Large-Scale AI Platform for Climate Modeling

By Kemal A. Delic - Martin Antony Walker

Climate modeling (Ref. 1,2) is a hypercomplex problem requiring huge computational resources and careful combination of several sub-models and disciplines into a coherent whole in order to produce climate forecasts. In order to establish policies for the development of adaptations to the changing climate, climate models have to produce trustworthy forecasts of sufficient accuracy. This is not the case today. The hope is that the application of AI-inspired methods will enable the development of trustworthy climate models.

Highly abstracted, climate models consist of physical models of the atmosphere and oceans based on hydrodynamics and thermodynamics, interplaying with chemical and biological models based on quantum mechanics of radiation and molecules. The many different orders of magnitude of length and time scales involved make direct computation from basic principles impracticable. Research groups around the world are attempting to help bridge the gaps between micro- and macro-scale science in climate models through application of methods involving inferences from artificial neural networks trained on observed data. Consequently, we postulate that the platform for global modeling should be decomposed into a hybrid architecture with sub-platforms differing in nature as they represent theory-rich and data-intensive parts of modeling.

We propose a hybrid architecture consisting of computational and cognitive subsystems corresponding to a decomposed climate model requiring different approaches (Fig. 1). Consequently, models have computation-intensive (HPC) and training-inference-heavy (AI) submodels which are executed and then combined/synthesized into complete solutions.

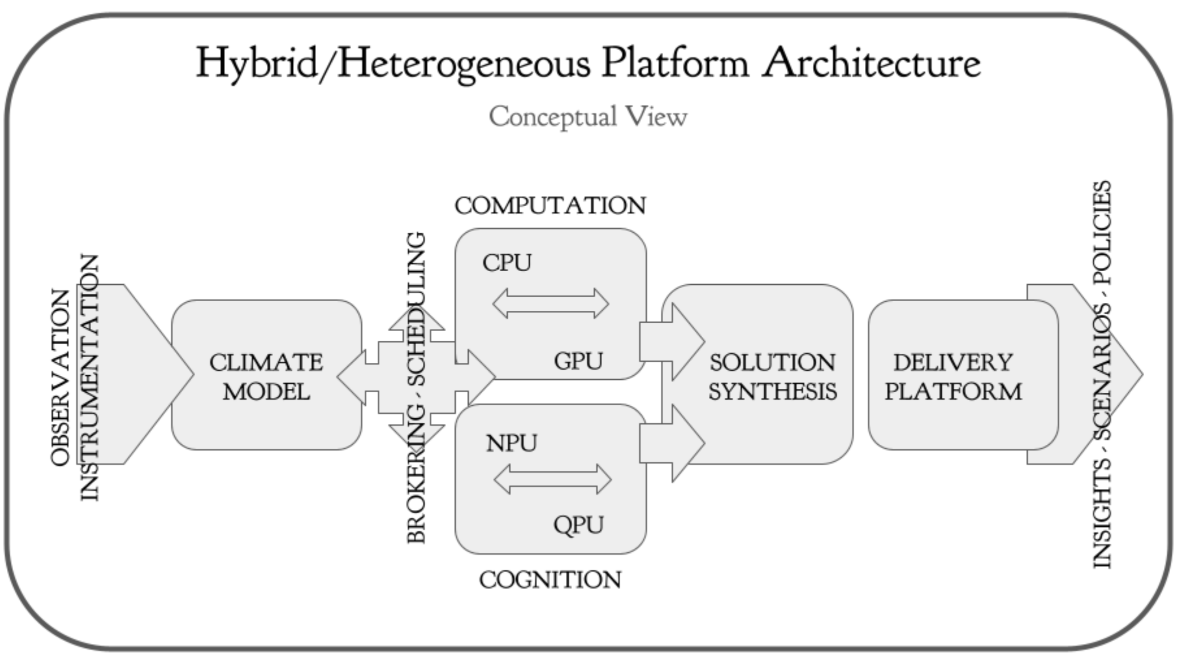

Fig.1 - Climate Modeling Platform Architecture

Goals and objectives for hybrid earth system modeling that have been proposed include (Ref. 3): better prediction of phenomena outside observed data used in training, and of extreme events in particular; simulations that obey physical conservation laws; include measures for self-validation and self-correction; and predictions that are reproducible and interpretable.

Requirements for infrastructure to support hybrid modeling that have been identified include (Ref. 4): design of standard methods to easily link Python and Fortran programs; develop benchmark training data sets tiered in complexity; make better use of heterogeneous hardware (CPUs, TPUs, GPUs); prepare data centers for larger data requests for training, and optimize infrastructure for training and inference close to the data; and enable use and sharing of pre-trained solutions.

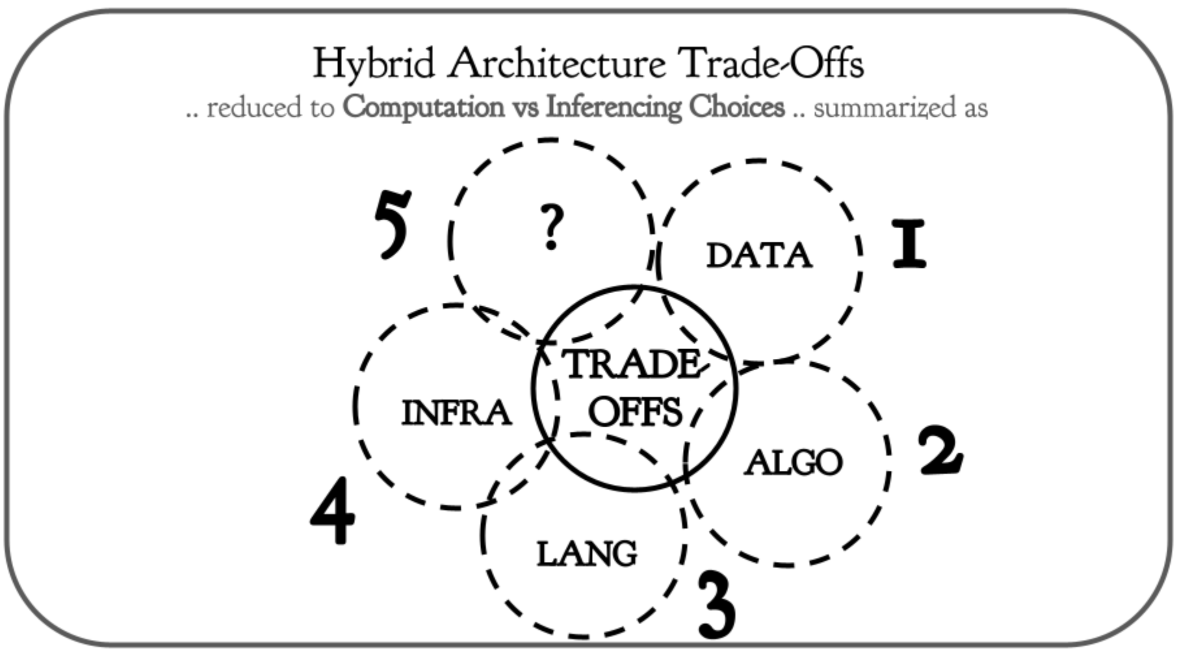

Fig.2 - Hybrid Architecture Trade-Offs

Climate modeling is an extremely complex exercise requiring several trade-offs to be made. We illustrate the possible flow of consideration in Fig. 2. Observational data (1) should be inspected first and checked for accuracy, quality and validity. Volume, velocity and veracity are additional dimensions to be checked. This will influence the choice of algorithms (2) to be deployed for processing. Part of this would be the choice of implementation language (3) since modeling from first principles is usually coded in Fortran, while new AI-related applications are written in Python. While assessing observational data and implementation algorithms we should estimate the scale and type of infrastructure (4) appropriate to previous choices. Finally (5), we should also take into account some non-specified trade-offs, as yet unnamed, but certainly to be found.

We outline the architecture of hyperscale scientific systems (HPC+AI) to address climate modeling challenges, and describe in more detail problem decomposition and synthesis as potentially novel approaches to climate modeling.

References

1. Thomas Tomkins Warner, Numerical Weather and Climate Prediction, Cambridge University Press, 2011

2. David Reichle, The Global Carbon Cycle and Climate Change Scaling Ecological Energetics from Organism to the Biosphere,Elsevier 1st Edition - November 8, 2019

3. Christopher Irrgang, Niklas Boers, Maike Sonnewald, Elizabeth A. Barnes, Christopher Kadow, Joanna Staneva, Jan Saynisch-Wagner: "Will Artificial Intelligence supersede Earth System and Climate Models?", DOI: 10.1038/s42256-021-00374-3; arXiv:2101.09126v1

4. Chantry M, Christensen H, Dueben P, Palmer T: "Opportunities and challenges for machine learning in weather and climate modelling: hard, medium and soft AI", Phil. Trans. R. Soc. A 379: 20200083. https://doi.org/10.1098/rsta.2020.0083