Hybrid (PDE+ML) models in the context of land ice modeling

Mauro Perego (SNL), Luca Bertagna (SNL), Matthew Hoffman (LANL),
Stephen Price (LANL), Sivasankaran Rajamanickam (SNL), Jerry Watkins (SNL)

Focal Area(s)

Primary focal areas are: predictive modeling through the use of AI-derived model components; advanced methods including network design/optimization/deep learning.

Science Challenge

The atmospheric, ocean and ice dynamics components of the Energy Exascale Earth System Model (E3SM) are governed by Partial Differential Equations (PDEs) and significant efforts have been made during the last decades to develop such computational models. Here we propose to *fundamentally improve* these PDE-based codes by *enhancing them with Machine Learning (ML) sub-models* for complex, poorly understood physical processes in the context of ice sheet modeling. We propose to train these models with a *novel approach* that allows the assimilation of the different sources of data available (direct/indirect observations and possibly simulation data), improving on existing simplified models. We also highlight computational challenges originating from the coexistence of PDE-based and ML-based models.

Rationale

The idea that deep learning models could entirely replace current climate models can be appealing. A possible barrier to deep learning models for the entire climate is the fact that most of the observations cover a relatively short period of time compared to the characteristic temporal scales of some important climate processes (e.g. ice sheet thermo-mechanics and deep ocean processes) and therefore they do not contain, without enforcing physics constraints, enough information for emulating the climate. Reichstein et al., 2019, advocate for hybrid climate models where traditional PDE based models coexist with, or are enhanced by, deep learning models.

While parametrizations and surrogate models have been around for a long time, the transition to hybrid (PDE+ML) models is *a paradigm shift* and poses multiple challenges, ranging from the *innovative design and training* of ML architectures to the efficient use and the *shaping* of future heterogeneous High Performance Computing (HPC) machines. Here we address some of these challenges in the context of hybrid models for land ice modeling.

Ice Sheets impact the water cycle and water cycle extremes in several ways. Despite the fact that land ice timescales can be on the order of thousands of years, some of the *most impactful effects may be initiated and/or will play out in coming years* or decades: 1. An increase in global mean sea level of a foot - as a result of glacier and ice sheet melting - could significantly worsen the effects of hurricanes and consequent floods. 2. The initiation of the West Antarctic marine ice sheet instability (see Joughin et al., 2014; Edwards et al., 2021) - which if triggered, would significantly increase the rate of sea level rise - could happen in the near future (see the AI4ESP idea “Early detection and uncertainty quantification of rapid sea-level rise from Antarctica” by S. Price et al.). 3. Glacier floods (Jokulhlaups, see Cuffey and Paterson, 2010) and glacier collapses (see Gilbert et al., 2015) are extreme events with often catastrophic consequences when happening near inhabited places, and they strongly depend on glacier

DOI: https://doi.org/10.2172/1769717
DOE BER, AI4ESP 2021 #AI4ESP1103 (https://www.ai4esp.org)

hydrology processes that we propose to model with ML. Further, increased freshwater flux to the oceans from ice sheets melting can have crucial, and hard to quantify, effects on ocean circulation (e.g. slowing down the *Atlantic Meridional Overturning Circulation*).

This work *will leverage the unique resources* (codes, expertises, computational power) that DOE has been supporting for several years, and it will help ensure that these resources will be harnessed and shaped for *defining the next era of climate modeling*.

Narrative

Land ice dynamics involve temporal scales up to tens of thousands of years that can hardly be captured by ML models relying only on observations covering a relatively short time span. To overcome this issue, we choose to rely on traditional thermo-mechanical PDEs and resort to ML models for constitutive relations or for other parts of the model (e.g. boundary conditions). In fact, several physical processes are not fully understood, lack validated theoretical models or are too expensive (e.g. because they feature different spatial/temporal scales) to be modeled. Among these are glacial hydrology (see e.g. Hoffman and Price, 2014; Flowers, 2015; de Fleurian et al., 2018), and in particular channelization of subglacial water flow, calving physics (see e.g. Alley R., 2008; Amundson and Truffer, 2017). These sub-models have a strong impact on the ice velocity. Hydrology, and in particular water pressure under the glacier, affects basal sliding, which is the most important control on ice velocity.

In order to train these models we need to harness the available observations. We often have few *direct* observations. Direct observations of the water pressure, which is a natural output of a hydrology model, are obtained using borehole measurements. These measurements alone are hardly enough to train an ML model for hydrology. However, there may be more than adequate *indirect* observational data that we can harness. In particular, satellite measurements of ice velocity are increasingly accurate and have high spatial and temporal resolution. Therefore we can use the velocity data to constrain the ML hydrology model. This requires an additional step, because once the water pressure at the base of the ice sheet is known, we need to solve the ice dynamics equations that depend on the water pressure and compute the ice velocity.

Similarly, in the case of iceberg calving physics, we typically do not have direct measurements of the ice calving rate, but we have observations of the calving front of the ice and we can track its position in time.

We identify three main research directions associated with these problems:

1. **Design of machine learning submodels**

A crucial part for the success of a deep learning approach is to design an architecture that can adequately represent the process we want to model. In the context of scientific ML, several efforts are being made to incorporate physics constraints into the ML models. Notable examples are the Physics Informed Neural Networks (PINNs), see Mazier et al. 2019, and Universal Differential Equations (UDE), see Rackauckas, C. et al. 2020. These approaches mainly focus on emulating standalone differential equations. UDEs also allow discovering the governing physics, but when working with very complex physics models one can hardly expect that a simple differential equation can properly represent the full model complexity. In order to emulate sub-models, we consider different approaches. We first need to distinguish between local and nonlocal sub-models (e.g., constitutive laws are often local, such as Glen's law for ice, which is a function of the local ice temperature and velocity gradient). However, more complex physical models, like the hydrology model, that involve the solution of differential equations are nonlocal (the solution does not depend only on the local values of the inputs). Local models can be represented by functions, whereas nonlocal models can be represented by operators that map fields into fields. The latter case is more complex and operator regression is becoming a very active field in ML (see Lu Lu et al. 2019; Bhattacharya et al. 2020; Trask et al. 2019; Patel et al. 2019).

2. Training with different sources of data

The typical setting of a regression problem consists of minimizing a loss function involving a mismatch term between the model output and the data. In our case we want to incorporate observations of derived quantities of the ML output. Consider the glacial hydrology problem where we have observations of the surface ice velocity. A hydrology ML model will provide as output the pressure of the water within the hydrology system at the base of a glacier. Water pressure affects the sliding of the ice and, in turn, the ice velocity, through the solution of the momentum equations. This means that for training the submodel we need to solve PDE problems. When using gradient-based optimization methods (e.g. stochastic gradient descent), we need to propagate gradients through the solution, which, in the context of PDE constrained optimization is often achieved using adjoint methods. This demands that the PDE code can efficiently compute gradients, which is often not the case.

Ideally, we would use all the observations available, and in particular direct observations of the model output. In the case of hydrology, borehole measurements have the potential to be transformative due to the amount of variables simultaneously observed and their measurement high-frequency. One can think of adopting transfer learning techniques (here intended in a broader sense than the typical ML meaning) to pre-train the submodel to match borehole measurements. Finally, one would use this pre-trained ML sub-model and further train it to match surface velocity observations.

3. Implementation of hybrid models

We consider the scenario where a large, efficient (e.g., parallel and scalable), and trusted PDE-based code is available. This is the case for many DOE funded production codes, developed and matured over several years, and optimized to run on emerging HPC machines. Among these is the E3SM land ice code MALI (see Hoffman et al., 2018; Tezaur et al., 2015). Ideally we want to enhance this model with improved ML-based capabilities with minimal disruptions, and we want the resulting code to be efficient in training and evaluation. Different solutions are available. On one end, we could couple the code with ML libraries like PyTorch or TensorFlow, so that the ML model is implemented in third party ML libraries. On the other end we could enhance the PDE-based code by providing implementation of ML architectures (e.g. feed forward or recurrent neural networks) and with tools for optimization / training. MALI already has built in optimization and automatic differentiation capabilities through the Trilinos software suite, but it would be a huge endeavor to achieve the flexibility of ML dedicated libraries, especially when training. An intermediate approach could be to couple the PDE-based code and a dedicated ML library, while designing and training the ML submodel in an offline phase, and eventually port the specific ML architecture of choice into the PDE-based application. Independent from the strategy adopted, significant additional efforts will be needed to implement these hybrid models and have them efficiently run on future DOE heterogeneous HPC machines.

References

- Reichstein, M. et al. 2019. Deep learning and process understanding for data-driven Earth system science. Nature, 566.
- Price, S.F., Payne, A.J., Howat, I.M., Smith, B.E., 2011. *Committed sea-level rise for the next century from Greenland ice sheet dynamics during the past decade.* Proc. Natl. Acad. Sci. U. S. A. vol 108.
- Levermann, a., Clark, P.U., Marzeion, B., Milne, G. a., Pollard, D., Radic, V., Robinson, a., 2013. *The multimillennial sea-level commitment of global warming.* Proc. Natl. Acad. Sci. 1–6.
- Joughin, I., Smith, B. E., & Medley, B. 2014. Marine ice sheet collapse potentially underway for the Thwaites Glacier basin, West Antarctica. Science, vol. 344.
- Edwards, T. et al. 2021. Antarctica dominates uncertainty in projected land ice contributions to sea level rise this century. Nature (in press).
- Cuffey, K..M. and Paterson, W.S.B., 2010, The physics of glaciers, Elsevier
- Gilbert, A., Vincent, C., Gagliardini, O., Krug, J. &Berthier, E. (2015). Assessment of thermal change in cold avalanching glaciers in relation to climate warming. Geophys. Res. Lett.
- Hoffman, M. and Price, S., 2014, Feedbacks between coupled subglacial hydrology and glacier dynamics, JGR Earth Surface, vol. 119, issue 3.
- Flowers, G.E., 2015. Modelling water flow under glaciers and ice sheets. Proc. R. Soc. A Math. Phys. Eng. Sci. 471, 20140907.
- Alley R. et al. (2008), A Simple Law for Ice-Shelf Calving, Science, Vol. 322, Issue 5906.
- De Fleurian et al., 2018, SHMIP The subglacial hydrology model intercomparison Project, Journal of Glaciology, vol. 64, issue 248.
- Amundson, J.M., Truffer, M., 2017, A unifying framework for iceberg-calving models, Journal of Glaciology, vol. 56, issue 199.
- Raissi, M., Perdikaris, P., Karniadakis, G. E., 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, Journal of Computational Physics, Volume 378, p. 686-707.
- Rackauckas, C. et al. 2020. Universal Differential Equations for Scientific Machine Learning, arXiv.
- Lu Lu et al. 2019. DeepONet: Learning nonlinear operators for identifying differential equations based on the universal approximation theorem of operators, arXiv.
- Bhattacharya,K. et al. 2020. Model Reduction and Neural Networks for Parametric PDEs, arXiv.
- Trask N., Patel, R. Gross B., Azberger P., 2019, GMLS-Nets: Scientific Machine Learning Methodsfor Unstructured Data, NeurIPS

- Patel R., Trask N., Woods M., Cyr E., 2019, A physics-informed operator regression framework for extracting data-driven continuum models, CMAME
- Hoffman, M., Perego, M., Price, S., Lipscomb, W., Zhang, T., Jacobsen, D., Tezaur, I., Salinger, A., Tuminaro, R., and Bertagna, L. 2018. MPAS-Albany Land Ice (MALI): A variable-resolution ice sheet model for Earth system modeling using Voronoi grids. Geoscientific Model Development, 11(9), 3747–3780. doi:10.5194/gmd-11-3747-2018
- Tezaur, I.K., Perego, M., Salinger, A.G., Tuminaro, R.S., Price, S.F., 2015, Albany/FELIX: a parallel, scalable and robust, finite element, first-order Stokes approximation ice sheet solver built for advanced analysis, Geoscientific Model Development, vol. 8, issue 4.