

Multiscale Reduced Order Modeling and Parameter Estimation for Climate Sciences

Arvind Mohan, Zhendong Cao, Marian Anghel, Mathew Maltrud, Christine Sweeney, Michael McKerns, Xylar Davis (Los Alamos National Laboratory, New Mexico)

Science challenge: Several problems in earth system modeling are dependent on highly multiscale phenomena, such as turbulence, where computational modeling is challenging and expensive. This issue is exacerbated in atmospheric and oceanic domains, due to inherent high-dimensionality of the problem. One approach to this problem has been reduced order modeling (ROM); which aims to represent the key physics of the phenomena as a low-dimensional system. AI methods have huge potential in building accurate, stable ROMs and parameter estimation for these ROMs, as it requires extracting nonlinearities and patterns from simulation and/or observational data. Developing physics-based AI approaches specialized for the complexities of multiscale data, along with strategies to account for uncertainties, will revolutionize rapid modeling, analysis and decision making for earth system problems of practical interest.

Focal Areas: Predictive modeling of complex systems with AI with tightly integrated component level models.

Narrative:

Model Reduction for ROM: Reducing dimensionality is the first step in constructing ROMs from a multiscale dataset [1-3]. This is often crucial to embed small- and meso-scale processes, such as the Cloud Resolving Models (CRMs), in each column of a large-scale model (ESM). ROMs will be constructed for both the *forward* and *backward* (adjoint) models, which are necessary for efficient parameter and state estimation. For example, consider the evolution of a non-linear, two-scale model with state variables (X, x) . The *coarse* physical process X evolves as $\dot{X} = F(X; o(x), \Lambda)$ and defines boundary conditions and driving forces $i(X)$ for the *fine* scale process. The fine-scale process evolves as $\dot{x} = f(x; i(X), \lambda)$. It provides information (feedback) $o(x)$ for the coarse-scale behavior through the computation of the response function $o(x) = M[i(X)]$, that takes coarse-scale driving forces as arguments and returns suitable microscopic averages. Here Λ and λ are coarse- and fine-scale parameters that have to be determined from observations. Some of the most popular model reduction techniques are Proper Orthogonal Decomposition (POD), including Dynamic Mode Decomposition (DMD), kernel PCA and more recently, neural network

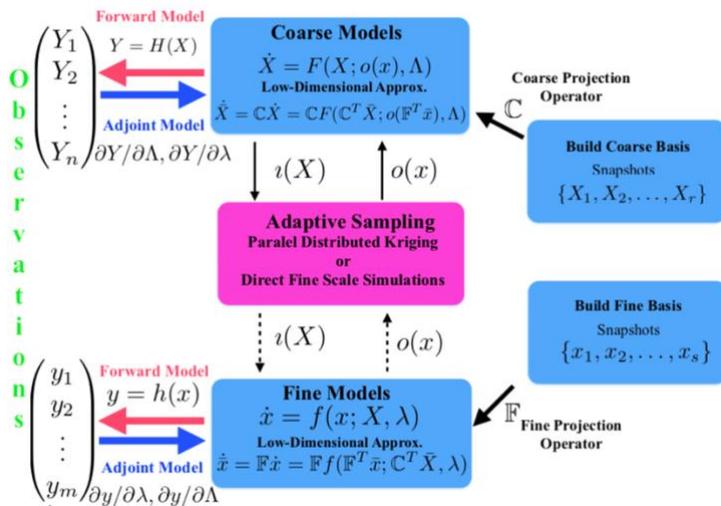


Figure 1 Two-scale hierarchical model with an adaptive sampling (AS) module inserted between the coarse-scale and the fine-scale. The AS module receives requests for fine-scale response information from the coarse-scale model.

autoencoders. These techniques identify the optimal basis vectors with a least-squares metric for a low dimensional representation of the system.

$$q(x, t) = \bar{q}(x) + \sum_{i=1}^{N_{\text{POD}}} a_i^P(t) \phi_i(x)$$

$$C^\theta = NN_\theta^C(f(q))$$

$$L^\theta = NN_\theta^L(f(q))$$

$$\dot{a}_i^R(t) = C_i^\theta + \sum_{j=1}^{N_{\text{gal}}} L_{ij}^\theta a_j^R(t) + \sum_{j,k=1}^{N_{\text{gal}}} Q_{ijk}$$

Figure 2 Differentiable Programming Based Galerkin-ROM

CRM and the reduction technique described above will be applied in parallel to each local (ESM, CRM) pair. The Galerkin Projection equation leads to the constant, linear and quadratic terms that are determined via an optimization and calibration process from the *data* (Fig.2). We remark that problems with this structure of unknown coefficients inside a known equation lend themselves well to be treated with *differentiable programming* techniques, among others. The unknown coefficients in the equations are represented by neural networks (NN) of appropriate dimensionality, and the adjoint-based backpropagation learns this NN accurately from the data [6,7]. This flexible abstraction lends itself to a variety of problems with arbitrary equation structures and smooth operators.

Distributed algorithms for coupled ROMs: In general, the ROMs for climate dynamics do not provide stable models that can predict its evolution over long time scales: the state evolution of these ROMs diverges to infinity. We expect to employ ML-based POD-Galerkin projection methods that generate stable local ROMs in polynomial form and distributed strategies to estimate the stability of coupled models of polynomial ODEs exist. These strategies use algebraic geometric techniques and can bound the dynamics of the coupled models without performing any dynamic simulations, using stability arguments from network control and safety verification methods. We can estimate bounds in the state space for the local dynamics given bounds on the states of the neighboring ROMs and we can learn structural corrections to the ROMs that provide dynamic stability, or detect when the ROMs become unstable, signaling a change in the climate state. Developing tools to discriminate the intrinsic instability of the climate from the instability due to the imperfection of ROMs will be necessary.

Adaptive Sampling (AS) strategies: Since the computational cost of directly embedding CRMs remains prohibitive, we can employ AS strategy (Fig.1) to learn the map $o(x) = M[i(X)]$ that provides (local) fine-scale feedback to the coarse ESM models. Since we expect the CRM to explore a large number of dynamic regimes, we expect to adaptively update their reduced order approximations. We can use AS [8] to substantially reduce the total number of expensive fine-scale updates and to produce a library of CRMs that is only updated when significant drift in the large-scale driving forces $i(X)$ is detected. We should implement an AS library to operate at the coarse scale, as dynamic drift is expected at this scale as well, and ROMs for the ESM have to be

Machine Learning based ROM: Reduced basis from model reduction is used to predict system evolution on a low-dimensional manifold. Popular methods involve Galerkin projection [4,5] of the reduced basis on the governing equations like the Navier-Stokes equations. A POD-Galerkin projection will produce ROMs for the coarse and fine scale evolution. Since an accurate, global, low-dimensional model does not exist, we propose to decompose the coarse-scale ESM model into a number of local interacting subsystems. Each such subsystem will be coupled locally to a

such subsystem will be coupled locally to a

updated as the system evolves. Therefore, distributed techniques are needed to detect where and when the ROM models become inadequate and should be corrected. This ties in with the distributed approach, where local ROMs can be tailored to retain/truncate basis contingent on local physics, while simultaneously adhering to global conservation laws. This can be dynamically learned from *differentiable programming*-based POD-Galerkin models described above.

Data assimilation for parameter estimation: For models at both global and local scales. We propose to pose the estimation of coarse, Λ , and fine-scale, λ , parameters as an optimization problem that seeks to minimize the difference (cost) between the predicted model behavior and observations. We embed the cost into a Bayesian framework [9], in order to get an estimate of parameters uncertainty. We will use a Hamiltonian Monte Carlo approach to sample efficiently over the posterior distribution in parameter space. This requires the computation of the derivatives of the cost function in order to predict the change in response due to change in parameters, and will be implemented efficiently using an *adjoint* formulation that simultaneously computes the sensitivities with respect to hundreds of parameters. Similar to the construction of two-scale, coupled ROMs for the forward evolution models, we will build POD bases and low-dimensional approximations in the *adjoint* space to speed up the adjoint computation. We envision *local* inference strategies to estimate these parameters in parallel given the local (ESM, CRM) pair models and data.

Scalable Machine Learning for ROMs: To identify optimal low dimensional representations of the forward and backward adjoint models from large datasets and optimally estimate parameters for two-scale coupled ROMs, sophisticated control strategies will be needed, such as adaptive control/learning and/or reinforcement learning. These will need to be scalable in order to optimize over a large parameter space, and distributed in order to be tailored to local physics. Optimization and control toolkits have been used to enable hybrid AI/physics simulation workflows as part of large scientific applications. Scalable reinforcement learning frameworks are being developed to enable distributed control on high-performance computers, where heterogeneous learning and inference tasks co-exist with model execution on shared resources. This maximizes resource utilization and minimizes communication costs.

Candidate Application: Biogeochemical (BGC) models [10,11] involve various simplified parameterizations to mathematically describe the complicated biogeochemical cycling processes in concern. We introduce free parameters to regulate the functioning of these algebraic parameterizations, which need to be tuned to match model results with real data. Unfortunately, traditional parameter tuning/estimation during model calibration becomes challenging primarily due to the curse of dimensionality. This manifests in: 1) Numerous tunable parameters, exploring a large parameter space. 2) Tuning all parameters simultaneously is intractable and expensive. 3) Parameter tuning has spatial limitations: a well calibrated model in one region may not be suitable for another region. 4) Comprehensive uncertainty quantification is nearly intractable. Nearly all these issues can be targeted with distributed, coupled ML-ROMs described above, since the equations are low-dimensional and considerably fewer parameters. Finally, we emphasize that these strategies are broadly applicable to ESM problems, as it can better learn from large observational/simulation datasets, while incorporating physics constraints.

References:

- [1] Ito, Kazufumi, and Sivaguru S. Ravindran. "A reduced-order method for simulation and control of fluid flows." *Journal of computational physics* 143.2 (1998): 403-425
- [2] Burkardt, John, Max Gunzburger, and Hyung-Chun Lee. "POD and CVT-based reduced-order modeling of Navier–Stokes flows." *Computer methods in applied mechanics and engineering* 196.1-3 (2006): 337-355.
- [3] Noack, Bernd R., Marek Morzynski, and Gilead Tadmor, eds. *Reduced-order modelling for flow control*. Vol. 528. Springer Science & Business Media, 2011.
- [4] Rapún, María-Luisa, and José M. Vega. "Reduced order models based on local POD plus Galerkin projection." *Journal of Computational Physics* 229.8 (2010): 3046-3063.
- [5] Kalashnikova, I., and M. F. Barone. "On the stability and convergence of a Galerkin reduced order model (ROM) of compressible flow with solid wall and far-field boundary treatment." *International journal for numerical methods in engineering* 83.10 (2010): 1345-1375.
- [6] Mohan, Arvind, Kaushik Nagarajan, and Daniel Livescu. "Learning Physics-based Galerkin models of turbulence with Neural Differential Equations." *Bulletin of the American Physical Society* (2020).
- [7] Rackauckas, Christopher, et al. "Universal differential equations for scientific machine learning." *arXiv preprint arXiv:2001.04385* (2020).
- [8] Knap, J., et al. "Adaptive sampling in hierarchical simulation." *International journal for numerical methods in engineering* 76.4 (2008): 572-600.
- [9] Wikle, Christopher K., and L. Mark Berliner. "A Bayesian tutorial for data assimilation." *Physica D: Nonlinear Phenomena* 230.1-2 (2007): 1-16.
- [10] Moore, J.K., Doney, S.C. and Lindsay, K., 2004. Upper ocean ecosystem dynamics and iron cycling in a global three-dimensional model. *Global Biogeochemical Cycles*, 18(4).
- [11] Burrows, S.M., Maltrud, M., Yang, X., Zhu, Q., Jeffery, N., Shi, X., Ricciuto, D., Wang, S., Bisht, G., Tang, J. and Wolfe, J., 2020. The DOE E3SM v1. 1 biogeochemistry configuration: Description and simulated ecosystem-climate responses to historical changes in forcing. *Journal of Advances in Modeling Earth Systems*, 12(9), p.e2019MS001766.