will have to be changed or extended to yield a better quality of the fit.

However, it is not the focus of this paper to further address this issue. Our aim clearly is to demonstrate the applicability of the proposed approach to the parameter optimization of the considered model. More specifically, the focus is to demonstrate that, by exemplary optimization runs, SBO is able to yield a solution close to the one obtained by a direct fine model optimization at low optimization costs.

It should be emphasized that, given attainable measurement data, direct fine model optimization is able to reconstruct the target and corresponding optimal parameters (i.e., the discrete model is well suited for parameter identification, see e.g. [19]). Secondly, the performance of SBO is similar, i.e., a solution which fits the observed quantities can be obtained at low computational costs. In [21], this has been verified using model-generated, attainable target data, where a surrogate as formulated in (22) has been employed in an illustrative SBO run.

Because the optimization problem under consideration is quite complex, we employ, for both the fine and coarse model optimization, a random search algorithm prior to the MATLAB's gradient-based `fmincon`. This turned out to be quite suitable for the considered problem to locate a rough solution initially at low computational costs (since no sensitivity data is used). More specifically, we use 500 model evaluations within this algorithm, which turned out to yield a reasonable trade-off between the accuracy of the solution and the optimization cost.

Furthermore, the initial point for the SBO algorithm run is the optimal solution of the coarse model. This is the best approximation of the optimum that we can obtain at a low cost, without involving the fine model at all.

7.1. Reference Fine Model

To be precise, we have to distinguish between two fine model responses and corresponding optimization problems.

Firstly, we consider the optimization of the *original* fine model for comparison which has been utilized in various optimization runs (see, e.g., [18, 19, 20]) and which we briefly described in Section 4.4. Let us recall that "original" denotes the fine model with a time step of $\tau_f = 1\,\mathrm{h}$ where no further operations such as down-sampling and smoothing have been applied (cf. Section 4.3).

However, the coarse model response, due to the employed coarser temporal discretization and applied smoothing, is supposed to provide an approximation of the down-sampled and smoothed fine model response $\mathbf{z}_f$. Moreover, the surrogate (29) is zero- and first-order consistent with the transformed fine model response $\mathbf{z}_f$ (cf. (24)). Consequently, in order to obtain a fair comparison, the *down-sampled* and *smoothed* fine model response and corresponding optimization has to be treated as the actual *reference*. A formulation of the corresponding reference cost function is provided below in Section 7.2. A well performing surrogate-based algorithm, exploiting the proposed surrogate, is thus expected converges to at least a local minimum of this *reference optimization problem* (cf. Section 3.3).

7.2. Cost Function – Reference Fine, Corse and Surrogate Model Optimization

The cost function used in conjunction with the original fine model response $\mathbf{y}_f$ has already been formulated in Section 4.4.

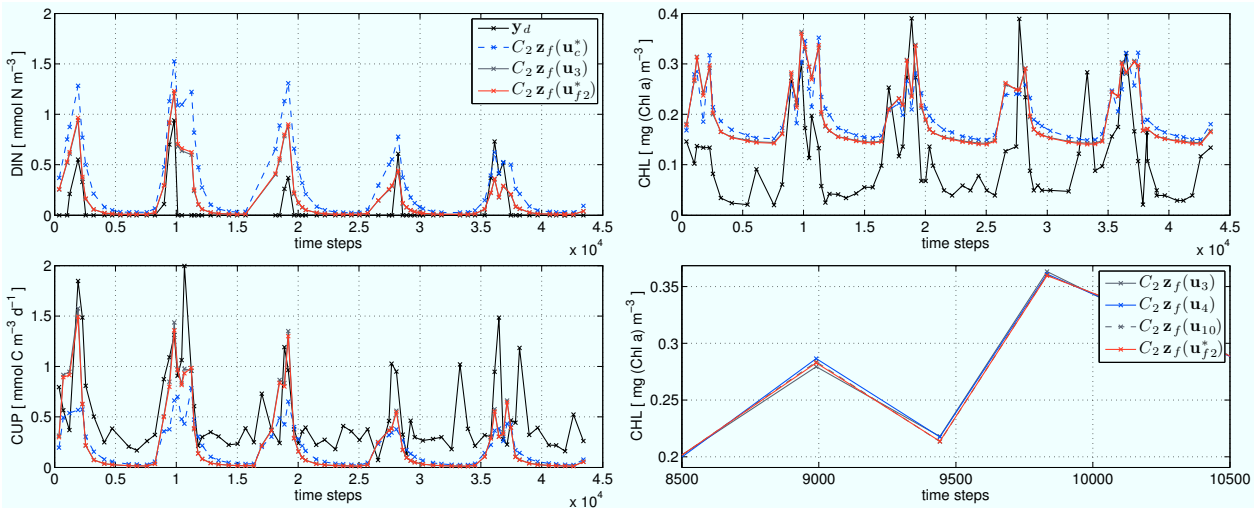
In this section we now propose a formulation for the reference fine, coarse model and for the surrogate-based optimization, which briefly reads

$$J_2(\mathbf{z}) := \| C_2\,\mathbf{z} - \mathbf{y}_d \|^2_\sigma,$$

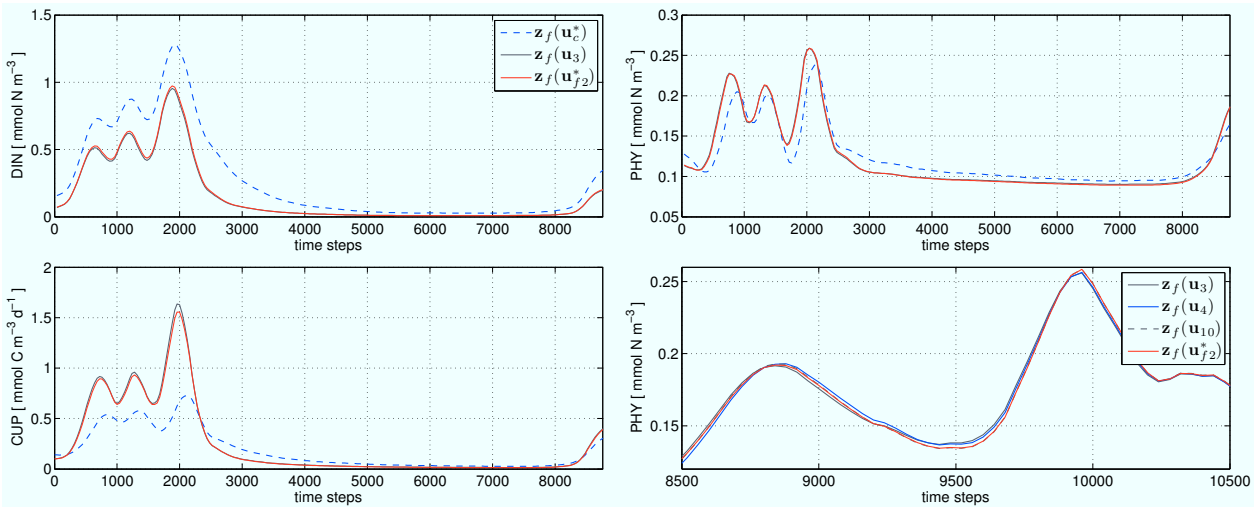
$$\mathbf{z} = \begin{cases} \text{reference fine model response,} & \mathbf{z} = \mathbf{z}_f \\ \text{smoothed coarse model response,} & \mathbf{z} = \mathbf{z}_c \\ \text{surrogate's response at iteration } k, & \mathbf{z} = \mathbf{s}_k \end{cases} \tag{30}$$

where, again, we choose an Euclidean norm weighted by assumed standard deviations of the measurements $\sigma = (\sigma_j)_{j=1\ldots,5}$ (see [18, 19]) and where the operator C_2, which has to be used to make the model response commensurable with the measurements, is similar to the one, C_1, used in the original cost function (cf. Section 4.4).

However, differently to the operator C_1 in (16), the constant temporal alignment is adjusted to the coarser temporal grid. As was motivated in Section 6.1.1, we further employ a simple division of the smoothed response $(\mathbf{z})_5$ (i.e., the response of state



(a) Transformed response using the operator C_2 to make it commensurable with the measurement data $\mathbf{y}_d$.



(b) Untransformed response to assess the overall quality.

Figure 4: Reference fine model response $\mathbf{z}_f$ at the solutions $\mathbf{u}_{f2}^*, \mathbf{u}_c^*, \mathbf{u}_3$ of an exemplary (reference) fine, coarse model and of a SBO run after three iterations (cf. Table 2). Responses are shown for three illustrative tracers, some depth layer and, in Figure (b), for the sake of better visibility, for a section of the whole time interval. Lower right plots: subsequent iterates $\mathbf{u}_4$ and $\mathbf{u}_{10}$ obtained in the SBO, here, by means of chlorophyll and phytoplankton and for an even smaller time section, since changes are small.

CUP) by the factor two, in order to obtain an approximation of the 24-hourly mean which is applied to the original hourly fine model response for this state. Again, for the sake of simplicity, we omit any detailed formulation of the operator C_2 here.

7.3. Optimization Problems and Comparison of Solutions

In the following, we account for the solutions of both an illustrative reference (cf. Section 7.1) and original fine model optimization (cf. Section 4.4), denoted as $\mathbf{u}_{f2}^*$ and $\mathbf{u}_{f1}^*$, respectively. We further consider the solution $\mathbf{u}_c^*$ of a coarse model optimization which is used as initial point for an illustrative SBO run. The solution obtained by SBO will be denoted by $\mathbf{u}_s^*$.

For the sake of brevity, results of the priorly performed random search as motivated above are omitted in the following. The underlying cost functions have been formulated in Sections 4.4 and 7.2. The four optimization problems (omitting the random search) are denoted as (O.1) - (O.4) and – to provide a clear overview – are again summarized in Table 2.

To verify the performance of the proposed method, we consider the reference fine model response $\mathbf{z}_f$ (cf. Section 7.1) and corresponding cost function value $J_2(\mathbf{z}_f)$ (cf. (30)) at the respective optima $\mathbf{u}_{f2}^*, \mathbf{u}_c^*$ and $\mathbf{u}_s^*$. In order to also account for the solution $\mathbf{u}_{f1}^*$ obtained by the original fine model optimization (cf. Section 4.4), we subsequently present the original fine model response $\mathbf{y}_f$ and corresponding cost function value $J_1(\mathbf{y}_f)$.

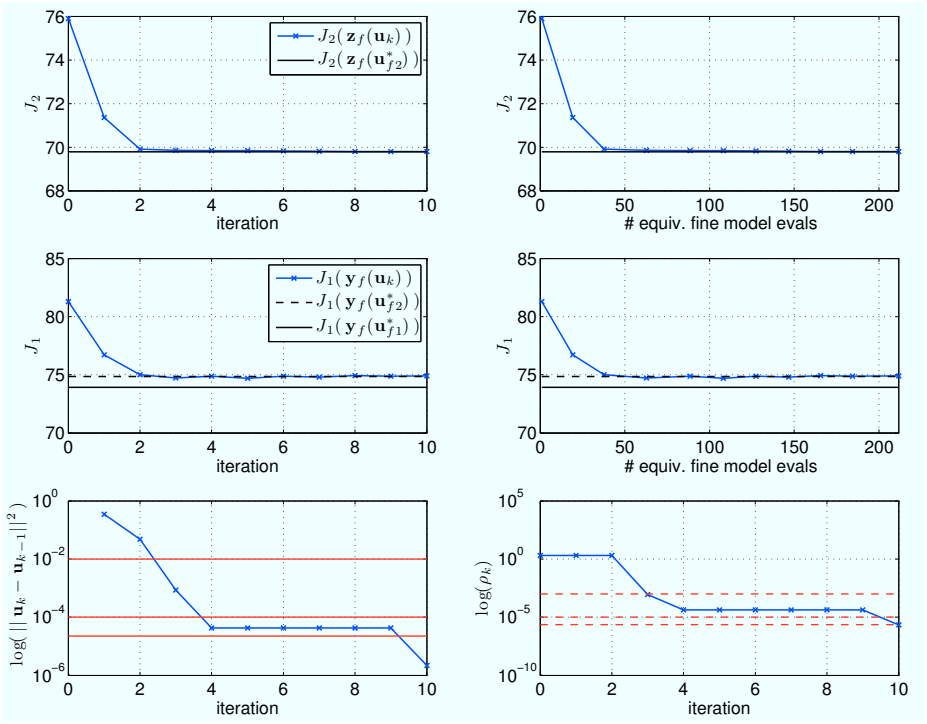


Figure 5: Optimization history (SBO) of the original and the reference cost function value J_1 and J_2, both versus number of iterations and the computational costs (measured in *equivalent* number of fine model evaluations). Also shown are the corresponding cost function values at optimal solutions $\mathbf{u}^*_{f1}$ and $\mathbf{u}^*_{f2}$ (obtained by the original and reference fine model optimization), the optimization history of the trust-region radius δ_k and of the squared step size norm. Note that, for the sake of simplicity, we omit explicit results of the prior coarse model optimization run. Horizontal red solid/dashed lines denote distinct termination conditions considered in the SBO. Legends also apply to the plots on the right, accordingly.

7.4. Trust Region Approach

For the SBO, we use a trust-region (TR) safeguard (cf. Section 3.3) to further increase the robustness of the optimization process, i.e., to guarantee convergence to at least a local minimum of the (reference) optimization problem. The TR radius δ_k is updated according to standard rules as follows [22, 23]

$$\delta_0 = 2, \quad \delta_k = \begin{cases} \delta_k / m_{\text{decr}}, & \text{if} \quad \rho_k < r_{\text{decr}} \\ \delta_k \cdot m_{\text{incr}}, & \text{if} \quad \rho_k > r_{\text{incr}} \end{cases}, \qquad (31)$$
$$r_{\text{incr}} = 0.75, \quad r_{\text{decr}} = 0.01, \quad m_{\text{incr}} = 3, \quad m_{\text{decr}} = 20,$$

with ρ_k denoting the *gain ratio* in iteration k defined as

$$\rho_k := \frac{f_{\text{new}} - f_{\text{old}}}{s_{\text{new}} - s_{\text{old}}}, \qquad (32)$$
$$f_{old} := J_2(\mathbf{z}_f(\mathbf{u}_k)), \; f_{new} := J_2(\mathbf{z}_f(\mathbf{u}_{k+1})),$$
$$s_{old} := J_2(\mathbf{s}_k(\mathbf{u}_k)), \; s_{new} := J_2(\mathbf{s}_k(\mathbf{u}_{k+1})).$$

Except for r_{decr} and m_{decr}, which are smaller, respectively larger than usual, all values specified above are fairly standard. It turned out that these values are typically more suitable for surrogate-based optimization schemes exploiting physics-based surrogates (see, e.g., [23]).

7.5. Stopping Criterion

As a termination condition for the SBO, we use the absolute step size (measured in the Euclidean norm) between two successive iterates $\mathbf{u}_k$ and $\mathbf{u}_{k-1}$ as well as a lower bound for the TR radius δ_k, in the following denoted by $\delta_k^{\min}$. In practice, we choose a smaller bound for TR radius than for gamma because of the large value of m_{decr}. The solution $\mathbf{u}^*_s$ obtained by SBO is thus defined as

$$\mathbf{u}^*_s := \left\{ \mathbf{u}_k \;\middle|\; \left(\| \mathbf{u}_k - \mathbf{u}_{k-1} \|^2 \leq \gamma \right) \vee \left(\delta_k \leq \delta_k^{\min} \right) \right\}. \qquad (33)$$

To test various trade-offs between the quality of the solution obtained by SBO and the corresponding computational cost we consider three distinct values for the threshold, more specifically

$$\{\gamma, \delta_k^{\min}\} = \{10^{-2}, 10^{-3}\}, \{10^{-4}, 10^{-5}\}, \{2.5 \cdot 10^{-5}, 2.5 \cdot 10^{-6}\}.$$

8. Numerical Results and Outlook

In Figure 4, the solutions of the exemplary optimization runs as described in Section 7.3 (cf. optimization problems (O.2) - (O.4) in Table 2) are presented, comparing the reference fine model response $\mathbf{z}_f$ at the respective optima. More specifically,

Table 3: Specific parameter values of the solutions obtained by coarse and fine (original and reference) model optimization and of the iterates 3, 4 and 10 obtained in a SBO run. Also shown are the considered upper and lower parameter bounds $\mathbf{b}_u$ and $\mathbf{b}_l$.

iterate	$u_{k,1}$	$u_{k,2}$	...									$u_{k,12}$
	Solution of a coarse model optimization (O.3)											
$\mathbf{u}_c^*$	1.000	0.522	0.051	0.040	0.010	4.000	4.000	0.007	0.059	0.010	0.747	5.000
	Solution of SBO (O.4):											
$\mathbf{u}_3$	1.000	1.195	0.052	0.035	0.024	4.000	4.000	0.003	0.089	0.010	1.000	5.000
$\mathbf{u}_4$	1.000	1.193	0.052	0.034	0.026	4.000	4.000	0.004	0.095	0.010	1.000	5.000
$\mathbf{u}_{10}$	1.000	1.176	0.048	0.035	0.018	4.000	4.000	0.004	0.094	0.010	1.000	5.000
	Solution of a (reference) fine model optimization (O.2)											
$\mathbf{u}_{f2}^*$	1.000	1.145	0.049	0.035	0.020	4.000	4.000	0.003	0.095	0.010	1.000	5.000
	Solution of a (original) fine model optimization (O.1)											
$\mathbf{u}_{f1}^*$	1.000	1.063	0.112	0.043	0.082	4.000	4.000	0.004	0.081	0.010	1.000	5.000
$\mathbf{b}_l$	0.300	0.200	0.001	0.000	0.010	0.025	0.040	0.000	0.010	0.010	0.100	2.000
$\mathbf{b}_u$	1.000	1.460	0.253	0.630	0.730	4.000	4.000	0.630	1.000	0.150	1.000	5.000

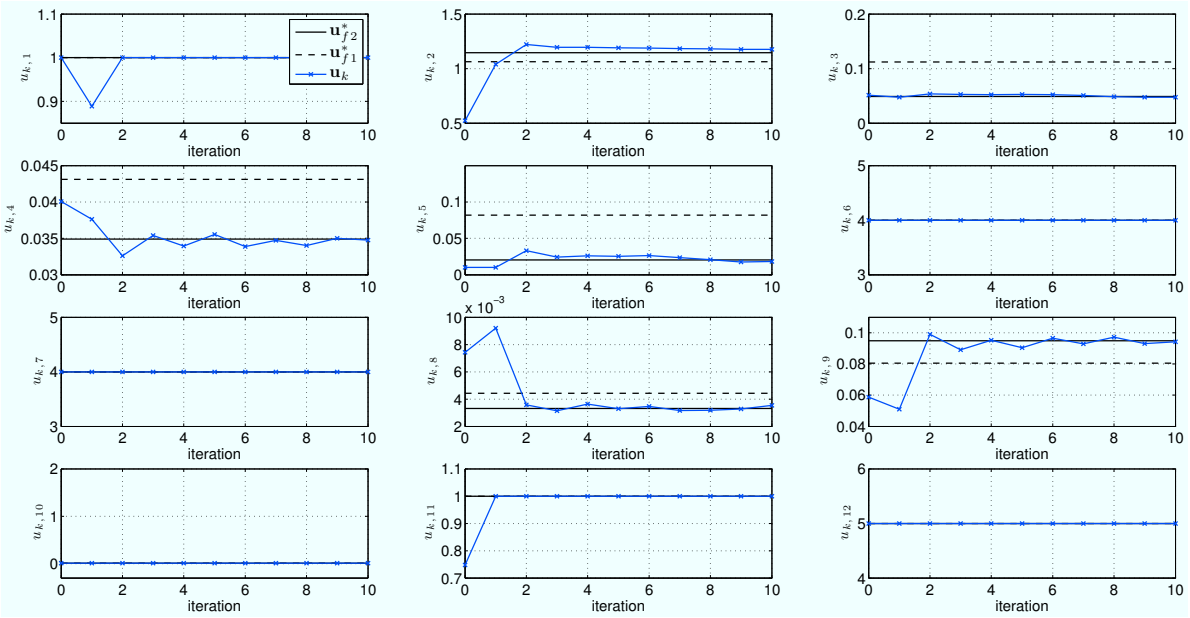
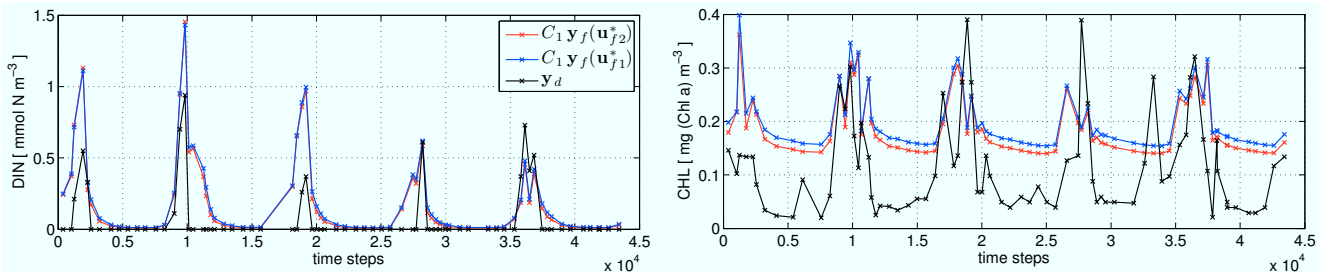


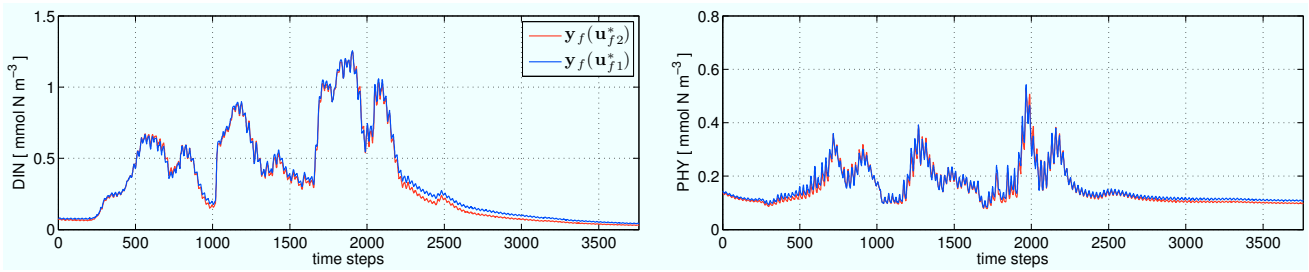
Figure 6: Optimization history of the parameters $\mathbf{u}_k$ obtained in the SBO as well as optimal parameter $\mathbf{u}_{f1}^*$ and $\mathbf{u}_{f2}^*$ obtained by direct fine model optimization (original and reference).

Figure 4a shows the model response which is transformed (using the operator C_2) to be commensurable with the given measurement data $\mathbf{y}_d$, i.e., $C_2\,\mathbf{z}_f$ (cf. Section 7.2). Shown are the solution $\mathbf{u}_3$ obtained by SBO after 3 iterations (corresponding to a stopping criterion of $\{\gamma, \delta_k^{\min}\} = \{10^{-2}, 10^{-3}\}$). Furthermore shown are – by means of the tracer Chlorophyll a – also the subsequent solutions $\mathbf{u}_4$ and $\mathbf{u}_{10}$ after 4 and 10 iterations (corresponding to $\{\gamma, \delta_k^{\min}\} = \{10^{-4}, 10^{-5}\}$ and $\{2.5 \cdot 10^{-5}, 2.5 \cdot 10^{-6}\}$, respectively). In order to verify that also the overall quality of the solution obtained by SBO is sufficiently close to the one obtained by fine model optimization, we present the corresponding "untransformed" response $\mathbf{z}_f$ in Figure 4b.

It can be observed that SBO converges to the optimal solution $\mathbf{u}_{f2}^*$ obtained by the reference fine model optimization as shown in Figures 4, 4b, 5 and 6 (see also Table 3). Whereas coarse model optimization provides a rather inaccurate solution (i.e., not close to the reference fine one), SBO is able to yield a remarkably accuracy already after 3 iterations – both in terms of quality of the solution (cf. Figure 4, 4b and 5) and parameter match (cf. Figure 6 and Table 3). Only approximately 63 equivalent fine model evaluations were required (cf. Figure 5) for these first three iterations of the SBO run, whereas the prior coarse model optimization requires 22 equivalent fine model evaluations which adds to the costs of the SBO. On the other



(a) Response is transformed using the operator C_1 to make it commensurable with the measurement data $\mathbf{y}_d$.



(b) Untransformed response to assess the overall quality.

Figure 7: Original fine model response $\mathbf{y}_f$, for comparison, at the solutions $\mathbf{u}^*_{f1}$, $\mathbf{u}^*_{f2}$ of an original and reference fine model optimization. Responses are shown for two illustrative tracers, some depth layer and, in (b), for the sake of better visibility, for a section of the whole time interval.

hand, fine model optimization requires 567 model evaluations. This corresponds to a significant reduction in the optimization cost down to approximately 15% of that of a direct fine model optimization.

Subsequent iterations within the SBO (or, equivalently, decreasing the thresholds used in the stopping criterion (33)), as shown in lower right plots in Figures 4 and 4b and in Figures 5, 6 and Table 3, only marginally increases the accuracy of its solution. However, proceeding to the subsequent iterate $\mathbf{u}_{10}$ in the SBO requires approximately 150 additional equivalent fine model evaluations (cf. upper right plot in Figure 5). Thus, terminating the SBO after the first three iterations seems appropriate for the considered problem.

Furthermore, the trade-offs between the solution accuracy and the extra computational overhead related to sensitivity calculation have been investigated by additional numerical experiments. It turns out that without using fine/coarse model sensitivity, the solution of SBO is not sufficiently accurate for the considered problem, whereas, clearly, the cost savings are higher. For the sake of brevity, the results of these additional experiments are not shown here.

8.1. Quality of Reference and Original Fine Model Solution

In order to assess the quality of the solutions obtained by the reference and original fine model optimization, Figures 7a and 7b furthermore present the original fine model response $\mathbf{y}_f$ at the two fine model solutions, $\mathbf{u}_{f1}$ and $\mathbf{u}_{f2}$. Shown is the transformed model response $C_1\,\mathbf{y}_f$ in Figure 7a to assess the quality and difference of the responses with respect to the measurement data. Also shown is the "untransformed" response $\mathbf{y}_f$ in Figure 7b to furthermore investigate the overall quality. For better visibility, a smaller illustrative time section is selected here. Figures 5 and 6 present the optimal parameters $\mathbf{u}^*_{f1}$ and corresponding cost function value $J_1(\mathbf{y}_f(\mathbf{u}^*_{f1}))$, for comparison with the corresponding quantities, $\mathbf{u}^*_{f2}$ and $J_1(\mathbf{y}_f(\mathbf{u}^*_{f2}))$, for the reference solution, respectively.

It can be observed that both solutions are fairly close in terms of the quality of the responses (further reflected by the corresponding quite close cost function values) and in terms of the parameter match. The rather small differences between the original and reference fine model solutions is mainly due to down-sampling and smoothing. These operations have been applied to the coarse and fine model response but not to the target data, since the real measurements are not given on the temporal grid of the fine model but are rather sparsely distributed. However, for dense model-generated data, it has been shown in [21] that these differences are marginal and the solution obtained by surrogate-based optimization converges to the original one as desired.

9. Conclusions

Computationally efficient calibration of a marine ecosystem model is presented. We exploit a surrogate-based optimization algorithm working with a low-fidelity (or coarse) model obtained from a temporal coarser discretization. We employ a multiplicative correction to the coarse model response which allows us to create a reliable approximation (the surrogate) of the computationally expensive original (or fine) model. The surrogate model is furthermore enhanced by using fine model

sensitivity information. The algorithm is embedded in the trust-region approach.

As a case study, we consider a selected representative of the class of one-dimensional marine ecosystem models. The complexity of the response of this specific model is comparably high. Thus, and because biochemistry mainly happens locally in space, this model serves as a suitable test case before investigating computationally more expensive three-dimensional models.

As demonstrated through numerical experiments, the presented approach yields a solution of high quality at a computational cost that is substantially lower than for the direct fine model optimization. Time savings are as high as 85 percent. A basic formulation of this approach (without exploiting fine model sensitivity) has already been investigated in [21] and demonstrated to yield an accurate solution at low computational costs for model-generated, synthetic target data.

The enhancements of the surrogate model and the optimization algorithm (the use of fine model sensitivity and a trust-region approach) introduced in this work are essential to calibrate the model against real, measurement data. Specifically, they ensure a suffiently accurate solution of the surrogate-based optimization while retaining the high computational savings when compared to a direct fine model optimization.

10. Acknowledgments

The authors would like to thank Andreas Oschlies, IFM Geomar, Kiel. This research was supported by the DFG Cluster of Excellence Future Ocean.

References

[1] K. McGuffie, A. Henderson-Sellers, A Climate Modelling Primer, Wiley, Chichester, 3rd edition, 2005.

[2] A. Majda, Introduction to PDE's and Waves for the Atmosphere and Ocean, AMS, 2003.

[3] A. E. Gill, Atmosphere - Ocean Dynamics, volume 30 of *International Geophysics Series*, Academic Press, San Diego et al., 1982.

[4] J. L. Sarmiento, N. Gruber, Ocean Biogeochemical Dynamics, Princeton University Press, Princeton et al., 2006.

[5] W. Fennel, T. Neumann, Introduction to the Modelling of Marine Ecosystems, Elsevier, Amsterdam et al., 2004.

[6] M. J. R. Fasham (Ed.), Ocean Biogeochemistry. The Role of the Ocean Carbon Cycle in Global Change., Global Change – The IGBP Series, Springer, Berlin et al., 2003.

[7] H. W. Engl, C. Flamm, P. Kügler, J. Lu, S. Müller, P. Schuster, Inverse problems in systems biology, Inverse Problems 25 (2009).

[8] A. Tarantola, Inverse Problem Theory and Methods for Model Parameter Estimation, SIAM, Philadelphia, 2005.

[9] M. Voßbeck, M. Clerici, T. Kaminski, T. Lavergne, B. Pinty, R. Giering, An inverse radiative transfer model of the vegetation canopy based on automatic differentiation, Inverse Problems 26 (2010) 095003.

[10] E. Weng, C. Gao, Y. Luo, R. Oren, Uncertainty analysis on forest carbon sink forecast with varying measurement errors: A data assimilation analysis, IOP Conference Series: Earth and Environmental Science 6 (2009).

[11] N. V. Queipo, R. T. Haftka, W. Shyy, T. Goel, R. Vaidyanathan, P. K. Tucker, Surrogate-based analysis and optimization, Progress in Aerospace Sciences 41 (2005) 1–28.

[12] T. W. Simpson, J. D. Poplinski, P. N. Koch, J. K. Allen, Metamodels for computer-based engineering design: Survey and recommendations, Engineering with Computers 17 (2001) 129–150.

[13] A. J. Smola, B. Schölkopf, A tutorial on support vector regression, Statistics and Computing 14 (2004) 199–222.

[14] J. Søndergaard, Optimization using surrogate models – by the space mapping technique, Ph.D. thesis, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, 2003.

[15] J. W. Bandler, Q. S. Cheng, S. A. Dakroury, A. S. Mohamed, M. H. Bakr, K. Madsen, J. Søndergaard, Space mapping: The state of the art, IEEE Transactions on Microwave Theory and Techniques 52 (2004) 337–361.

[16] A. I. J. Forrester, A. J. Keane, Recent advances in surrogate-based optimization, Progress in Aerospace Sciences 45 (2009) 50–79.

[17] L. Leifsson, S. Koziel, Multi-fidelity design optimization of transonic airfoils using physics-based surrogate modeling and shape-preserving response prediction, Journal of Computational Science 1 (2010) 98–106.

[18] M. Schartau, A. Oschlies, Simultaneous data-based optimization of a 1d-ecosystem model at three locations in the north atlantic: Part I - method and parameter estimates, Journal of Marine Research 61 (2003) 765–793.

[19] J. Rückelt, V. Sauerland, T. Slawig, A. Srivastav, B. Ward, C. Patvardhan, Parameter optimization and uncertainty analysis in a model of oceanic CO_2-uptake using a hybrid algorithm and algorithmic differentiation, Nonlinear Analysis B Real World Applications 10 (2010) 3993–4009.

[20] M. Schartau, Data-assimilation studies of marine, nitrogen based, ecosystem models in the North Atlantic Ocean, Ph.D. thesis, Christian-Albrechts-Universität Kiel, 2001.

[21] M. Prieß, S. Koziel, T. Slawig, Surrogate-based optimization of climate model parameters using response correction, Journal of Computational Science 2 (2011) 335–344.

[22] A. R. Conn, N. I. M. Gould, P. L. Toint, Trust-region methods, MPS-SIAM series on optimization, Society for Industrial and Applied Mathematics, Philadelphia, 2000.

[23] S. Koziel, J. W. Bandler, Q. S. Cheng, Robust trust-region space-mapping algorithms for microwave design optimization, IEEE Transactions on Microwave Theory and Techniques 58 (2010) 2166–2174.

[24] P. Kunkel, V. Mehrmann, Differential-algebraic equations: analysis and numerical solution, EMS, 2006.

[25] S. M. Griffies, Fundamentals of Ocean Climate Models, Princeton University Press, Princeton, New Jersey, 2004.

[26] J. D. Mahlman, W. J. Moxim, Tracer simulation using a global general circulation model – results from a midlatitude instantaneous source experiment, Journal of the Atmospheric Sciences 35 (1978) 1340–1374.

[27] S. Khatiwala, M. Visbeck, M. Cane, Accelerated simulation of passive tracers in ocean circulation models, Ocean Modelling 9 (2005) 51–69.

[28] D. Echeverria, P. Hemker, Manifold mapping: A two-level optimization technique, Computing and Visualization in Science 11 (2008) 193–206.

[29] S. Koziel, Shape-preserving response prediction for microwave design optimization, IEEE Transactions on Microwave Theory and Techniques 58 (2010) 2829–2837.

[30] N. M. Alexandrov, J. E. Dennis, R. M. Lewis, V. Torczon, A trust-region framework for managing the use of approximation models in optimization, Structural and Multidisciplinary Optimization 15 (1998) 16–23.

[31] A. Oschlies, V. Garçon, An eddy-permitting coupled physical-biological model of the North Atlantic 1. Sensitivity to advection numerics and mixed layer physics, Global Biogeochemical Cycles 13 (1999) 135–160.

[32] J. Liebig, L. Playfair, J. Webster, Chemistry in its application to agriculture and physiology, J. Owen, 1842.

[33] G. T. Evans, J. S. Parslow, A model of annual plankton cycles, Biological Oceanography 3 (1985) 328–347.

[34] A. Redfield, B. Ketchum, F. Richard, The influence of organisms on the composition of sea water, in: M. Hill (Ed.), The Sea, Wiley, New York, 1963, pp. 26–77.

[35] G. I. Marchuk, Methods of Numerical Mathematics, Springer, New York et al., 2nd edition, 1982.

[36] C. A. J. Fletcher, Computational Techniques for Fluid Dynamics, volume 1, Springer, 2nd edition, 1991.

[37] M. Prieß, S. Koziel, T. Slawig, 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, pp. 449–457.

[38] J. Rückelt, A. Oschlies, T. Slawig, Optimization of Parameters and Initial Values in a Marine NPZD-Type Ecosystem Model, Technical Report 1013, CAU Kiel, Institut für Informatik, 2010.

Accelerated Parameter Identification in 3D Marine Ecosystem Models Using Surrogate-Based Optimization

M. Prieß[a,1,*], J. Piwonski[a,1], S. Koziel[b], T. Slawig[a]

[a] *Institute for Computer Science, Cluster The Future Ocean, Christian-Albrechts Universität zu Kiel, 24098 Kiel, Germany*
[b] *Engineering Optimization & Modeling Center, School of Science and Engineering, Reykjavik University, Menntavegur 1, 101 Reykjavik, Iceland*

Abstract

We show the acceleration of parameter identification runs in a marine ecosystem model using an offline computation with transport matrices and the method of Surrogate-based Optimization (SBO). Aim is to identify parameters of the biogeochemical model such that the model output fits given data. The SBO approach replaces the original model by a surrogate, based on a coarser model and further aligned by a special multiplicative point-wise correction technique. In our example problem, the model computes a steady annual cycle for a two-tracer biogeochemical model, which is coupled to pre-computed ocean circulation. The steady annual cycle of the ecosystem is computed in a classical spin-up by integrating the model some thousands of years model time until no significant temporal changes in the system are measurable. The coarse model for the SBO approach is constructed by reducing the steps significantly in this fixed point type iteration. As a test case we use model-generated, synthetic data. We study an appropriate choice of the coarse model, give details on the multiplicative correction scheme and compare the computational effort of the SBO approach to the one of a direct fine model optimization. The obtained reduction in computing time is significant. The method is quite general and can be easily applied to other, also more complex biogeochemical models.

Keywords: climate models, marine ecosystem models, accelerated parameter identification, surrogate-based optimization, response correction, low-fidelity models

1. Introduction

To determine or identify the parameters of the biogeochemical source-minus-sink terms in a given marine ecosystem model is still a challenging task in ocean modeling, especially when a spatially three-dimensional ocean circulation is considered as forcing. Since there is even no agreement upon what is *the* correct ecosystem model or model structure, an assessment of the different models highly depends on their validation against given observational data. This validation process intrinsically requires parameter optimization runs to estimate the model's capability in representing the data and thus to be appropriate and valid for prognostic simulations.

Solving this nonlinear optimization problem, whether deterministic (e.g., gradient-based) or stochastic (e.g., metaheuristics) algorithms are used, typically requires a large number of expensive objective function evaluations, which translates into prohibitively high computational cost. Straightforward attempts by employing the model under consideration directly in an optimization loop using conventional optimization algorithms is thus often tedious or even infeasible. As a consequence, methods that accelerate either the optimization process itself or the underlying simulation are highly appreciated.

With respect to the latter, the *Transport Matrix Method (TMM)* introduced in Khatiwala et al. (2005) allows to compute the distribution of biogeochemical tracers for a given climatological ocean circulation. Pre-computed by an ocean model, the discretized diffusion and advection operators are stored in the transport matrices which can be used later on for an *offline* computation of the biogeochemical tracers. Consequently, using the TMM all biogeochemical tracers are regarded to be passive.

When applied to simulate a steady annual cycle in the marine ecosystem, the TMM can be efficiently used in a spin-up simulation (i.e., a pseudo-time stepping or fixed point iteration) or using Newton's method, as described for example in Khatiwala (2008). Using sophisticated numerical libraries and parallelization strategies, a flexible environment for simulation of a whole class of biogeochemical models has been developed, see Piwonski and Slawig (2010, 2011).

But still, with this numerical acceleration technique, a parameter optimization run is very time-consuming even on high performance computers when directly applying standard methods of nonlinear optimization. In this paper, we present a strategy called *Surrogate-based Optimization (SBO)* that significantly reduces the computational effort of a parameter optimization run. The method is widely used in engineering applications (see, among others, Bandler et al. (2004); Forrester and Keane (2009); Leifsson and Koziel (2010); Queipo et al.

*Corresponding author (*phone:* +49-(0)431 880 7452, *fax:* +49-(0)431 880 7618)
Email addresses: mpr@informatik.uni-kiel.de (M. Prieß), jpi@informatik.uni-kiel.de (J. Piwonski), koziel@ru.is (S. Koziel), ts@informatik.uni-kiel.de (T. Slawig)

[1] Research supported by DFG Cluster The Future Ocean

(2005)). One SBO approach employing a response corrected physical coarse model has already been successfully applied on parameter optimization in a biogeochemical model of NPZD type in a single water column, see Prieß et al. (2011c,b). The reduction in computation time achieved there was up to 95%. Motivated by this results, we now investigate SBO for a spatially three-dimensional marine ecosystem model.

The idea of the surrogate-based optimization method is to replace the original (in the SBO framework also called *fine* or *high-fidelity*) model in the optimization process by a computationally cheaper but yet reasonably accurate representation, the surrogate. In our application, the term *model* refers to the complete coupled marine ecosystem model consisting of the precomputed ocean circulation represented by the transport matrices and the biogeochemical source-minus-sink terms. When looking for a steady annual cycle, one model run thus means a spin-up of the coupled ecosystem model into a steady cycle (at least up to a desired numerical accuracy).

The surrogate can be created by approximating sampled fine model data (so-called *function-approximation surrogates*, see Queipo et al. (2005); Simpson et al. (2001); Smola and Schölkopf (2004)) or by employing a physics-based *low-fidelity* or *coarse* model, a computationally cheap but less accurate representation of the fine model. The latter approach is used in this paper.

Here, at each iteration of the SBO algorithm, the surrogate is build from the physics-based coarse model and, in order to improve its accuracy, by updating its output/response using information from the fine model and a special alignment or correction procedure. A next iterate, a prediction of the fine model optimum is obtained by optimizing this surrogate. This optimization process only incorporates evaluations of the computationally much cheaper coarse model. This process of updating the surrogate and subsequent optimization is iterated in order to keep the surrogate close to the original fine model and locate the fine model optimum as precisely as possible while, most importantly, the overall computational effort is ideally retained small. Since every iteration step of the proposed SBO algorithm just requires *one* fine model evaluation, the SBO approach can be very efficient in terms of total number of function evaluations of the fine model necessary to yield a reasonably accurate optimum, when compared to a direct fine model optimization.

In this paper, we show how the SBO method using a truncated spin-up to construct a coarse model and a special multiplicative alignment or correction technique to obtain the surrogate leads to a significant reduction in overall time for parameter optimization in a exemplary marine ecosystem model, namely the N-DOP model described in Kriest et al. (2010). To prove feasibility of the method, we used synthetic, model-generated twin data and performed illustrative optimization runs. In this application, the total reduction of optimization cost achieved by the SBO method is again higher than 90% while still yielding a reasonable optimum. We have thus successfully transferred the SBO method from the engineering context to an application in ocean modeling and parameter identification.

Subsequent steps will include enhancement of the present approach by coarse/fine model sensitivity, which would allows us to locate fine model optimum more accurately. The trade-offs between the solution accuracy and the extra computational overhead related to sensitivity calculation will be investigated. Furthermore, the application to real measurement data will be necessary to demonstrate the full capabilities of this approach. This has already been successfully done for the one-dimensional NPZD model in Prieß et al. (2011a). Since the used software framework in Piwonski and Slawig (2011) is rather flexible with respect to the choice of the biogeochemical model, an application of the SBO method to other, also more complex ecosystem models is computationally tractable.

The structure of the paper is as follows: In Section 2, we briefly describe the general structure of marine ecosystem models, the N-DOP model considered as an example in this paper, and the numerical solution method based on the TMM. A general formulation of the parameter identification or optimization problem is presented in Section 3. Section 4 provides a short overview on surrogate-based optimization. To show how an appropriate low-fidelity or coarse model can be chosen, we investigate distinct coarse models in Section 5. Our actual choice and the multiplicative response correction technique which we employ in this paper are motivated in Section 6. Section 7 provides an initial validation of the proposed approach which was performed prior to a full optimization run. The details of the SBO setup are described in Section 8, and numerical results and discussion of our SBO runs are presented in Section 9. Section 10 concludes the paper with a summary and an outlook.

2. Model description

Marine ecosystem models mainly consist of two parts, namely the ocean circulation and the biogeochemical model. For the investigation of the oceanic part of the global carbon cycle, especially the latter is subject of current research. Ocean biota plays an import role therein, and due to its complex and partly unknown organic and inorganic cycles a formulation of a comprehensive biogeochemical model is rather difficult. Here parameter identification and optimization helps assessing biogeochemical models in a satisfactory manner.

2.1. Coupled marine ecosystem models

A fully coupled marine ecosystem model is a system of equations modeling the ocean circulation including temperature and salinity distributions coupled to equations governing transport and reaction of biogeochemical tracers. The coupling reflects the fact that tracer concentrations are advected by the ocean circulation, their diffusion is dominated by the turbulent mixing of marine water, and vice versa a tracer concentration may effect the ocean circulation. A fully coupled simulation is computationally expensive since the simulation of both systems must be performed simultaneously. A single model evaluation in three space dimensions can be performed on high-performance computers only, even more if steady annual cycles – whose simulation requires long-term spin-ups – are looked for.

In contrast, an *off-line* model or computation is a simplified approach for tracers that are (or are regarded as) *passive*,

i.e., they do not affect the ocean physics or this influence is neglected. This results in a one-way coupling only, namely from the ocean circulation to the tracer dynamics, i.e., the pre-computed circulation data enter the tracer transport equations as forcing terms. These are the advection velocity vector field $\mathbf{v} = \mathbf{v}(\mathbf{x}, t)$, mixing coefficient $\kappa = \kappa(\mathbf{x}, t)$, temperature and optionally salinity. Here $(\mathbf{x}, t)$ denotes a point in the space-time cylinder $\Omega \times [0, T]$ with $\Omega \in \mathbb{R}^3$ being the spatial domain (i.e., the ocean) with boundary $\Gamma = \partial\Omega$ and $[0, T], T > 0$, the time interval.

With this data given, the marine ecosystem model considered in an offline computation is the following system of parabolic partial differential equations (here for n tracers, with $y_i = y_i(\mathbf{x}, t)$):

$$\frac{\partial y_i}{\partial t} = \nabla \cdot (\kappa \nabla y_i) - \nabla \cdot (\mathbf{v} y_i) + q_i(y_1, \ldots, y_n, \mathbf{u}), \quad i = 1, \ldots, n \quad (1)$$

Additionally, homogeneous Neumann boundary conditions on Γ for all tracers y_i are imposed. The source-minus-sink terms q_i in general are nonlinear and represent growth, dying, and reaction models. Usually each of the q_i depends on several other tracers, reflecting the coupling between them. Here, we neglect the additional dependency on the space and time coordinates in the notation for brevity. The q_i also include the model parameters (as growth and dying rates, sinking velocities etc.) that are subject to identification. They are spatially and temporally constant and summarized in the vector $\mathbf{u} \in \mathbb{R}^m$.

2.2. The N-DOP model as a biogeochemical model example

For our example application of the parameter identification using the SBO method we choose the N-DOP model as the biogeochemical part of the considered marine ecosystem. The model incorporates two tracers, namely phosphate (nutrients, N) and dissolved organic phosphorus (DOP) denoted by y_N and y_{DOP} respectively. We give here a short model description and basically use the notation of Kriest et al. (2010), where more details can be found. The vertical coordinate is denoted by x_3 here.

In the model, the biological production (the net community productivity) is calculated as a function f of nutrients and light I. The production is limited using a half saturation function, also known as Michaelis-Menten kinetics, and a maximum production rate parameter α as

$$f(y_N, I) = \alpha \, \frac{y_N}{y_N + K_N} \, \frac{I}{I + K_I} \, .$$

Light, here, is a portion of short wave radiation I_{SWR}, which is computed as a function of latitude and season following the astronomical formula of Paltridge and Platt (1976). The portion depends on the photo-synthetically available radiation σ_{PAR}, the ice cover σ_{ice} and the exponential attenuation of water

$$I = I_{SWR} \, \sigma_{PAR} \, (1 - \sigma_{ice}) \, \exp(-x_3 \, K_{H2O}) \, .$$

A fraction of the biological production σ remains suspended in the water column as dissolved organic phosphorus, which remineralizes with a rate λ. The remainder of the production sinks as particulate to depth where it is remineralized according to the empirical power law relationship determined by Martin et al. (1987). Similar modeling of biological production can be found in Dutkiewicz et al. (2005); Parekh et al. (2005); Yamanaka and Tajika (1997).

Moreover the model formulation consists of a production (sun lit, euphotic) zone, with a depth of l', and a noneuphotic zone, Ω_1 and Ω_2 respectively. The source-minus-sink terms read

$$q_N(y_N, y_{DOP}, \mathbf{u}) = \begin{cases} -f(y_N, I) + \lambda \, y_{DOP} & \text{in } \Omega_1 \\ \bar{\sigma} \, \frac{\partial}{\partial x_3} F(y_N, I) + \lambda \, y_{DOP} & \text{in } \Omega_2 \end{cases}$$

$$q_{DOP}(y_N, y_{DOP}, \mathbf{u}) = \begin{cases} \sigma \, f(y_N, I) - \lambda \, y_{DOP} & \text{in } \Omega_1 \\ -\lambda \, y_{DOP} & \text{in } \Omega_2 \end{cases}$$

where

$$F(y_N, I) = (x_3 / l')^{-b} \int_0^{l'} f(y_N, I) \, dx_3 \, .$$

The parameters to be identified are summarized in the vector $\mathbf{u}$. They are are given in Table 1.

2.3. Transport matrices

In this section we briefly describe the usage of the *Transport Matrix Method (TMM)* introduced in Khatiwala et al. (2005) which significantly accelerates an offline tracer transport simulation. Since in this case the ocean circulation data is only used as pre-computed input for the tracer transport equations (1), the spatial differential operators therein can be represented as a linear operator and the equations can be formally written as

$$\frac{\partial y_i}{\partial t} = A(\kappa, \mathbf{v}) \, y_i + q_i(y_1, \ldots, y_n, \mathbf{u}), \quad i = 1, \ldots, n, \quad (2)$$

where $A(\kappa, \mathbf{v})$ is the (time-dependent) linear operator comprising the whole transport, i.e., diffusion and advection, for the given ocean circulation data κ and $\mathbf{v}$. Note that A is identical for all tracers if the molecular diffusion of the tracers is small compared to the turbulent mixing, which is a reasonable simplification.

Instead of using the ocean circulation data $\kappa, \mathbf{v}$ itself and discretizing the corresponding diffusion and advection operators in the tracer transport simulation, the TMM builds up a certain number of temporally averaged matrices A using an ocean model, i.e., with its numerical diffusion and advection scheme. This is done basically by applying the diffusion and advection parts of the ocean model to special kinds of tracer basis functions, thus building up the matrix A columnwise. Since normally in ocean models an operator splitting between explicit and implicit parts is used to guarantee stability, the TMM also builds two sets of matrices, one for either part of the numerical scheme. Note that hence also the application of the implicit part of the scheme is now performed just by a matrix-vector multiplication (namely with the implicit transport matrix) instead of solving a linear system. A comprehensive discussion of the temporal and spatial discretization as well as the pro-

Table 1: Element in parameter vector, variable name, description and units for the N-DOP model parameters.

u_i	Name	Description	Unit
u_1	λ	remineralization rate of DOP	$1/d$
u_2	α	maximum community production rate	$1/d$
u_3	σ	fraction of DOP, $\bar{\sigma} = (1 - \sigma)$	–
u_4	K_N	half saturation constant of N	$m\,molP/m^3$
u_5	K_I	half saturation constant of light	W/m^2
u_6	K_{H2O}	attenuation of water	$1/m$
u_7	b	sinking velocity exponent	–

cess of evaluating transport matrices, especially in combination with operator splitting schemes can be found in Khatiwala et al. (2005). For our results we used twelve implicit and twelve explicit transport matrices, which represent monthly averaged diffusion and advection. The matrices are interpolated linearly to the corresponding discrete time step during simulation.

We now introduce a time discretization for (2) and denote by $\mathbf{y}_j$ the appropriately arranged vector of the values $(y_i(\mathbf{x}_k, t_j))_{i,k}$ of all n tracers on all spatial grid points $\mathbf{x}_k \in \Omega$ at the time step j. In the same way, we denote the vector of the discretized source-minus-sink terms q_i at all spatial grid points $\mathbf{x}_k$, evaluated at fixed time t_j, by $\mathbf{q}_j(\mathbf{y}_j, \mathbf{u})$. Using the TMM and for simplicity a fixed time step τ, the time integration scheme for (2) now reads

$$\begin{aligned} \mathbf{y}_{j+1} &= \mathbf{A}_{imp,j}\,(\mathbf{A}_{exp,j}\,\mathbf{y}_j + \tau\,\mathbf{q}_j(\mathbf{y}_j, \mathbf{u})) \\ &=: \varphi_j(\mathbf{y}_j, \mathbf{u}), \qquad\qquad j = 0, \ldots, n_\tau - 1\,. \end{aligned} \tag{3}$$

Here n_τ is the total number of time steps and $\mathbf{A}_{imp,j}, \mathbf{A}_{exp,j}$ are the implicit and explicit transport matrices at time step j. The matrices are block-diagonal and usually sparse, depending on the used numerical scheme of the ocean model. Starting from a vector $\mathbf{y}_0$ of initial values for the tracers, each step in the time integration scheme to solve the tracer transport equations (1) just consists of the evaluation of the source-minus-sink term and two matrix-vector multiplications.

2.4. Computation of a steady annual cycle

In our exemplary application, we use precomputed ideal or synthetic data denoted by $\mathbf{y}_d$ that have been generated by running the model into a (up to a certain numerical threshold) steady annual cycle. A model run in the optimization process thus means to compute a periodic solution of the discretized system (3) with a given fixed period of one year. Setting the end point T of the considered time interval to one year, we are looking for a fixed point of the mapping

$$\mathbf{y}_{n_\tau} = \Phi(\mathbf{y}_0, \mathbf{u}),$$

where $\Phi := \varphi_{n_\tau - 1} \circ \cdots \circ \varphi_0$ with the φ_j defined in (3), i.e., for a trajectory $(\mathbf{y}_j)_{j=0,\ldots,n_\tau}$ with

$$\mathbf{y}_{n_\tau} = \Phi(\mathbf{y}_0, \mathbf{u}) = \mathbf{y}_0\,. \tag{4}$$

In this setting one application of the mapping Φ corresponds to the computation of one year model time. Thus we will also refer to a period as a *model year* in the following. In the sequel we set the number of steps per year to $n_\tau = 45$. Assuming 360 days a year this time step corresponds to 192 hours. Both, the time step and the step count is kept fixed for our analysis and hence is not explicitly specified again.

The whole fixed point iteration now consists of a repeated application of the mapping Φ, i.e., we set

$$\mathbf{y}^{l+1} = \Phi(\mathbf{y}^l, \mathbf{u}), \quad l = 0, \ldots, n_l - 1, \tag{5}$$

where n_l is the total number of iterations (model years) necessary to compute a steady annual cycle and $\mathbf{y}^l$ denotes the vector of discretized tracer after l years, i.e., $\mathbf{y}^l := \mathbf{y}_{l \cdot n_\tau}$. The iteration starts with a constant distribution $\mathbf{y}^0$ of all tracers.

It is implemented as part of the simulation package of Metos3D (Marine Ecosystem Toolkit for Simulation and Optimization in 3-D), see Piwonski and Slawig (2011). From several computations it can be observed that after $n_l = 3000$ iterations (model years), a numerical steady solution (up to an accuracy of more than 10^{-2} in Euclidean norm, compare Figure 1) is obtained. Thus we refer to this as a converged steady annual cycle and take it as the *reference high-fidelity (or fine) model* output/response.

As shown in Figure 1, the residual in the solution of (4) can be further decreased by using a higher number n_l of model years used in the fixed point iteration (5). However, the number $n_l = 3000$ of steps (already used for example in Kriest et al. (2010)) provides a satisfactory accuracy.

We add the subscript f to distinguish the fine model state and corresponding number of model years, i.e., $\mathbf{y}_f, n_{f,l}$, from the corresponding coarse model ones.

3. The optimization problem

In order to identify the parameters in the biogeochemical model, we solve the following nonlinear optimization problem with given data $\mathbf{y}_d$:

$$\min_{\mathbf{u} \in U_{ad}} J(\mathbf{y}(\mathbf{u})), \tag{6}$$

where

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

$$\mathbf{b}_l, \mathbf{b}_u \in \mathbb{R}^m, \quad \mathbf{b}_l < \mathbf{b}_u,$$

where $\mathbf{y}(\mathbf{u})$ denotes a discrete solution of (4) for given parameters $\mathbf{u}$ and where m denotes the total number of model parameters to be identified (in this paper we have $m = 7$). The inequalities in the definition of the set U_{ad} of admissible parameters are meant component-wise. The functional J may additionally include a regularization term for the parameters which was not necessary for the application considered in this paper. Additional constraints on the state variable of tracers, $\mathbf{y}$, might also be necessary, e.g., to ensure non-negativity. This was also not necessary in this application, negative values were treated as zero. The norm Y in the cost function is a (maybe weighted) Euclidean vector norm where optional weights may be taken as the inverse values of the variances of the measurements. Since we use synthetic data in this paper, such variances have not been applied.

4. Surrogate-based optimization

Solving nonlinear optimization problems where computation of the objective function involves time consuming computer simulations may be quite challenging. The fundamental bottleneck is that 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, which translates into prohibitively high computational cost. For such problems, where optimization of complex three-dimensional climate models is a representative example, development of methods that would reduce the number of expensive simulations necessary to yield a satisfactory solution becomes critical.

4.1. Surrogate-based optimization: overview

Computationally efficient optimization of expensive simulation models can be realized using surrogate-based optimization (SBO), see Bandler et al. (2004); Forrester and Keane (2009); Leifsson and Koziel (2010); Queipo et al. (2005). The principal idea of SBO is to replace the direct optimization of the original (fine) model $\mathbf{y}_f$ by iterative updating and re-optimization of its surrogate model, which is a computationally cheap and yet reasonably accurate representation of $\mathbf{y}_f$. Typically, the SBO algorithm constructs, in each iteration k, the new/updated surrogate model $\mathbf{s}_k$ which is subsequently optimized to yield a prediction of the fine model optimum. The surrogate is created using available fine model data. For a well performing SBO algorithm the number of iterations is significantly smaller than for majority of conventional techniques, and each of these iteration normally involves just a few (in many cases only just one) evaluations of the high-fidelity model. This allows us to significantly reduce the overall optimization cost.

4.2. Surrogate model construction

Possible ways to create a surrogate include approximating sampled fine model data using a suitable technique, e.g., polynomial regression (Queipo et al., 2005), kriging (Simpson et al., 2001) or support-vector regression (Smola and Schölkopf, 2004). These so-called *function-approximation surrogates* do not inherit any physical characteristics of the original fine model – they are constructed without any particular knowledge of the system. They usually require a large amount of fine model evaluations so that their use to ad-hoc optimization may be questionable. On the other hand they are easily transferrable to other application areas.

Another possibility, which we explore in this paper, is to construct the surrogate from a physics-based low-fidelity (or coarse) model. Since the accuracy of the coarse model is usually not sufficient to directly replace the fine model in an optimization loop, it is often necessary to use suitable alignment/correction techniques to reduce the misalignment between the coarse and fine model responses. The specific correction technique exploited in this work is described in detail in Section 6. These so-called *physics-based surrogates* Søndergaard (2003), provided that the underlying coarse model is chosen properly, inherit the relevant physical characteristics of the original fine model so that only a few fine model data is necessary to ensure their good alignment with the fine model. As a results, SBO schemes working with physics-based surrogates normally require small number of fine model evaluations to yield a satisfactory solution. On the other hand, their transfer to other applications is less straightforward since the underlying coarse model and chosen correction approach is rather problem specific.

Possible ways to create a physics-based coarse model include using a coarser discretization (while employing the same simulation tool as for the fine model), simplified physics or different ways of describing the same physical phenomenon or even by using analytical formulas if available. Another straightforward way to construct a coarse model is to use a relaxed convergence criterion (or equivalently cutting down the number of iterations) used in the solution of the model state. This approach is used in this paper for the considered marine ecosystem model, where the model response is a steady annual cycle and a fixed point iteration is applied.

The surrogate model is updated at each iteration of the optimization algorithm, typically using available fine model data from the current and/or also from previous iterates. The particular surrogate model exploited in this paper is constructed by using the fine model response at the current optimization variable vector $\mathbf{u}_k$ only.

4.3. Updating solution, consistency conditions and SBO performance

In a typically SBO scheme, also exploited in this paper, the next iterate, $\mathbf{u}_{k+1}$, is obtained by optimizing the surrogate $\mathbf{s}_k$, i.e.,

$$\mathbf{u}_{k+1} = \operatorname*{argmin}_{\mathbf{u}\in U_{ad}} J(\mathbf{s}_k(\mathbf{u})) . \tag{7}$$

The process of aligning the coarse model to obtain the surrogate and subsequent optimization of this surrogate is repeated until a user-defined termination condition is satisfied, which can be based on certain convergence criteria, assumed level of cost function value or a specific number of iterations (particularly

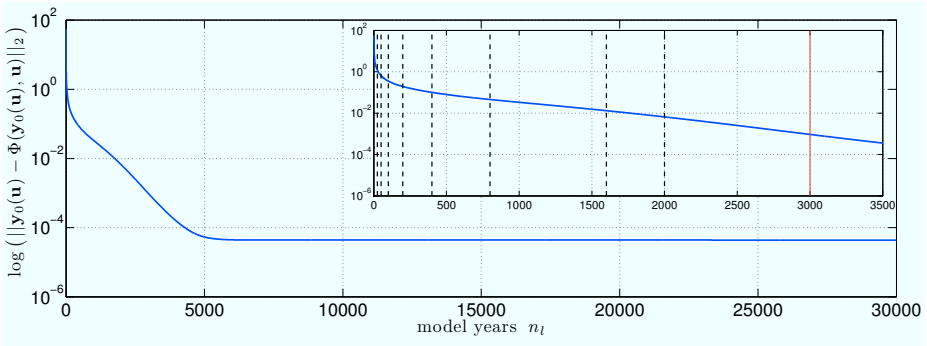


Figure 1: Convergence of the fix point iteration towards a solution $\mathbf{y}(\mathbf{u})$ of (4), for some illustrative parameter vector $\mathbf{u}$. Shown is the Euclidean norm of the residual (cf. Section 2.4). Inset: detailed section. In this paper, we consider a reduced number of fix point iterations (or, equivalently, number of model years n_l) to create a low-fidelity (or coarse) model. Initially considered coarse models are indicated by vertical dashed black lines whereas the solution after $n_l = 3000$ model years (vertical red line) is considered as the reference fine model solution.

if the computational budget of the optimization process is limited).

Key prerequisites to ensure that the SBO algorithm performs well, both in terms of low computational complexity and the quality of the final solution, are a cheap and yet reasonably accurate coarse model as well as a properly selected and low-cost alignment procedure (i.e., using a limited number of fine model evaluations, preferably just one).

Provided that the surrogate $\mathbf{s}_k$ satisfies so-called 0- and 1st-order consistency conditions with the original fine model $\mathbf{y}_f(\mathbf{u}_k)$ at the iterate $\mathbf{u}_k$, i.e, agreement between function values and the 1st-order derivatives at the current iteration point as

$$\mathbf{s}_k(\mathbf{u}_k) = \mathbf{y}_f(\mathbf{u}_k), \quad \mathbf{s}'_k(\mathbf{u}_k) = \mathbf{y}'_f(\mathbf{u}_k), \tag{8}$$

the surrogate-based scheme (7) is provable convergent to at least a local optimum of (6), under mild conditions regarding the coarse and fine model smoothness, and provided that the surrogate optimization scheme is enhanced by the trust-region (TR) safeguard, i.e.,

$$\mathbf{u}_{k+1} = \underset{\substack{\mathbf{u}\in U_{ad},\\ \|\mathbf{u}-\mathbf{u}_k\|\leq\delta_k}}{\operatorname{argmin}} J(\mathbf{s}_k(\mathbf{u})), \tag{9}$$

with δ_k being the trust-region radius updated according to the TR rules. We refer the reader to e.g. Conn et al. (2000); Koziel et al. (2010) for more details.

Ensuring the 1st-order consistency requires including fine/coarse model sensitivity which clearly increases cost for the coarse model alignment. Thus, for a given problem, the trade-offs between the solution accuracy and the extra computational overhead related to sensitivity calculation would have to be assessed.

By definition, the surrogate proposed in this paper satisfies 0-order consistency only. Formally, this is not sufficient to ensure the convergence of the surrogate-based scheme to a (local) minimum of the fine model optimization problem. However, as pointed out before, since the surrogate is physics-based, it inherits substantial knowledge about the fine marine ecosystem model under consideration and thus, its derivatives are expected to be at least similar to those of the fine model. Furthermore, because of being constructed from a physics-based coarse model, the surrogate exhibits quite good generalization capability, which means thats it provides a reliable approximation of the fine model when moving from one parameter vector to another. Numerical results presented in Section 9 further confirm this, demonstrating that the 0-order consistent surrogate is able to yield remarkably good results at the cost of a few evaluations of the fine model only.

5. Low-fidelity models

There are various ways to create a physics-based coarse model. Some straightforward methods include neglecting second order terms in the model equations, using simplified physics or different ways of describing the same physical phenomenon or even by using analytical formulas if available instead of performing simulation.

Climate models, and also the three-dimensional coupled marine ecosystem model that we consider here as a representative sub-class, are typically given as time-dependent partial differential or differential algebraic equations (PDE/PDEAs), compare Gill (1982); Majda (2003); McGuffie and Henderson-Sellers (2005). A straightforward way to introduce a coarse model for these types of models thus is to reduce the spatial or temporal resolution. When the model response is a steady stationary or periodic state and a fixed point iteration is applied as it is the case for the considered model in this paper (cf. Section 2.4), another way of constructing a coarse model is possible by reducing the number of iterations in this fix point iteration (or, equivalently, by employing a relaxed stopping criterion). The implementation of this approach is very straightforward and convenient, and it is the method of choice for this paper.

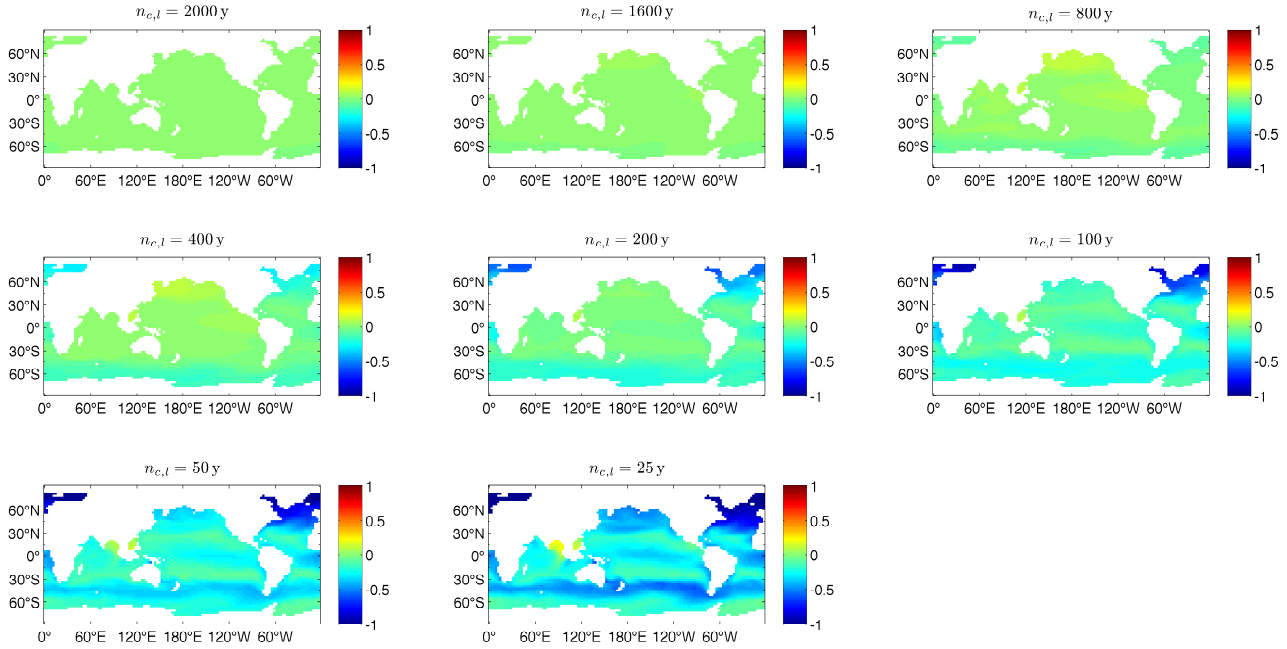


Figure 2: Difference in fine and coarse model responses, $\mathbf{y}_f - \mathbf{y}_c$, for illustratiion, here for the tracer N, for the uppermost depth layer and at some point in time. The coarse models are obtained by a reduced number of fix point iterations (or equivalently number of model years n_l) which are employed to solve for a steady annual cycle in (4). The reference fine model solution $\mathbf{y}_f$ is obtained with $n_{f,l} = 3000$ model years. Selected depth layer and time are representative for the overall model behavior.

5.1. Specific choice of a low-fidelity model

We follow the approach of a relaxed fix point convergence for the three-dimensional N-DOP model to obtain a coarse model employing a reduced number of fixed point iterations, or, equivalently, number of model years n_l, to solve for an approximation of the steady annual cycle (cf. Section 2.4). As noted before, the time step τ employed in the underlying time integration scheme (3) is fixed and the same as for the reference fine model. For the sake of further analysis, the coarse model state and number of iterations employed will be denoted by $\mathbf{y}_c$ and $n_{c,l}$, respectively to be distinguishable from those used for the fine model.

The evaluation time for the coarse models is, compared to the one for the fine model, reduced by the factor α_{eval} which is simply given as

$$\alpha_{eval} = (n_{f,l}/n_{c,l})\,. \tag{10}$$

Whereas the evaluation time for the fine model is several minutes on a 48-processor cluster, the time required for one coarse model evaluation could be significantly reduced to a few seconds if $n_{c,l}$ is sufficiently small.

For initial experiments, we consider distinct coarse models with various values of $n_{c,l}$, more epsifically

$$n_{c,l} = \{2000, 1600, 800, 400, 200, 100, 50, 25\}\,. \tag{11}$$

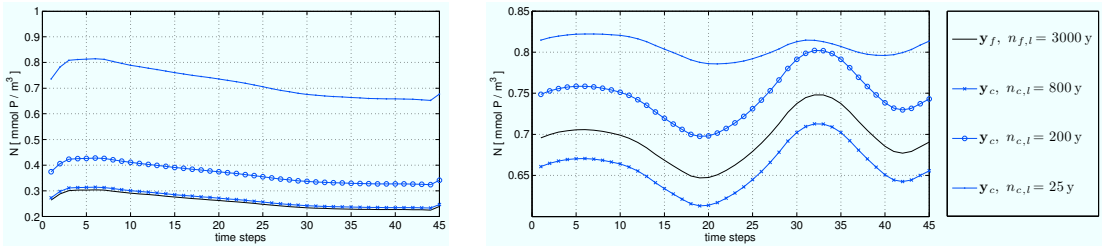
Figure 1 shows the convergence of the fix point iteration towards a solution $\mathbf{y}(\mathbf{u})$ of (4) as well as the obtained residual in the fix point iteration for the fine model and for the coarse models employing the distinct number of model years given above.

To further assess the quality of approximation of these different coarse models we compare their responses with the one of the reference fine model. For this purpose, Figure 2 shows differences in the fine and coarse model responses for one illustrative tracer (here, N) and some point in time and at the uppermost depth layer.

Note that Figure 2, for illustration, shows one selected tracer for one chosen point in time in the whole time interval (here, one year) and at one chosen depth layer only. The total number of depth layers is 15 and the entire discrete time scale is 45 so that it is impossible to present a full model response here. We emphasize that shown responses are "representative" which means that the qualitative behavior of the responses under consideration is similar for the second tracer, other points in time and depth layers. This also holds for all subsequent plots shown in this paper and – for the sake of brevity – will not be mentioned explicitly again.

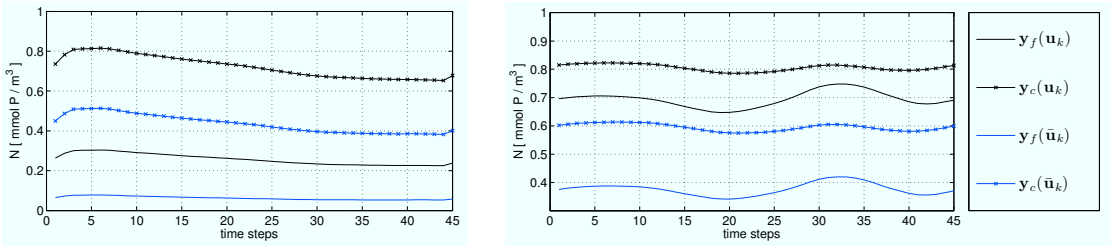
It can be observed that the differences between the fine and coarse model response become quite noticeable for the coarse models with $n_{c,l} \leq 200$. This is confirmed in Figure 3 which shows the entire trajectories at selected spatial locations for the corresponding model responses.

It can be seen that more or less all coarse model responses share the relevant characteristics of the fine model one such as local minima and maxima. Clearly, with decreasing number of $n_{c,l}$, the accuracy of the corresponding coarse model response decreases accordingly. However, even with $n_{c,l} = 25$, the coarse model response still accounts for the main features of the fine



(a) $x = 30.9375°$W, $y = 30.9375°$N. (b) $x = 90°$E, $y = 0°$.

Figure 3: Fine and coarse model responses $\mathbf{y}_f, \mathbf{y}_c$, corresponding to Figure 2, here, for a whole trajectory (one year or, equivalently, 45 discrete time steps) and at two spatial locations. For the sake of better visibility, only the coarse model with $n_{c,l} = 3000, 800, 200, 25$ model years are shown.



(a) $x = 30.9375°$W, $y = 30.9375°$N. (b) $x = 90°$E, $y = 0°$.

Figure 4: Fine and coarse model responses $\mathbf{y}_f, \mathbf{y}_c$ (here, for the sake of brevity, only the one with $n_{c,l} = 25$ is shown) for the illustrative tracer N, at the uppermost depth layer, a whole trajectory and at two spatial locations. Shown are the responses at a reference design $\mathbf{u}_k$, and a neighboring point $\bar{\mathbf{u}}_k$ in order assess their qualitative relation and to choose a suitable correction approach.

model one.

The coarse models with $n_{c,l} \leq 200$ seems to be sufficiently accurate while, at the same time, sufficiently cheap. In the following analysis, we will thus concentrate on these models only.

6. Surrogate construction

The surrogate model utilized in this paper falls into the category of physics-based models (cf. Section 4) as it is constructed from the underlying coarse model which is in turn based on the same model equations as the fine one.

As motivated in Section 4, those surrogates, if the underlying coarse model is chosen properly, inherit the relevant physical characteristics of the fine model, so that a reasonable accuracy can be obtained by applying a suitable correction while using a limited number of fine model data. For the same reason, the generalization capability of physics-based models is typically very good, which is in contrast to function-approximation surrogates. Below, we motivate the choice of one specific coarse model and corresponding correction approach.

6.1. Choice of a low-fidelity model and correction approach

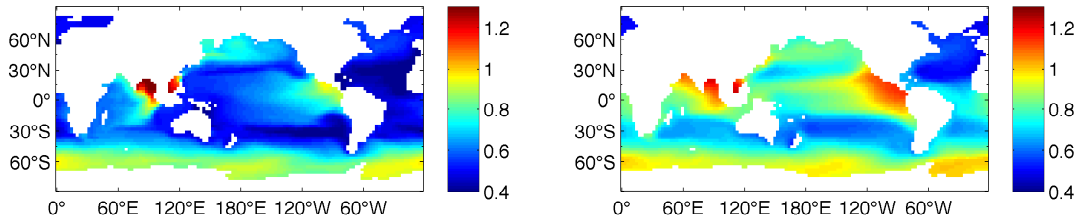
As described in Section 4, the surrogate is established at each iteration k of the SBO optimization loop. The surrogate is set up at the parameter vector $\mathbf{u}_k$ being the outcome of the previous iteration, using the coarse model response, the fine model response at $\mathbf{u}_k$, and a suitable correction technique.

In order to select a suitable correction method of the coarse model responses in focus (i.e., the ones with $n_{c,l} \leq 200$ model years as was motivated in the last section), we investigate the fine and coarse model responses at a randomly selected reference point $\mathbf{u}_k$ and its neighborhood, represented by another randomly selected point $\bar{\mathbf{u}}_k$. Here, $\| \bar{\mathbf{u}}_k - \mathbf{u}_k \| \approx 6$, i.e., $\bar{\mathbf{u}}_k$ lies in a rather close vicinity of $\mathbf{u}_k$.

Figure 4 shows the fine and coarse model responses $\mathbf{y}_f$ and $\mathbf{y}_c$ for the same illustrative tracer and spatial locations as in Figure 3, at the reference and neighboring point $\mathbf{u}_k$ and $\bar{\mathbf{u}}_k$. For the sake of brevity, we only show the coarse model responses with $n_{c,l} = 25$. The qualitative behavior for the other coarse models, i.e., with $n_{c,l} = 50, 100, 200$ looks similar. Actually, their accuracy is even higher (see previous Section for details), i.e., their responses lie even closer to the one of the fine model.

It can be seen that the overall "shape" of the coarse model response resembles that of the fine one. Furthermore, the qualitative relation of the fine and coarse model response is rather well preserved (at least locally) for the two selected parameter vectors. In particular, the high-value outputs for both models are corresponding to each other on the time scale, which is the consequence of the coarse model being physics-based. This even holds for the "coarsest" of the coarse models under consideration here, i.e. the one using $n_{c,l} = 25$, which indicates that this very model will be suitable to construct the surrogate.

The relationship between the fine and coarse model response indicates that the natural way of constructing the surrogate would be *multiplicative response correction*. More specifically,



(a) Uppermost depth layer.

(b) 5th depth layer.

Figure 5: Surface plots of the correction factors in $\mathbf{a}_k$, for one illustrative tracer (here, N), for selected depth layers, at some representative point in time and parameter vector $\mathbf{u}_k$. Here, and for the further analysis we focus on the coarse model with $n_{c,l} = 25$.

multiplicative response correction is a convenient way of adjusting the response level without distorting the tracer shapes while moving from one parameter vector to another. This technique has already been investigated and successfully applied to a one-dimensional marine ecosystem model in Prieß et al. (2011c), where a motivation for it was similar.

6.2. Surrogate model formulation

The surrogate at iteration k of the optimization process, $\mathbf{s}_k$ (cf. Section 4), is generated through a multiplicative correction of the coarse model response (see also Prieß et al. (2011c)). The *correction vector*, denoted as $\mathbf{a}_k$, is simply given as the point-wise division of the fine by the coarse model response at the point $\mathbf{u}_k$, i.e.,

$$\mathbf{a}_k := \frac{\mathbf{y}_f(\mathbf{u}_k)}{\mathbf{y}_c(\mathbf{u}_k)}, \quad k = 1, 2, \ldots . \tag{12}$$

Having computed the correction factors, summarized in the *correction vector* $\mathbf{a}_k$, the surrogate model is defined as

$$\mathbf{s}_k(\mathbf{u}) := \mathbf{a}_k \, \mathbf{y}_c(\mathbf{u}), \tag{13}$$

where the multiplication in (13) is again meant point-wise. Note that the surrogate model is constructed using just one evaluation of the fine model.

Occasionally, when using the surrogate as given in (12), it might occur a situation where the coarse model response is close to zero (and maybe even negative and/or a few magnitudes smaller than the fine one, which leads to large, possibly negative, entries in the corresponding correction vector $\mathbf{a}_k$). Resulting "spikes" appearing in the surrogate's response can be viewed, in a way, as numerical noise that slows down the algorithm convergence and makes the fine model optimum more difficult to locate. This has already been observed in Prieß et al. (2011b), where a different marine ecosystem model was considered.

In order to estimate a typical magnitude of the correction factors for the given coarse and fine model, we calculate the correction $\mathbf{a}_k$ at some randomly chosen parameter vector $\mathbf{u}_k$. Figure 5 shows 2D surface plots of the correction factors in $\mathbf{a}_k$ at this parameter vector $\mathbf{u}_k$, for one representative tracer (here, N), selected depth layers and for some point in time. It can be observed that this particular correction vector – at least for the given parameter vector $\mathbf{u}_k$ – does not contain any large (negative) entries.

Nevertheless, we apply some simple modifications (as have been proposed in Prieß et al. (2011b)) that allow us to eliminate any possible influence of the problems described above. These modifications do not require any extra computational overhead, and include: (i) upper bounds a_{ub} for the correction factors in $\mathbf{a}_k$,(ii) setting the fine and coarse model response values to zero (and the correction factor to one) if their values lie below a certain threshold δ, which is supposed to be of the order of the discretization error of the model. For the considered problem, we use $\delta = 5 \cdot 10^{-3}$.

The aforementioned modifications can be formally written as follows:

$$\begin{aligned}
(i)\ \mathbf{y}_c(\mathbf{u}_k) &= \begin{cases} 0; & \text{if } \mathbf{y}_c \le \delta \\ \mathbf{y}_c; & \text{else} \end{cases}, \\
(ii)\ \mathbf{y}_f(\mathbf{u}_k) &= \begin{cases} 0; & \text{if } \mathbf{y}_f \le \delta \\ \mathbf{y}_f; & \text{else} \end{cases}, \\
(iii)\ \mathbf{a}_k &= \begin{cases} 0; & \text{if } \mathbf{a}_k \ge a_{\text{ub}} \\ \mathbf{a}_k; & \text{else} \end{cases},
\end{aligned} \tag{14}$$

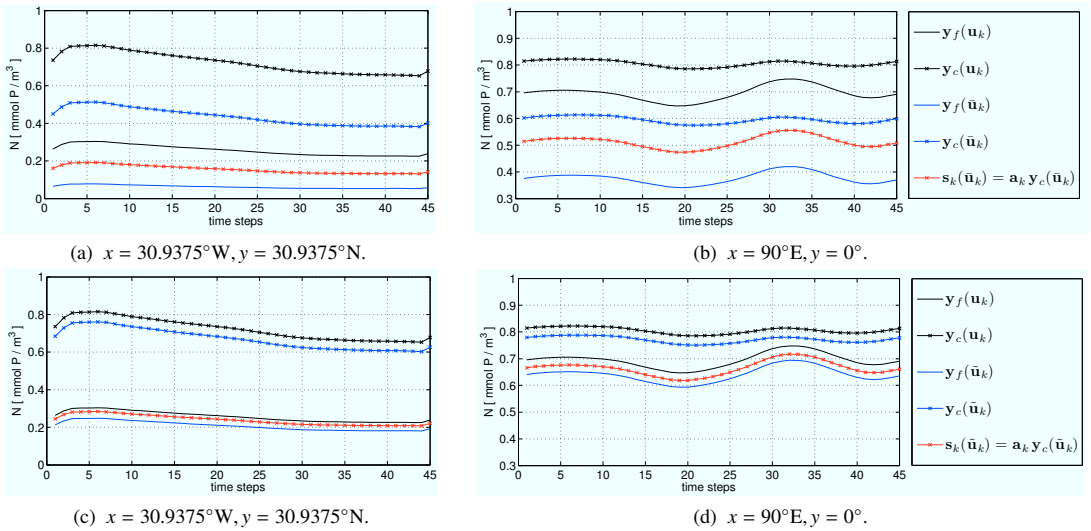
where the operations are again meant point-wise. These simple means can further improve the accuracy of the surrogate as well as the performance of the optimization algorithm, which has been investigated in Prieß et al. (2011b) for a similar response correction approach and another exemplary model. In the following we choose $a_{\text{ub}} = 5$ which, from numerical experiments, turned out to be a reasonable choice.

6.3. Consistency conditions

It should be noted that the surrogate model (12) satisfies, by definition, the 0-order consistency condition in (8) in the point of alignment $\mathbf{u}_k$, i.e.,

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

Our surrogate model does not use fine model sensitivity data. Hence, the 1st-order consistency condition in (8) cannot be satisfied exactly. Nevertheless, the surrogate exhibits quite good



(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 6: Fine, coarse model and surrogate's response $\mathbf{y}_f, \mathbf{y}_c$ and $\mathbf{s}_k$ for one illustrative tracer (here, N), at the uppermost depth layer, a whole trajectory and at two spatial locations. Shown are the responses at one representative "reference point" $\mathbf{u}_k$, at some point $\bar{\mathbf{u}}_k$ in the vicinity of $\mathbf{u}_k$ (6a, 6b) and at another closer neighboring point $\tilde{\mathbf{u}}_k$ (6c, 6d), in order to assess the generalization capability of the proposed surrogate. The surrogate's response at the reference point is omitted, since, by definition, the model alignment is perfect at this point.

generalization capability, which means that the surrogate provides a reasonable approximation of the fine one in the neighborhood of $\mathbf{u}_k$. As noted before, this is a result of the surrogate model being physics-based and since its derivatives are expected to be at least similar to those of the fine model (cf. Section 4.3).

The generalization capability will be analyzed in the next section in more detail. Furthermore, as demonstrated in Section 9, our SBO scheme exploiting the surrogate in (13) is able to yield remarkably good results, not only with respect to the quality of the final solution, but, most importantly, in terms of the low computational cost of the optimization process.

7. Initial validation

In order to validate the multiplicative response correction approach proposed in the last section, we analyze the generalization capability of the surrogate. More specifically, we check whether the surrogate provides a reasonable approximation of the fine model in a neighborhood of the "reference point" $\mathbf{u}_k$, i.e., the parameter vector where the surrogate is established. Recall, that the model alignment is perfect at $\mathbf{u}_k$ by definition. Thus, in the following we omit to show the corresponding responses in this reference point.

It should be noted that the fact of possible accuracy loss while moving away from the reference point is not a major concern with respect to the robustness of the surrogate-based optimization process (9). This is because the distance between the reference point and the updated parameter vector obtained by optimizing the surrogate will normally decrease upon convergence of the algorithm, either naturally or forcefully due to the reduction of the trust-region radius (cf. Section 4.3), which improves the accuracy of the surrogate model $\mathbf{s}_k$.

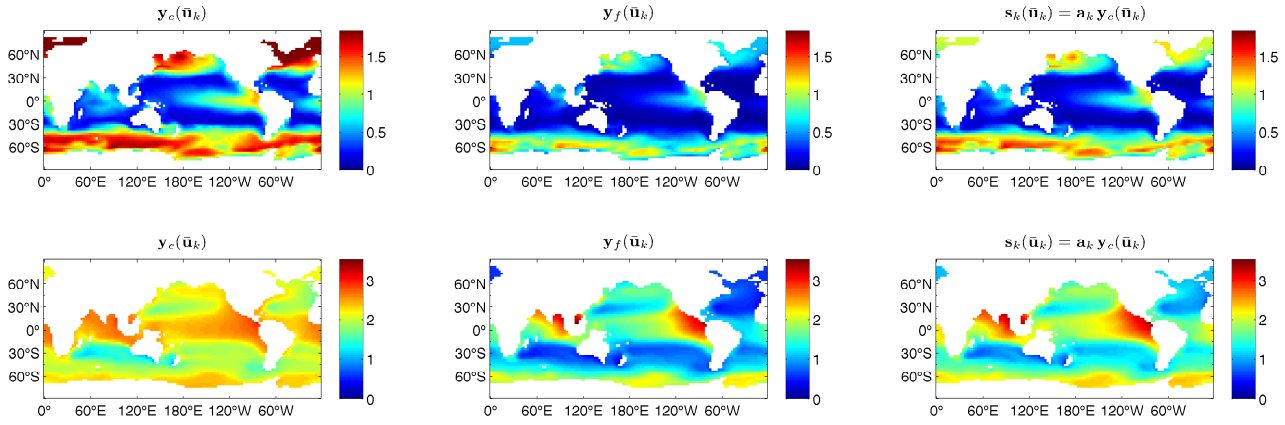
To analyze the properties addressed above we consider the same parameter vectors $\mathbf{u}_k$, and $\bar{\mathbf{u}}_k$ as in the last section (see also Figure 4) where we analyzed the qualitative relation of the fine and coarse model responses at different parameter vectors with $\| \bar{\mathbf{u}}_k - \mathbf{u}_k \| \approx 6$. Additionally, we consider another point $\tilde{\mathbf{u}}_k$ in a closer vicinity of $\mathbf{u}_k$, satisfying $\| \tilde{\mathbf{u}}_k - \mathbf{u}_k \| \approx 1$. Figures 6 shows the fine, coarse and surrogate's response at the reference point $\mathbf{u}_k$, its neighborhood $\bar{\mathbf{u}}_k$ (Figure 6a, 6b) and at $\mathbf{u}_k$ and the closer neighboring point $\tilde{\mathbf{u}}_k$ (Figure 6c, 6d). Shown are the model responses for the same illustrative tracer and spatial locations as in Figure 4.

It can be observed that the surrogate provides a reasonable approximation of the fine model also at the neighboring point $\bar{\mathbf{u}}_k$ whereas its accuracy is even increased at the closer point $\tilde{\mathbf{u}}_k$. As an additional evidence, we also present the corresponding model responses on the whole 2D spatial grid (i.e., with the vertical dimension z kept fixed), at two illustrative depth layers, some point in time and for one tracer (here, N) in Figures 7. Again, responses at the reference point $\mathbf{u}_k$ are omitted.

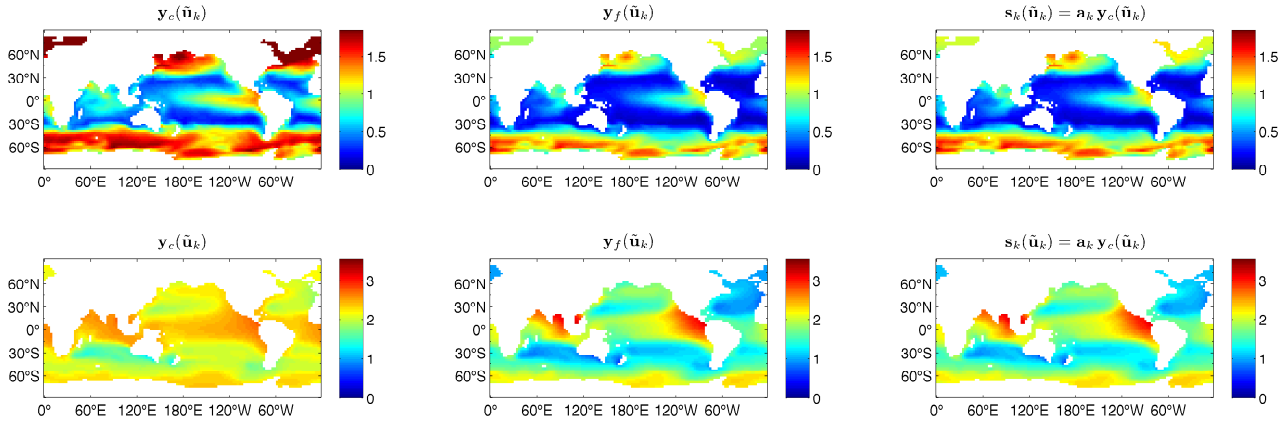
The qualitative validation carried out in this section confirms that the multiplicative response correction approach in conjunction with the coarse model under consideration (i.e., using $n_{c,y} = 25$) is a reasonable choice to construct a reliable surrogate. Exploiting the latter in a SBO seems very promising.

8. SBO – optimization setup

The optimization approach in this work has been verified using model-generated, attainable target data $\mathbf{y}_d$ which is obtained



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



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

Figure 7: Shown are – from left to right columns – the coarse, fine model and surrogate's response $\mathbf{y}_c, \mathbf{y}_f$ and $\mathbf{s}_k$, corresponding to Figure 6, here, for the uppermost and 5th depth layer and at some point in time. Again, responses at the reference point $\mathbf{u}_k$ are omitted.

by evaluating the fine model at some randomly chosen parameter vector, in the following denoted by $\mathbf{u}_d$, i.e.,

$$\mathbf{y}_d := \mathbf{y}_f(\mathbf{u}_d)\,. \tag{15}$$

For this "synthetic setup", since in this case the target or optimal parameter vector $\mathbf{u}_d$ is known, we can assess properties such as the parameter match obtained after employing SBO. This helps us to validate the applicability of our approach. Of course, in a usual setup, i.e., when considering real measurement data, the optimal parameters are unknown.

It is furthermore worth noticing that, for the considered problem, a direct fine model optimization would require immense computational effort (most likely several weeks) which, in practice, is a rather tedious process.

The solution of the surrogate-based optimization is compared to the target data by inspection of the corresponding fine model response and cost function values at this solution. The performance of the SBO process is assessed through investigating the accuracy of matching the target data by the final solution found by the algorithm as well as the computational costs. The latter is measured in terms of *equivalent* fine model evaluations (cf. Section 8.4 for details.)

The surrogate model is optimized using the MATLAB[1] function `fmincon`, exploiting the active-set algorithm.

8.1. *Cost function*

We define the following discrete cost function, measuring the difference between the discrete model response and the target in a squared Euclidean norm (cf. (6)) as

$$J(\mathbf{z}) := \|\mathbf{z} - \mathbf{y}_d\|_2^2 = \sum_{i=1}^{p} \left(z_j - (y_d)_j\right)^2, \tag{16}$$

with a general state vector $\mathbf{z}$ of dimension p and where $\mathbf{z}_j$ denotes the value (i.e., the concentration) of the state $\mathbf{z}$ at one

[1] MATLAB is a registered trademark of The MathWorks, Inc., http://www.mathworks.com

discrete (spatial and temporal) point and for one tracer. In this paper we particularly use a whole trajectory of the model solution in the cost function, which is one year (or, equivalently, 45 discrete time steps).

For SBO, we now consider the following optimization problem (cf. (9)):

$$\mathbf{u}_{k+1} = \underset{\substack{\mathbf{u}\in U_{ad},\\ \|\mathbf{u}-\mathbf{u}_k\|^2\le\delta_k}}{\operatorname{argmin}} J(\mathbf{s}_k(\mathbf{u})), \quad k=0,1,\ldots$$
$$U_{ad} := \{\mathbf{u}\in\mathbb{R}^m : \mathbf{b}_l \le \mathbf{u} \le \mathbf{b}_u\}, \quad \mathbf{b}_l, \mathbf{b}_u \in \mathbb{R}^m, \quad \mathbf{b}_l < \mathbf{b}_u, \tag{17}$$

where, for the example model treated in this paper we have $m = 7$ and where the specific initial and optimal parameter vector $\mathbf{u}_0$ and $\mathbf{u}_d$, the lower and upper parameter bounds $\mathbf{b}_l$ and $\mathbf{b}_u$ are explicitly given in Table 2. In (17), we enhance each optimization step by employing the trust-region convergence safeguard as introduced in Section 4. See the next subsection for further details.

At each iteration k of the optimization process (17), the surrogate is obtained by evaluating both the coarse and the fine model using the correction as given in (13) and the simple modification proposed in (14). The optimization of each surrogate $\mathbf{s}_k$ requires evaluations of the underlying (computationally cheap) coarse model only. For the considered approach, since we do not use fine model sensitivity data (cf. Section 6.3), the fine model has to be evaluated only once at the beginning of each iteration k in (17) which keeps the overall optimization cost low.

8.2. Trust-region convergence safeguards

As described in Section 4, the surrogate-based scheme (17) is provable convergent to at least a local optimum of (6), if the surrogate $\mathbf{s}_k$ satisfies 0- and 1st-order consistency conditions (Conn et al., 2000; Koziel et al., 2010) with the fine model at $\mathbf{u}_k$ (cf. (8)) and provided that both the coarse and the fine model are sufficiently smooth and the algorithm is enhanced by the trust-region (TR) safeguard as given in (9).

In (17), the trust-region radius δ_k is updated after each iteration, i.e., decreased if the new design was rejected or if the improvement of the fine model objective function was too small compared to the prediction given by the surrogate, or increased otherwise. We use classical updating rules (Conn et al., 2000; Koziel et al., 2010) with slightly modified parameters as

$$\delta_0 = 2, \quad \delta_k = \begin{cases} \delta_k/m_{\text{decr}}, & \text{if} \quad \rho_k < r_{\text{decr}} \\ \delta_k \cdot m_{\text{incr}}, & \text{if} \quad \rho_k > r_{\text{incr}} \end{cases},$$
$$r_{\text{incr}} = 0.75, \quad r_{\text{decr}} = 0.01, \quad m_{\text{incr}} = 3, \quad m_{\text{decr}} = 20, \tag{18}$$

where ρ_k denotes the *gain ratio* in iteration k defined as follows:

$$\rho_k := \frac{f_{\text{new}} - f_{\text{old}}}{s_{\text{new}} - s_{\text{old}}},$$
$$f_{old} := J(\mathbf{y}_f(\mathbf{u}_k)), \quad f_{new} := J(\mathbf{y}_f(\mathbf{u}_{k+1})),$$
$$s_{old} := J(\mathbf{s}_k(\mathbf{u}_k)), \quad s_{new} := J(\mathbf{s}_k(\mathbf{u}_{k+1})). \tag{19}$$

Note that values specified above are fairly standard except r_{decr} and m_{decr}; the first one is smaller than, whereas the latter is larger than usual. It was found out that these values are typically more suitable for surrogate-based optimization schemes working with physics-based surrogates.

As pointed out in Section 6.3, the surrogate defined in (12) does not satisfy the 1st-order consistency condition in (8) exactly. Still, applying a TR safeguard is reasonable, which is because the physics-based surrogate inherits substantial knowledge about the marine model under consideration so that its derivatives are expected to be at least similar to those of the fine model. Moreover, the accuracy of the surrogate model increases with the descrease of the TR radius as briefly validated in Section 7. Numerical results of an illustrative SBO run provided in the next section support the applicability of a TR safeguard, even without using fine model sensitivity data to ensure the 1st-order consistency exactly.

8.3. Stopping criteria

The process (17) of aligning the coarse model to obtain the surrogate and subsequent optimization of this surrogate is repeated until a user-defined termination condition is satisfied (cf. Section 4).

For the considered problem, we use the absolute step size (measured in the Euclidean norm) between two successive iterates $\mathbf{u}_k$ and $\mathbf{u}_{k-1}$ as well as a lower bound for the TR radius δ_k, in the following denoted by $\delta_k^{\min}$. In practive, we chosen a smaller bound for TR radius than for gamma because of the large value of m_{decr}.

We define the overall solution of the SBO – in the following denoted as $\mathbf{u}_s^*$ – as the design obtained when either the absolute step size is equal or below a given threshold γ or if the updated TR radius δ_k is below the given lower bound $\delta_k^{\min}$, i.e.,

$$\mathbf{u}_s^* := \left\{ \mathbf{u}_k \;\middle|\; \left(\|\mathbf{u}_k - \mathbf{u}_{k-1}\|^2 \le \gamma \right) \vee \left(\delta_k \le \delta_k^{\min} \right) \right\}. \tag{20}$$

It might not be necessary to run the SBO until convergence: an approximate solution might be sufficient as the surrogate model is not perfectly accurate anyway so that using a rather relaxed stopping criterion could allow us to obtain a sufficiently accurate solution at rather low computational cost.

To trade the quality of the solution obtained by SBO against the corresponding computational costs, we consider three distinct values for the threshold, more specifically

$$\{\gamma, \delta_k^{\min}\} = \{10^{-1}, 10^{-2}\}, \{10^{-2}, 10^{-3}\}, \{10^{-4}, 10^{-5}\}. \tag{21}$$

8.4. Optimization cost

This cost of the SBO is measured in terms of the total number of *equivalent* fine model evaluations. The cost for the evaluation of the coarse model is, compared to the cost for the corresponding fine model evaluation, reduced by the factor α_{eval} as given in (10), which is $\alpha_{eval} = 120$ for the selected coarse model under consideration. In other words, α_{eval} evaluations of the coarse model are equivalent to (or, as expensive as) one fine model evaluation. On the other hand, the cost of one iteration of the surrogate-based optimization procedure (17) equals to

Table 2: Initial, optimal and final parameters $\mathbf{u}_0$, $\mathbf{u}_d$ and $\mathbf{u}_s^*$ for the exemplary SBO run. Solution $\mathbf{u}_s^*$ of the SBO problem given in (17) is determined by the stopping criterion (20). The iterates $\mathbf{u}_2$, $\mathbf{u}_5$ and $\mathbf{u}_{10}$ correspond to thresholds $\gamma = 10^{-1}, 10^{-2}, 10^{-4}$.

iterate	u_1	u_2	...				u_7
$\mathbf{u}_0$	0.3	5.0	0.4	0.8	25	0.04	0.78
$\mathbf{u}_2$	0.502	3.328	0.633	0.845	24.886	0.036	0.92
$\mathbf{u}_5$	0.482	2.562	0.652	0.856	24.99	0.027	0.885
$\mathbf{u}_{10}$	0.485	2.334	0.659	0.745	25.076	0.025	0.864
$\mathbf{u}_d$	0.5	2.0	0.67	0.5	30.0	0.02	0.858
$\mathbf{b}_l$	0.25	1.5	0.05	0.25	10.0	0.01	0.7
$\mathbf{b}_u$	0.75	200.0	0.95	1.5	50.0	0.05	1.5

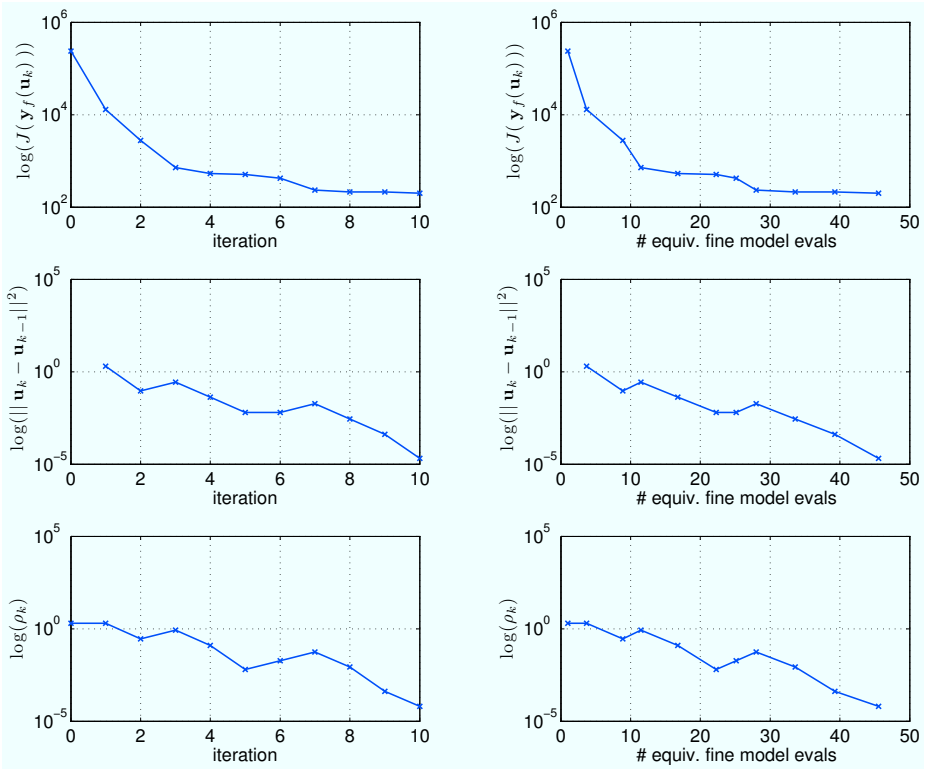


Figure 8: Convergence of the cost function value $J(\mathbf{y}_f)$ (cf. (16)), the step size norm and of the trust-region radius δ_k (both versus number of iterations and equivalent number of fine model evaluations) for an illustrative SBO run. *Updated* TR radii according to (18) and (19) and – for the sake of better visibility – semi-log plots are shown.

the number of coarse model evaluations necessary to optimize the surrogate model divided by this factor α_{eval}, and increased by one (since only one fine model evaluation is requires for the correction/alignment of the coarse model response).

9. SBO – results and discussion

The operation and performance of the proposed algorithm is illustrated through the results of an exemplary test run with the reference fine model as defined in Section 2.4, the coarse model using $n_{c,l} = 25$ and the correction approach as motivated in Section 6.1).

Figure 8 and 9 show corresponding convergence plots for the cost function value $J(\mathbf{y}_f)$ (cf. (16)), the squared step size norm, the trust-region radius δ_k (both versus number of iterations and equivalent number of fine model evaluations) and for the single parameter values $u_{k,i}$.

Figure 10 and 11 present the fine model response at the solution $\mathbf{u}_s^*$ of the SBO run, with $\mathbf{u}_s^*$ considered as one of the iterates $\mathbf{u}_2, \mathbf{u}_5$ and $\mathbf{u}_{10}$ corresponding to different values for the threshold γ used in the stopping criterion (cf. Section 8.3). Table 2 shows the corresponding parameter values.

It can be observed that a reasonably accurate solution, $\mathbf{y}_f(\mathbf{u}_2)$, i.e., a solution that is sufficiently close to the target $\mathbf{y}_d$, can be obtained after two iterations of the SBO, which corresponds to a termination condition employing a threshold of $\gamma = 10^{-1}$.

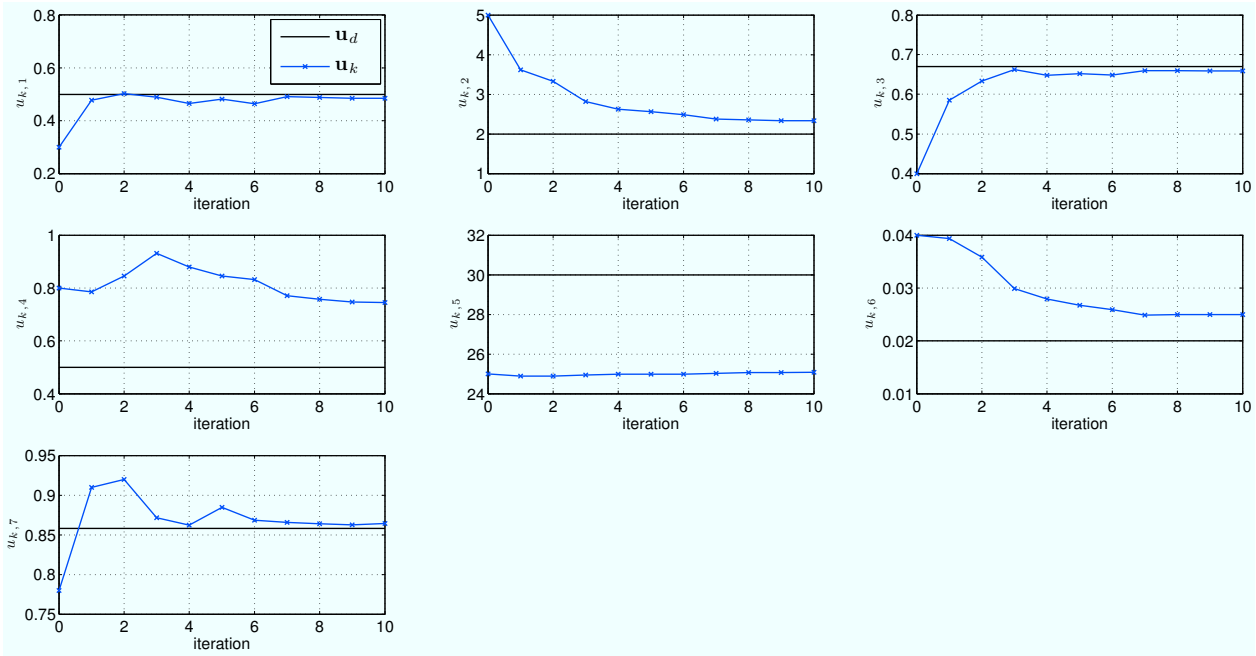
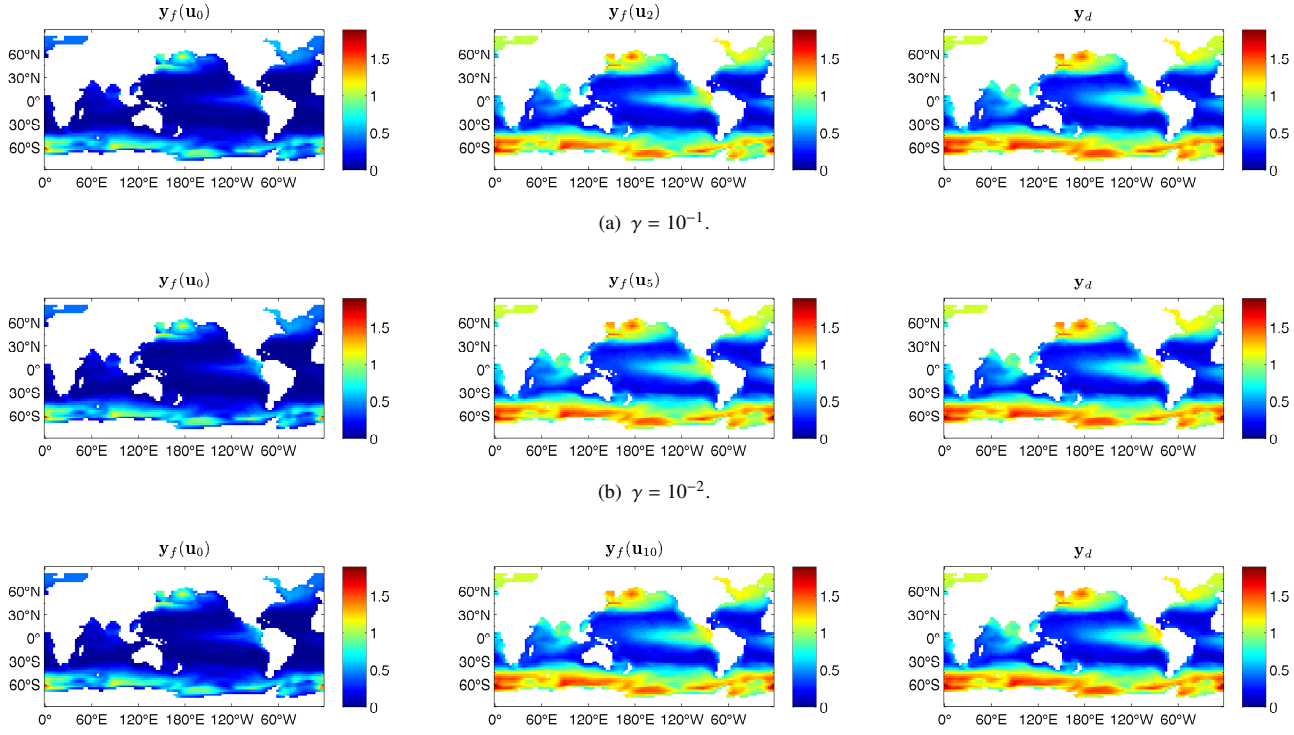


Figure 9: Convergence plots for the single parameter values $u_{k,i}$ for each iteration k of an illustrative SBO run.

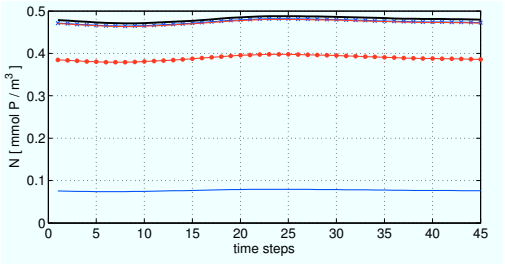


(a) $\gamma = 10^{-1}$.

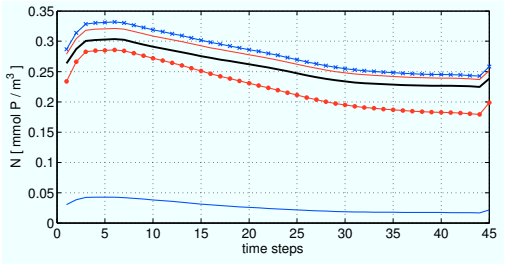
(b) $\gamma = 10^{-2}$.

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

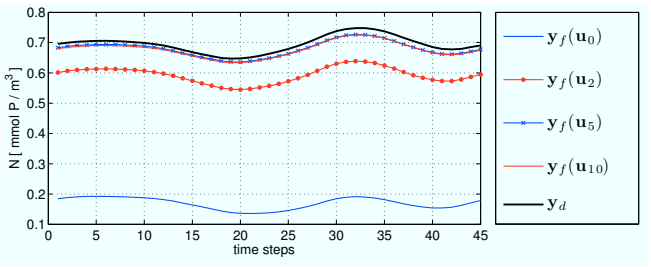
Figure 10: Solutions $\mathbf{y}_f(\mathbf{u}_2)$, $\mathbf{y}_f(\mathbf{u}_5)$ and $\mathbf{y}_f(\mathbf{u}_{10})$ obtained by an illustrative SBO run after two 10a, five 10b and ten 10c iterations – corresponding to different thresholds employed in the termination condition in (20). Shown are, from left to right, the fine model response $\mathbf{y}_f$ at the initial parameter vector $\mathbf{u}_0$, at the solution $\mathbf{u}_k$ and the target response $\mathbf{y}_d$, for one representative tracer (here, N), some point in time and at the uppermost depth layer.



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



(b) $x = 120.94°\text{W}, y = 30.94°\text{N}$.



(c) $x = 90°\text{E}, y = 0°$.

Figure 11: Shown are the solutions $\mathbf{y}_f(\mathbf{u}_2), \mathbf{y}_f(\mathbf{u}_5), \mathbf{y}_f(\mathbf{u}_{10})$, the initial response $\mathbf{y}_f(\mathbf{u}_0)$ and the target $\mathbf{y}_d$, corresponding to Figure 10, here, whole trajectories at three distinct locations.

After five iterations (corresponding to $\gamma = 10^{-2}$), both the quality of the solution $\mathbf{y}_f(\mathbf{u}_5)$ and the parameter match can be further increased. This is further supported by the obtained decrease in the corresponding cost function value as shown in Figure 8. A threshold of $\gamma = 10^{-4}$ which is reached after ten iterations in the SBO process, with a solution $\mathbf{y}_f(\mathbf{u}_{10})$, yields only slight improvements.

The crucial point is that these solutions by SBO can be obtained at the cost of a very few fine model evaluations only, more specifically 9, 22 and 46 equivalent fine model evaluations corresponding to approximately one or a few evaluations of the fine model gradient were required to obtain the solutions $\mathbf{u}_2, \mathbf{u}_5$ and $\mathbf{u}_{10}$.

A dramatic reduction of the optimization costs can be achieved by SBO. While direct optimization of the fine model was not performed, its cost can be estimated based on the number of model evaluations necessary to optimize the surrogate model, which is a few hundred to a thousand. Assuming approximately 30 minutes for a single fine mode evaluation on a 48-processor cluster, a direct optimization approach could hence 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).

As pointed out before, we do not use fine model sensitivity data in the definition of the surrogate (cf. Section 6). Thus, 1st-order consistency is not ensured exactly and the obtained accuracy in terms of parameter match for some parameters is clearly not perfect. This could possibly be because of low sensitivity of the model with respect to some of the parameters which has already been discovered in Rückelt et al. (2010) for another marine ecosystem model. A similar investigation for the N-DOP model in this paper would be useful. On the other hand, the solution is definitely sufficiently accurate in terms of matching the target output as indicated in Figure 10. However, by including fine model sensitivity, the algorithm might be capable of locating the solution even more precisely in terms of model parameters.

Concluding, SBO using a multiplicative response correction approach in conjunction with the proposed coarse model is able to yield a remarkably accurate solution – most importantly in terms of the quality of the final solution – whereas achieving significant cost savings from 93 up to 99% (estimated values). It is worth noticing, that yielding such cost savings will become even more important when considering for example a finer temporal and/or spatial resolution for the reference fine model if a more accurate fine model solution is considered as the reference.

10. Conclusions

Identification of model parameters in climate science can be computationally very expensive or even beyond the capabilities of modern computer power. A direct approach, where the high-fidelity (or fine) model under consideration is optimized to make its output match a given target data is often infeasible when such an identification process is performed using conventional optimization algorithms that require a large number of model evaluations. The development of faster methods becomes critical, particularly for handling complex three-dimensional models.

Surrogate-based optimization (SBO) addresses these issues by replacing the original fine model in the optimization loop by

a surrogate, its computationally cheap but yet reasonably accurate representation. In this paper we exploit physics-based surrogates constructed by suitable correction of a low-fidelity (or coarse) models. We demonstrate that a combination of properly selected low-fidelity model and simple, multiplicative response correction, allows us to build a reliable prediction tool that helps us locating the fine model optimum at a low computational cost.

As a case study, we considered parameter optimization for a selected representative of the class of global three-dimensional coupled marine ecosystem models, simulating the concentrations of phosphate (nutrients) and dissolved organic phosphorus. For this example problem, a physics-based coarse model is obtained from a truncated spin-up which is used to solve for a stead annual cycle. We demonstrate that the proposed multiplicative response correction approach can reduce the misalignment between the coarse and fine model leading to a reliable approximation of the original, fine ecosystem model.

We furthermore presented the results of an illustrative SBO run using model-generated, attainable target data to verify our approach. Solutions demonstrate that by iteratively updating and re-optimization of the proposed surrogate we are able to obtain a remarkably accurate solution at the cost of a very few evaluations of the fine model only. Whereas a direct optimization run using the fine model in a classical optimization loop would typically require several weeks on a 48-processor cluster, the computational cost using such a surrogate can be significantly reduced down to a few hours.

Enhancement of the present approach by coarse/fine model sensitivity would allow us to locate fine model optimum more accurately. Employing those improvements, an investigation of the trade-offs between the solution accuracy and the extra computational overhead related to sensitivity calculation will be crucial. Also, considering other ways of reducing the CPU cost of the coarse model such as coarsening the temporal and/or spatial discretization will be of great interest since this may further improve the cost savings when compared to the results presented here. Furthermore, the application to real measurement data will be necessary to demonstrate the full capabilities of this approach.

11. Acknowledgments

The authors would like to thank Andreas Oschlies, IFM Geomar, Kiel. This research was supported by the DFG Cluster of Excellence Future Ocean.

References

Bandler, J.W., Cheng, Q.S., Dakroury, S.A., Mohamed, A.S., Bakr, M.H., Madsen, K., Søndergaard, J., 2004. Space mapping: The state of the art. IEEE T. Microw. Theory. 52.

Conn, A.R., Gould, N.I.M., Toint, P.L., 2000. Trust-region methods. Society for Industrial and Applied Mathematics, Philadelphia, PA.

Dutkiewicz, S., Follows, M., Parekh, P., 2005. Interactions of the iron and phosphorus cycles: A three-dimensional model study. Global Biogeochem. Cycles 19, 1–22.

Forrester, A.I.J., Keane, A.J., 2009. Recent advances in surrogate-based optimization. Prog. Aerosp. Sci. 45, 50–79.

Gill, A.E., 1982. Atmosphere - Ocean Dynamics. volume 30 of *International Geophysics Series*. Academic Press.

Khatiwala, S., 2008. Fast spin up of ocean biogeochemical models using matrix-free newton-krylov. Ocean Modelling 23, 121–129.

Khatiwala, S., Visbeck, M., Cane, M., 2005. Accelerated simulation of passive tracers in ocean circulation models. Ocean Modelling 9, 51–69.

Koziel, S., Bandler, J.W., Cheng, Q.S., 2010. Robust trust-region space-mapping algorithms for microwave design optimization. IEEE T. Microw. Theory. 58, 2166–2174.

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.

Leifsson, L., Koziel, S., 2010. Multi-fidelity design optimization of transonic airfoils using physics-based surrogate modeling and shape-preserving response prediction. Journal of Computational Science 1, 98–106.

Majda, A., 2003. Introduction to PDE's and Waves for the Atmosphere and Ocean. AMS.

Martin, J.H., Knauer, G.A., Karl, D.M., Broenkow, W.W., 1987. Vertex: carbon cycling in the northeast pacific. Deep Sea Research Part A. Oceanographic Research Papers 34, 267–285.

McGuffie, K., Henderson-Sellers, A., 2005. A Climate Modelling Primer. Wiley. 3nd edition.

Paltridge, G.W., Platt, C.M.R., 1976. Radiative processes in meteorology and climatology, 318 pp.

Parekh, P., Follows, M.J., Boyle, E.A., 2005. Decoupling iron and phosphate in the global ocean. Global Biogeochemical Cycles 19.

Piwonski, J., Slawig, T., 2010. The Idea and Concept of Metos3D - A Marine Ecosystem Toolkit for Optimization and Simulation in 3-D. Technical Report 1060. Christian-Albrechts-Universität zu Kiel, Institute for Computer Science.

Piwonski, J., Slawig, T., 2011. Metos3d: A Marine Ecosystem Toolkit for Optimization and Simulation. CAU Kiel, Institut für Informatik. `http://www.informatik.uni-kiel.de/co2/software/meteos3d`.

Prieß, M., Koziel, S., Slawig, T., 2011a. A Fast and Robust Optimization Methodology for a Marine Ecosystem Model Using Surrogates. Technical Report 1110. Christian-Albrechts-Universität zu Kiel, Institute for Computer Science.

Prieß, M., Koziel, S., Slawig, T., 2011b. 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. pp. 449–457.

Prieß, M., Koziel, S., Slawig, T., 2011c. Surrogate-based optimization of climate model parameters using response correction. Journal of Computational Science , 1877–7503.

Queipo, N.V., Haftka, R.T., Shyy, W., Goel, T., Vaidyanathan, R., Tucker, P.K., 2005. Surrogate-based analysis and optimization. Prog. Aerosp. Sci. 41, 1–28.

Rückelt, J., Sauerland, V., Slawig, T., Srivastav, A., Ward, B., Patvardhan, C., 2010. Parameter optimization and uncertainty analysis in a model of oceanic CO_2-uptake using a hybrid algorithm and algorithmic differentiation. Nonlinear Analysis B Real World Applications 10, 3993–4009.

Simpson, T.W., Poplinski, J.D., Koch, P.N., Allen, J.K., 2001. Metamodels for computer-based engineering design: Survey and recommendations. Eng. Comput. 17, 129–150. 10.1007/PL00007198.

Smola, A.J., Schölkopf, B., 2004. A tutorial on support vector regression. Stat. Comput. 14, 199–222. 10.1023/B:STCO.0000035301.49549.88.

Søndergaard, J., 2003. Optimization using surrogate models - by the space mapping technique. Ph.D. thesis. Informatics and Mathematical Modelling, Technical University of Denmark, DTU. Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby. Supervisor: Kaj Madsen.

Yamanaka, Y., Tajika, E., 1997. Role of dissolved organic matter in the marine biogeochemical cycle: Studies using an ocean biogeochemical general circulation model. Global Biogeochemical Cycles 11, 599–612.

BIBLIOGRAPHY

Armstrong, R.A., Sarmiento, J.L., Slater, R.D., 1995. Monitoring ocean productivity by assimilating satellite chlorophyll into ecosystem models, in: Powell, T.M., Steele, J.H. (Eds.), Ecological Time Series. Chapmann & Hall, New York, pp. 371–390.

Bandler, J.W., Biernacki, R.M., Chen, S.H., Grobelny, P.A., Hemmers, R.H., 1994. Space mapping technique for electromagnetic optimization. IEEE Transactions on Microwave Theory and Techniques 42, 2536–2544.

Bandler, J.W., Cheng, Q.S., Dakroury, S.A., Mohamed, A.S., Bakr, M.H., Madsen, K., Søndergaard, J., 2004. Space mapping: The state of the art. IEEE Transactions on Microwave Theory and Techniques 52, 337–361.

Bärwolff, G., 2007. Numerik für Ingenieure, Physiker und Informatiker. Elsevier, München.

Boden, T., Blasing, T., 2011. Record high 2010 global carbon dioxide emissions from fossil-fuel combustion and cement manufacture. US DOE (ORNL), Carbon Dioxide Information Analysis Center (CDIAC).

Boyd, P.W., Jickells, T., Law, C.S., Blain, S., Boyle, E.A., Buesseler, K.O., Coale, K.H., Cullen, J.J., de Baar, H.J.W., Follows, M., Harvey, M., Lancelot, C., Levasseur, M., Owens, N.P.J., Pollard, R., Rivkin, R.B., Sarmiento, J., Schoemann, V., Smetacek, V., Takeda, S., Tsuda, A., Turner, S., Watson, A.J., 2007. Mesoscale iron enrichment experiments 1993-2005: Synthesis and future directions. Science 315, 612–617.

Braembussche, R.A., 2008. Numerical optimization for advanced turbomachinery design, in: Thevenin, D., Janiga, G. (Eds.), Optimization and Computational Fluid Dynamics. Springer, Berlin et al., pp. 147–189.

Braman, L.M., Suarez, P., van Aalst, M.K., 2010. Climate change adaptation: Integrating climate science into humanitarian work. International Review of the Red Cross 92, 693–712.

Broyden, C.G., 1965. A class of methods for solving nonlinear simultaneous equations. Mathematics of Computation 19, 577–593.

Conn, A.R., Gould, N.I.M., Toint, P.L., 2000. Trust-region methods. MPS-SIAM series on optimization, Society for Industrial and Applied Mathematics, Philadelphia.

Denman, K.L., 2008. Climate change, ocean processes and ocean iron fertilization. Marine Ecology Progress Series 364, 219–225.

Denman, K.L., Hofmann, E.E., Marchant, H., 1996. Marine biotic responses to environmental change and feedbacks to climate, in: Houghton, J.T., Meira Filho, L.G., Callander, B.A., Harris, N., Kattenberg, A., Maskell, K. (Eds.), Climate Change 1995. Cambridge University Press, Cambridge et al., pp. 483–516.

Drieschner, F., 2011. Auschub mit Zeitansage – Zum Stand der Klimadiplomatie vor der Konferenz in Durban: Die Sünder mauern. DIE ZEIT 48, 46.

Dumas, L., 2008. CFD-based optimization for automotive aerodynamics, in: Thevenin, D., Janiga, G. (Eds.), Optimization and Computational Fluid Dynamics. Springer, Berlin et al., pp. 191–214.

Echeverría, D., Hemker, P., 2005. Space mapping and defect correction. Computational Methods in Applied Mathematics 5, 107–136.

Echeverria, D., Hemker, P., 2008. Manifold mapping: A two-level optimization technique. Computing and Visualization in Science 11, 193–206.

European Environment Agency, 2005. Vulnerability and Adaptation to Climate Change in Europe. Technical Report.

Evans, G.T., Garçon, V.C., 1997. One-dimensional models of water column biogeochemistry. Report of a workshop held in Toulouse, France, November-December 1995. GOFS Report N °23/97, JGOFS Bergen, Norway.

Fasham, M.J.R. (Ed.), 2003. Ocean Biogeochemistry. The Role of the Ocean Carbon Cycle in Global Change. Global Change – The IGBP Series, Springer, Berlin et al.

Fennel, K., Losch, M., Schröter, J., Wenzel, M., 2001. Testing a marine ecosystem model: Sensitivity analysis and parameter optimization. Journal of Marine Systems 28, 45–63.

Fennel, W., Neumann, T., 2004. Introduction to the Modelling of Marine Ecosystems. Elsevier, Amsterdam et al.

Forrester, A.I.J., Keane, A.J., 2009. Recent advances in surrogate-based optimization. Progress in Aerospace Sciences 45, 50–79.

Ghil, M., Malanotte-Rizzoli, P., 1991. Data assimilation in meteorology and oceanography, in: Dmowska, R., Saltzman, B. (Eds.), Advances in Geophysics. Academic Press, London. Volume 33 of *Advances in Geophysics*, pp. 141–266.

Gill, A.E., 1982. Atmosphere - Ocean Dynamics. Volume 30 of *International Geophysics Series*. Academic Press, San Diego et al.

Glover, F., Kochenberger, G.A., 2003. Handbook of metaheuristics. International series in operations research & management science, Kluwer Academic Publishers.

Golub, G.H., Ortega, J.M., 1995. Wissenschaftliches Rechnen und Differentialgleichungen. Eine Einführung in die Numerische Mathematik. Volume 6 of *Berliner Studienreihe zur Mathematik*. Heldermann, Berlin.

Griffies, S.M., 2004. Fundamentals of Ocean Climate Models. Princeton University Press, Princeton, New Jersey.

Großmann, C., Roos, H.G., 2005. Numerischen Behandlung partieller Differentialgleichungen. Teubner Studienbücher Mathematik, Wiesbaden. 3rd edition.

Güssow, K., Proelss, A., Oschlies, A., Rehdanz, K., Rickels, W., 2010. Ocean iron fertilization: Why further research is needed. Marine Policy 34, 911–918.

Hackbusch, W., 2010. Elliptic Differential Equations: Theory and Numerical Treatment. Springer Series in Computational Mathematics, Springer, Berlin.

Hansjürgens, B., 2009. Internationale Klimapolitik nach Kyoto: Bausteine und Architekturen. Zeitschrift für Umweltpolitik und Umweltrecht 32, 123–152.

Hicks, R.M., Henne, P.A., 1978. Wing design by numerical optimization. Journal of Aircraft 15, 407–412.

Hurtt, G.C., Armstrong, R.A., 1999. A pelagic ecosystem model calibrated with bats and owsi data. Deep Sea Research Part I: Oceanographic Research Papers 46, 27–61.

IPCC, 2007a. in: Core Writing Team, Pachauri, R.K., Reisinger, A. (Eds.), Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the

Fourth Assessment Report of the Intergovernmental Panel on Climate Change. IPCC, Geneva, Switzerland, p. 104 et seq.

IPCC, 2007b. Summary for Policymakers, in: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M., Miller, H.L. (Eds.), Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge and New York, pp. 1–18.

Jameson, A., 1988. Aerodynamic design via control theory. Journal of Scientific Computing 3, 233–260.

Kelley, C.T., 1999. Iterative methods for optimization. Society for Industrial and Applied Mathematics, Philadelphia.

Kelley, C.T., 2003. Solving nonlinear equations with Newton's method. Society for Industrial and Applied Mathematics, Philadelphia.

Khatiwala, S., 2008. Fast spin up of ocean biogeochemical models using matrix-free newton-krylov. Ocean Modelling 23, 121–129.

Khatiwala, S., Visbeck, M., Cane, M., 2005. Accelerated simulation of passive tracers in ocean circulation models. Ocean Modelling 9, 51–69.

Körtzinger, A., Wallace, D.W.R., 2002. Der globale Kohlenstoffhaushalt und seine anthropogene Störung – Eine Betrachtung aus mariner Perspektive. promet 28, 64–70.

Kosmol, P., 1993. Methoden zur numerischen Behandlung nichtlinearer Gleichungen und Optimierungsaufgaben. Teubner, Stuttgart. 2nd edition.

Koziel, S., 2010a. Multi-fidelity multi-grid design optimization of planar microwave structures with sonnet. International Review of Progress in Applied Computational Electromagnetics , 719–724.

Koziel, S., 2010b. Shape-preserving response prediction for microwave design optimization. IEEE Transactions on Microwave Theory and Techniques 58, 2829–2837.

Koziel, S., Bandler, J.W., Cheng, Q.S., 2010. Robust trust-region space-mapping algorithms for microwave design optimization. IEEE Transactions on Microwave Theory and Techniques 58, 2166–2174.

Koziel, S., Yang, X. (Eds.), 2011. Computational Optimization, Methods and Algorithms. Volume 356 of *Studies in Computational Intelligence*. Springer, Berlin et al.

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.

Lampitt, R.S., Achterberg, E.P., Anderson, T.R., Hughes, J.A., Iglesias-Rodriguez, M.D., Kelly-Gerreyn, B.A., Lucas, M., Popova, E.E., Sanders, R., Shepherd, J.G., Smythe-Wright, D., Yool, A., 2008. Ocean fertilization: A potential means of geoengineering? Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 366, 3919–3945.

Lawson, L.M., Hofmann, E., Spitz, Y., 1996. Time series sampling and data assimilation in a simple marine ecosystem model. Deep Sea Research Part II: Topical Studies in Oceanography 43, 625–651.

Leifsson, L., Koziel, S., 2010. Multi-fidelity design optimization of transonic airfoils using physics-based surrogate modeling and shape-preserving response prediction. Journal of Computational Science 1, 98–106.

Luenberger, D.G., 2008. Linear and Nonlinear Programming. Springer, New York. 3rd edition.

Mahlman, J.D., Moxim, W.J., 1978. Tracer simulation using a global general circulation model – results from a midlatitude instantaneous source experiment. Journal of the Atmospheric Sciences 35, 1340–1374.

Malanotte-Rizzoli, P. (Ed.), 1996. Modern Approaches to Data Assimilation in Ocean Modeling. Volume 61 of *Elsevier Oceanography Series*. Elsevier, New York.

Marchuk, G.I., 1982. Methods of Numerical Mathematics. Springer, New York et al.. 2nd edition.

Matear, R.J., 1995. Parameter optimization and analysis of ecosystem models using simulated annealing: A case study at Station P. Journal of Marine Research 53, 571–607.

McGuffie, K., Henderson-Sellers, A., 2005. A Climate Modelling Primer. Wiley, Chichester. 3rd edition.

Meehl, G.A., Stocker, T.F., Collins, W.D., Friedlingstein, P., Gaye, A.T., Gregory, J.M., Kitoh, A., Knutti, R., Murphy, J.M., Noda, A., Raper, S.C.B., Watterson, I.G., Weaver, A.J., Zhao, Z.C., 2007. Global climate projections, in: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M., Miller, H.L. (Eds.), Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the

Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge and New York.

Nocedal, J., Wright, S.J., 2000. Numerical Optimization. Springer, New York.

O'Neill, B.C., Oppenheimer, M., 2002. Dangerous climate impacts and the kyoto protocol. Science 296, 1971–1972.

Oschlies, A., 2004. Feedbacks of biotically induced radiative heating on upper-ocean heat budget, circulation, and biological production in a coupled ecosystem-circulation model. Journal of Geophysical Research 110.

Oschlies, A., Garçon, V., 1999. An eddy-permitting coupled physical-biological model of the North Atlantic 1. Sensitivity to advection numerics and mixed layer physics. Global Biogeochemical Cycles 13, 135–160.

Oschlies, A., Koeve, W., Rickels, W., Rehdanz, K., 2010. Side effects and accounting aspects of hypothetical large-scale southern ocean iron fertilization. Biogeosciences 7, 4017–4035.

Pacanowski, R., Dixon, K., Rosati, A., 1991. The G.F.D.L Modular Ocean Model Users Guide Version 1. Technical Report 2. Geophysical Fluid Dynamics Laboratory, Ocean Group. Princeton.

Percival, S., Hendrix, D., Noblesse, F., 2001. Hydrodynamic optimization of ship hull forms. Applied Ocean Research 23, 337–355.

Piwonski, J., Slawig, T., 2011. METOS3D: A Marine Ecosystem Toolkit for Optimization and Simulation. CAU Kiel, Institut für Informatik. http://www.informatik.uni-kiel.de/co2/software/meteos3d.

Queipo, N.V., Haftka, R.T., Shyy, W., Goel, T., Vaidyanathan, R., Tucker, P.K., 2005. Surrogate-based analysis and optimization. Progress in Aerospace Sciences 41, 1–28.

Raven, J.A., Falkowski, P.G., 1999. Oceanic sinks for atmospheric CO_2. Plant, Cell & Environment 22, 741–755.

Rickels, W., Klepper, G., Dovern, J., Betz, G., Brachatzek, N., Cacean, S., Güssow, K., Heintzenberg, J., Hiller, S., Hoose, C., Leisner, T., Oschlies, A., Platt, U., Proelß, A., Renn, O., Schäfer, S., M., Z., 2011. Gezielte Eingriffe in das Klima? Eine Bestandsaufnahme der Debatte zu Climate Engineering. Sondierungsstudie für das Bundesministerium für Bildung und Forschung.

Rückelt, J., Sauerland, V., Slawig, T., Srivastav, A., Ward, B., Patvardhan, C., 2010. Parameter optimization and uncertainty analysis in a model of oceanic CO_2-uptake using a hybrid algorithm and algorithmic differentiation. Nonlinear Analysis B Real World Applications 10, 3993–4009.

Sarmiento, J.L., Gruber, N., 2002. Sinks for anthropogenic carbon. Physics Today 55, 30–36.

Sarmiento, J.L., Gruber, N., 2006. Ocean Biogeochemical Dynamics. Princeton University Press, Princeton et al.

Sarmiento, J.L., Slater, R., Barber, R., Bopp, L., Doney, S.C., Hirst, A.C., Kleypas, J., Matear, R., Mikolajewicz, U., Monfray, P., Soldatov, V., Spall, S.A., Stouffer, R., 2004. Response of ocean ecosystems to climate warming. Global Biogeochemical Cycles 18.

Schartau, M., 2001. Data-assimilation studies of marine, nitrogen based, ecosystem models in the North Atlantic Ocean. Ph.D. thesis. Christian-Albrechts-Universität Kiel.

Schartau, M., Oschlies, A., 2003. Simultaneous data-based optimization of a 1d-ecosystem model at three locations in the north atlantic: Part I - method and parameter estimates. Journal of Marine Research 61, 765–793.

Schartau, M., Oschlies, A., Willebrand, J., 2001. Parameter estimates of a zero-dimensional ecosystem model applying the adjoint method. Deep Sea Research II, 1769–1800.

Schmedders, K., 2008. Numerical optimization methods in economics, in: Durlauf, S.N., Blume, L.E. (Eds.), The New Palgrave Dictionary of Economics. Palgrave Macmillan, Basingstoke, 2nd edition.

Simpson, T.W., Poplinski, J.D., Koch, P.N., Allen, J.K., 2001. Metamodels for computer-based engineering design: Survey and recommendations. Engineering with Computers 17, 129–150.

Smola, A.J., Schölkopf, B., 2004. A tutorial on support vector regression. Statistics and Computing 14, 199–222.

Søndergaard, J., 2003. Optimization using surrogate models – by the space mapping technique. Ph.D. thesis. Informatics and Mathematical Modelling, Technical University of Denmark, DTU.

Spitz, Y.H., Moisan, J.R., Abbott, M.R., Richman, J.G., 1998. Data assimilation and a pelagic ecosystem model: Parameterization using time series observations. Journal of Marine Systems 16, 51–68.

Stoer, J., Bulirsch, R., 2002. Introduction to Numerical Analysis. Springer, New York. 3rd edition.

Storch, H., Güss, S., Heimann, M., 1999. Das Klimasystem und seine Modellierung. Eine Einführung. Springer, Berlin et al.

Talbi, E.G., 2009. Metaheuristics: From Design to Implementation. John Wiley & Sons, Inc., Hoboken.

Tarantola, A., 2005. Inverse Problem Theory and Methods for Model Parameter Estimation. Society for Industrial and Applied Mathematics, Philadelphia.

UN, 1992. United Nations Framework Convention on Climate Change. International Legal Materials 31, 849.

UN, 1998. Kyoto Protocol to the United Nations Framework Convention on Climate Change. International Legal Materials 37, 22.

UN, 2010. Report of the conference of the parties on its sixteenth session, held in Cancún from November 29 to December 10, 2010 – addendum, March 15, 2011, decision 1/CP.16, outcome of the work of the ad hoc working group on long-term cooperative action under the convention. FCCC/CP/2010/7/Add.1 , para. I.4.

UN, 2011. Status of Ratification of the Kyoto Protocol to the United Nations Convention on Climate Change. http://unfccc.int/kyoto_protocol/status_of_ratification/items/2613.php.

Ward, B., 2009. Marine Ecosystem Model Analysis Using Data Assimilation. Ph.D. thesis. http://web.mit.edu/benw/www/Thesis.pdf.

Ward, B., Anderson, M., Friedrichs, T., Oschlies, A., 2010. Parameter optimisation techniques and the problem of underdetermination in marine biogeochemical models. Journal of Marine Systems and Control Letters 81, 34–43.

Warren, R., 2011. The role of interactions in a world implementing adaptation and mitigation solutions to climate change. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 369, 217–241.

Weicker, K., 2007. Evolutionäre Algorithmen. Teubner, Stuttgart et al.. 2nd edition.

Acknowledgements

First of all, I want to say thank you to my supervisor Prof. Dr. Thomas Slawig for his encouragement and enduring support during the last three years. I am indebted to Thomas for offering me a great position in his working group. I want to thank Thomas for all the inspiring discussions and for sharing his profound knowledge with me during the last years. He also supported me with various decisions throughout the whole process. There have been a few times where I felt quite frustrated about the work. Thomas always took the time for listening and encouraging me. Thomas, thank you for all your support.

The position in Thomas' working group, I partially owe to Prof. Dr. Andreas Oschlies from the GEOMAR in Kiel. In 2008, I was looking for a position within the area of marine sciences. I want to thank Andreas for all his crucial support and advices at that time. I would like to thank him that he decided to be the co-supervisor and suggested to be the second reviewer of my thesis. Thank you, Andreas, for the great collaboration and your support during the last years.

I am also indebted to Prof. Slawomir Koziel, Ph.D. from Reykjavik University. It has been great fortune that I could start a collaboration with Slawomir in the year 2010. This cooperation rapidly grew from that time and has developed to a highly valuable teamwork. I want to thank you, Slawomir, for inspiring me with your profound experience and for the various intensive discussions via Email and at diverse conferences and research stays. They have been most valuable and contributed greatly. Finally, I want to thank Slawomir that he suggested to be a further reviewer of my work.

This work was supported as a part of the Cluster of Excellence *The Future Ocean* and has been furthermore assisted by the graduate school therein, the *Integrated School of Ocean Sciences*. I am particularly thankful for the financial support and the distinct highly valuable courses of studies and workshops. I also want to thank Dr. Avan Antia for the diverse inspiring discussions and her support.

Moreover, I want to thank all my colleagues in the working group for all the valuable discussions and thoughts and, last but not least, for a great time. Particularly, I have to thank my officemate Jaroslaw Piwonski who shared countless fruitful discussions with me. Thank you that I could bombard you with all my questions during the last years, particularly within the last months. Clearly, I also would like to thank Dr. Henrike Mütze, the project coordinator in my working group. Thank you for all the organization and, furthermore, all the support you provided during the last years. I would like to express my thanks to Willi Burmeister and Peter Pichol who commendably kept the technical side in our department running. Thanks both of you for all your great support.

Furthermore, I would like to thank Leifur Leifsson, Ph.D., from Reykjavik University for the great organization of highly valuable workshops and mini symposia and for the

invitations. Last but not least I want to thank you, Leifur, for your support and for spending a great time at diverse conferences and during my research stays in Iceland.

I am particularly indebted to Kerstin Güssow, Jaroslaw Piwonski, Joscha Reimer and Rasmus Prieß who spent various hours in proof-reading my thesis. Without their precious comments, many parts of this work would probably still be up to the confusions of my head. Thank you for your support.

I also would like to thank all my friends for encouraging me during the last years and for patiently bearing all the frustrations about my work in the past.

Finally, I want to say thank you to Kerstin, my parents and my brother for all their support during the last years. Without all their help, writing this thesis would clearly not have been possible.

Erklärung

Hiermit versichere ich,

I. dass diese Abhandlung – abgesehen von der Beratung durch die Betreuer Thomas Slawig, Andreas Oschlies und Slawomir Koziel – nach Inhalt und Form meine eigene Arbeit ist,

II. dass diese Arbeit zum Teil an einer anderen Stelle bereits veröffentlicht, bzw. zur Veröffentlichung eingereicht wurde, im Detail:

II.a. Aggressive space mapping for optimisation of a marine ecosystem model:
International Journal of Mathematical Modelling and Numerical Optimisation 3, 98–116 (2012), Inderscience Publishers, akzeptiert (veröffentlicht)

II.b. Surrogate-based optimization of climate model parameters using response correction:
Journal of Computational Science 2, 335–344 (2011), Elsevier, akzeptiert (veröffentlicht)

II.c. Marine Ecosystem Model Calibration through Enhanced Surrogate-Based Optimization:
Advances in Intelligent and Soft Computing Series, Springer, akzeptiert (im Druck)

II.d. Low-Cost Marine Ecosystem Model Calibration Using Surrogate-Based Optimization and Trust Regions:
Inverse Problems, IOP Publishing, eingereicht

II.e. Accelerated Parameter Identification in 3D Marine Ecosystem Models Using Surrogate-Based Optimization:
Ocean Modelling, Elsevier, eingereicht

III. dass diese Arbeit unter Einhaltung der Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft entstanden ist

IV. und dass die eingereichten Veröffentlichungen einen Eigenanteil gemäß des Schreibens meines Betreuers vorweisen.

Malte Prieß

12. Dezember 2011