Procedia Computer Science

Volume 29, 2014, Pages 2121–2126

ICCS 2014. 14th International Conference on Computational Science

Hierarchical emulation & data assimilation into the sediment transport model

Nugzar Margvelashvili[1], Eddy Campbell[2], Laurence Murray[2] and Emlyn[1] Jones

[1] CSIRO Marine and Atmospheric Research, Australia

[2]CSIRO Mathematical & Information Science, Australia

Nugzar.margvelashvili@csiro.au

Abstract

Synthetic observations of the suspended sediment concentration in an idealised macro-tidal estuary are assimilated into the 3d sediment transport model. The assimilation scheme relies on fast and cheap surrogates of the complex model (called emulators) to update the model's state variables and its 2 parameters. A scenario with a hierarchically structured emulator is contrasted to the scenario with a more conventional non-hierarchical emulator. Numerical experiments indicate that for a given size of the ensemble an emulator which replicates a hierarchical structure of the model tends to provide a better approximation of that model. Improving the quality of the emulator translates into the improved quality of the assimilation products.

1 Introduction

Fine-resolution nonlinear multi-disciplinary ocean models are computationally expensive and take days, weeks and longer to simulate ocean processes over the time-scales relevant to managemenet and research practices. These models are complex and difficult to calibrate. Often only a handful of the model trajectories is available and large number of scenarios relevant to the management or research objectives remain untested. Within the eReefs project, for example (the five-year collaborative study aiming at a comprehensive coastal information system for the iconic Great Barrier Reef, http://www.bom.gov.au/environment/activities/coastal-info.shtml), integrated hydrodynamic, sediment transport and biogeochemical models can take weeks to simulate one

Selection and peer-review under responsibility of the Scientific Programme Committee of ICCS 2014

doi:10.1016/j.procs.2014.05.196

year. These computational expenses translate into the limited capacity of such models to underpin decision-support for a wide range of management scenarios.

Emulators are fast and cheap approximations of complex models that run orders of magnitude faster than the original model (simulator). Emulators have been successfully used as a substitute for complex models in many industrial optimisation problems (Sacks et al., 1989). In the field of the ocean modeling the progress in this direction is much less advanced largely because of the high complexity of the ocean models. Despite this impediment, in recent years a number of promising approaches to emulation of ocean models have been published in scence literature (van der Merwe et al. 2007; Mattern et al., 2012; Margvelashvili and Campbell , 2012).

In this paper we employ emulators to assimilate synthetic observations of the suspended sediment concentration in an idealised macro-tidal estuary into the 3d sediment transport model. A scenario with a hierarchically structured emulator is contrasted to the scenario with a more conventional non-hierarchical emulator. Numerical experiments indicate that replicating a hierarchical structure of the model into its emulator reduces uncertainty of the emulator and improves the quality of the data-assimilation products.

2 Method and scenarios

2.1 Sediment transport model

The sediment transport model simulates sinking, deposition and resuspension of multiply size-classes of suspended sediment (Margvelashvili et al., 2009). The model solves advection-diffusion equations of the mass conservation of suspended and bottom sediments, and is driven by a fine-resolution, non-linear, non-stationary, free-surface, hydrostatic approximation hydrodynamic model (Herzfeld., 2006). Sediment particles settle on the seabed due to the gravity force and resuspend into the water column whenever the bottom shear stresses, exerted by waves and currents, exceed the critical shear stress of erosion. The resuspension and deposition fluxes are parameterised with the Ariathurai and Krone (1976) formula. Estimates of the bottom shear stress, required by this formula, are derived through the Grant and Madsen boundary layer model (Madsen, 1994). Bottom roughness is scaled by ripple dimensions (Grant and Madsen, 1982) which are considered the model input parameters. Sediments in benthic layers undergo vertical mixing due to bioturbation, represented by local diffusion.

The numerical grid for sediment variables in the water column coincides with the numerical grid for the hydrodynamic model. Within the bottom sediments, the model utilises a time-varying sediment-thickness-adapted grid, with the top active sediment layer having constant thickness, and the thickness of deeper layers varying with time to accommodate the deposited sediment. Horizontal resolution within sediments follows the resolution of the water column grid.

The sediment transport model is initialised with a prescribed uniform distribution of sediments in benthic layer, and zero initial concentration of the suspended sediment. The concentration of the suspended sediment is set to a constant value at the upstream river boundary. At the marine boundary the model utilises a free-flow boundary condition when the direction of currents is downstream and no inflow of sediments when the direction of currents reverses.

2.2 Emulation and assimilation scheme

The emulation and assimilation scheme, in general, follows the approach of Margvelashvili & Campbell (2012). The method involves ensemble forecasting step followed by the dimension reduction, error-subspace propagation and Monte Carlo Markov Chain (MCMC) sampling steps delivering updated parameters of the model into the final, analysis step. Observations are assimilated sequentially in time over the predefined time-window (t_i,t_{i+1}). The model is initialised with a prior distribution of parameters. After the forecast step the model dimension is reduced through the Singular Value Decomposition (SVD) and then the Gaussian Process Modelling (GPM) is employed to map decomposition coefficients from time t_i to time t_{i+1} (i.e. GPM implements an error-subspace propagation step). GPM and SVD give a fast and computationally inexpensive approximation of the model called an emulator. Observations are assimilated into this emulator through the combined Residual Sampling (RS, see Lui and Chen, 1998) and Differential Evolution (DE, see Ter Braak, 2006) MCMC steps. Updated parameters initialize the model analysis.

2.3 Numerical experiments

The model simulates suspended sediment transport in an idealised macrotidal estuary comprised of a 2d vertical channel and 3d open coastal water (Figure 1). The water column is coupled to a 3d benthic layer of sediment deposits through the sediment resuspension and deposition fluxes. Two classes of sediments are considered (C1 and C2) each having specific initial distribution in the water column and in the seabed and specific boundary conditions.

Observations (obtained from the twin model run) are available only for the sum of C1 and C2 concentrations called Total Suspended Solids (TSS). These observations are given as 2 daily snapshots of surface TSS (truth) contaminated with a measurement error (stdev 50% of the value).

The model ensemble comprises 24 members. Spatial distributions are approximated through the decomposition of the model solution into 3 basis functions per a state variable and per a time-window. The assimilation scheme updates 4 state variables (2 classes of sediment in the water column and 2 classes of sediment in the seabed) and two model parameters (settling velocity and ripple height).

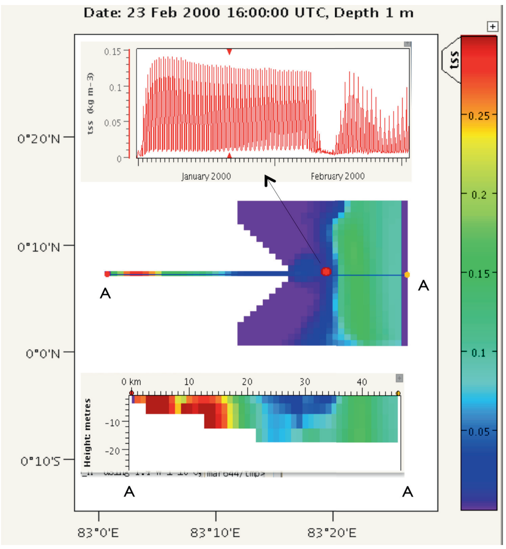

Figure 1: Time series of simulated TSS (top); a snapshot of simulated surface TSS (middle); and a snapshot of simulated vertical cross section of TSS (bottom).

2.4 Scenarios

Two assimilation scenarios are considered, one with a hierarchical emulator of TSS and another one with a non-hierarchical emulator (fig. 2). The nonhierarchical emulator expresses TSS at observation time (t_{i+1}) as a function of 4 state variables (suspended and benthic concentrations of C1 and C2 at time t_i) and 2 parameters (settling velocity and ripple height). Each state variable is approximated by 3 basis functions which results in a 14 dimensional design space (4x3+2).
A hierarchical emulator represents TSS at observation time (t_{i+1}) as a function of C1 and C2 concentrations both taken at the same time (t_{i+1}). C1 and C2 are in turn represented as functions of the corresponding initial states and parameter at t_i . C1, for example, is given as a function of its initial distribution in the water column and in the seabed and 2 parameters. The maximum dimension of the design space in this case is 8 (2x3+2).

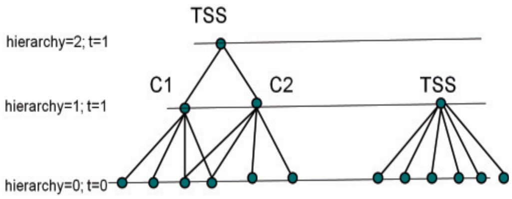

Figure 2: Hierarchical (left) and non-hierarchical structures of the emulator

3 Results and Conclusions

Figure 3 illustrates comparison between emulated and true values of the SVD coefficients for the first three modes of the decomposition. The quality of both hierarchical and non-hierarchical emulators tends to deteriorate towards the higher modes of the decomposition (from left to right). The hierarchical emulator provides persistently better fit to the model as compared to the non-hierarchical emulator.

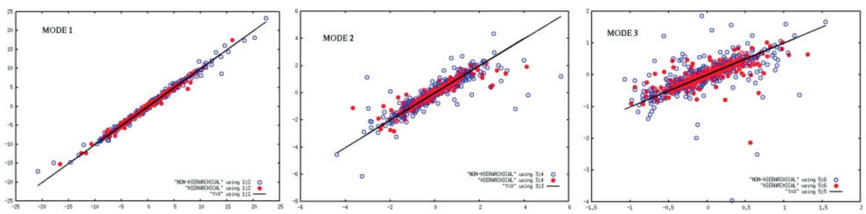

Figure 3: Emulated decomposition coefficients of TSS fields vs true values for the first three modes of SVD (red filled circles represent scenario with a hierarchical emulator).

The quality of the assimilation products have been assessed by estimating uncertainty of the model parameters (Fig. 4) and evaluating ensemble mean error of the inferred TSS (Fig. 5). A hierarchical emulator tends to produce lower spread of the estimated parameters and smaller bias of the mean values. Replicating a hierarchical structure of the model into its emulator also improves the quality of the inferred TSS fields (Fig. 5).

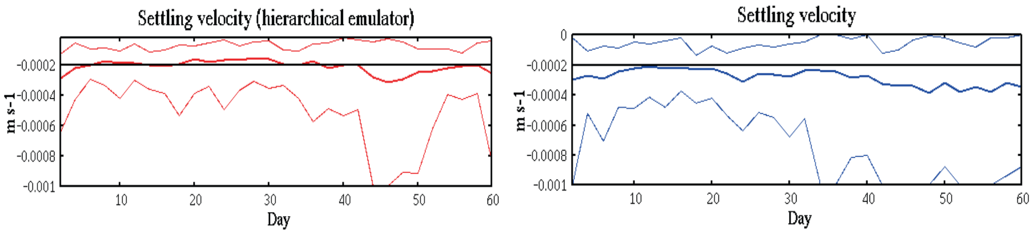

Figure 4: Estimated mean (thick colored lines) and the range of settling velocity vs truth (black lines) for scenarios with hierarchical (left) and non-hierarchical emulators.

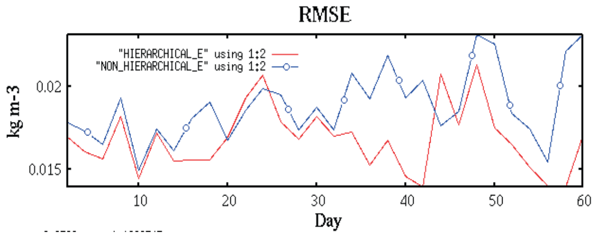

Figure 5: Ensemble mean error of the estimated TSS for the scenario with a hierarchical emulator (red) and scenario with a non-hierarchical emulator (blue line with circle markers)

To conclude, the results of our study suggest that having imposed additional constraints on the emulator through the hierarchical structuring of the dependencies between the state variables and the parameters, improves the quality of the emulator and reduces uncertainty of the assimilation products.

Acknowledgments

This work was funded through CSIRO CSS TCP "Biogeochemical Cycles" project and CSIRO WfO Flagship. We also thank CMAR Coastal Modelling Group for access to an idealized estuary model http://www.emg.cmar.csiro.au/www/en/emg.html

References

Ariathurai, R., Krone, R. B., (1976). Finite element model for cohesive sediment transport. J. Hydraul. Div. ASCE, 104, HY2, pp. 323-328.

Grant, D., Madsen, O., (1982). Movable bed roughness in unsteady oscillatory flow. J. of Geophysical Research, V. 87, N C1, pp. 469-481.

Herzfeld, M., (2006). An alternative coordinate system for solving finite difference ocean models. Ocean Modelling, 14, pp. 174-196.

Lui, J., Chen, R., (1998). Sequential MonteCarlo methopds for dynamical systems. J. Amer. Stat. Assoc., 90, 567-576.

Madsen, O. S., (1994). Spectral wave-current bottom boundary layer flows, in Coastal Engineering 1994 Proceedings, 24th International Conference Coastal Engineering Research Council/ASCE, pp. 384-398.

Margvelashvili, N., (2009). Stretched Eulerian coordinate model for coastal sediment transport. Computers and Geosciences, 35, 1167-1176.

Margvelashvili, N., Campbell, E., (2012). Sequential data assimilation in fine resolution models using error-subspace emulators: theory and preliminary evaluation. Journal of Marine Systems, 90, 13-22.

Margvelashvili N., Andrewartha J., Herzfeld M., Robson B., Brando V. (2013). Satellite data assimilation and estimation of a 3d coastal sediment transport model using error-subspace emulators. Journal Env. Modelling and Software, 40, 191-201.

Mattern, J.P., Fennel, K., Dowd, M., (2012). Estimating time-dependent parameters for a biological ocean model using an emulator approach. Journal of Marine Systems 96-97, 32-47.

Sacks, J., Welch, W., Mitchell, T., Wynn, H., (1989). Design and analysis of computer experiments. Statistical Science 4 (4), 409-423.

Ter Braak, C., (2006) A Markov Chain Monte Carlo version of the genetic algorithm Differential Evolution: easy Bayesian computing for real parameter spaces Stat Comput 16, 239–249.

Van der Merwe, R., Leen, T., Lu, Z., Frolov, S., Baptista A. (2007). Fast neural network surrogates for very high-dimensional physics-based models in computational oceanography. Neural Networks 20, 462-478.