Bridging idealized and operational models to improve the Earth system simulations

Pouria Behnoudfar[1], Charlotte Moser[1], Marc Bocquet[2], Sibo Cheng[2], Nan Chen[1*]

[1]Department of Mathematics, University of Wisconsin-Madison, Wisconsin, USA.
[2]CEREA, ENPC, EDF R&D, Institut Polytechnique de Paris, Ile-de-France, France.
Email addresses:
behnoudfar@wisc.edu (Pouria Behnoudfar);
crmoser2@wisc.edu (Charlotte Moser);
marc.bocquet@enpc.f (Marc Bocquet);
sibo.cheng@enpc.fr (Sibo Cheng);
chennan@math.wisc.edu (Nan Chen).

The results in this article summarize those in [1] "Bridging Idealized and Operational Models: An Explainable AI Framework for Earth System Emulators" (arXiv:2510.13030, https://arxiv.org/abs/2510.13030, 2025) by Pouria Behnoudfar, Charlotte Moser, Marc Bocquet, Sibo Cheng, and Nan Chen.

Background. Computer models are indispensable tools for understanding the Earth system. While high-resolution operational models have achieved many successes, they exhibit persistent biases, particularly in simulating extreme events and statistical distributions. In contrast, coarse-grained idealized models (developed by applied mathematicians) isolate fundamental processes and can be precisely calibrated to excel in characterizing specific dynamical and statistical features. However, different models remain siloed by disciplinary boundaries. See Figure 0.1 that illustrates the chasm between different modeling communities.

New AI approach bridging the gap. By leveraging the complementary strengths of models of varying complexity, an explainable AI framework is developed for Earth system emulators. It bridges the model hierarchy through a reconfigured latent data assimilation technique, uniquely suited to exploit the sparse output from the idealized models. The resulting bridging model inherits the high resolution and comprehensive variables of operational models while achieving global accuracy enhancements through targeted improvements from idealized models. Crucially, the AI's mechanism provides a clear rationale for these advancements, moving beyond black-box correction to physically insightful understanding in a computationally efficient framework that enables effective physics-assisted digital twins and uncertainty quantification. An overview of the method is shown in Figure 0.2.

Correcting El Niño simulation. The power of the framework is demonstrated by substantially improving simulations of El Niño-Southern Oscillation (ENSO) spatiotemporal diversity and complexity in state-of-the-art CMIP6 climate models. ENSO, a dominant mode of interannual climate variability, exhibits diverse spatial patterns and temporal evolutions that have proven challenging for operational models to capture accurately. One of the most widely used operational model, CESM2 (Community Earth System Model 2), is a high-resolution, comprehensive model that simulates the entire Earth system, including the atmosphere, ocean, land, and ice. It provides a detailed

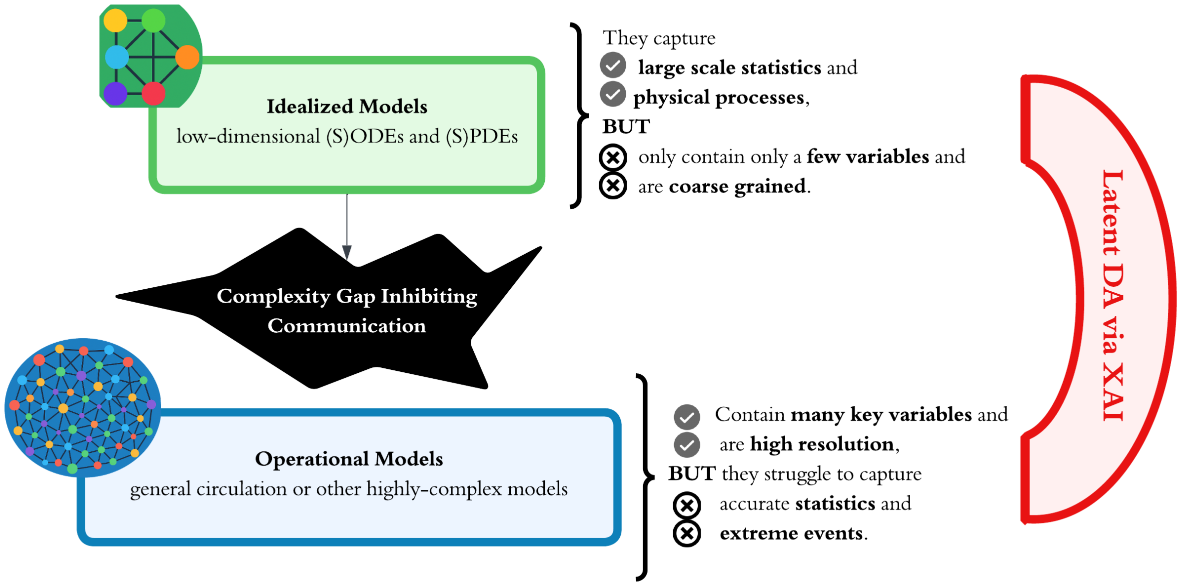

Figure 0.1: Visual summarizing the model hierarchy and the complementary strengths of idealized and operational models. These models have historically been isolated by the large gap in complexity which can now be bridged by the proposed explainable AI framework.

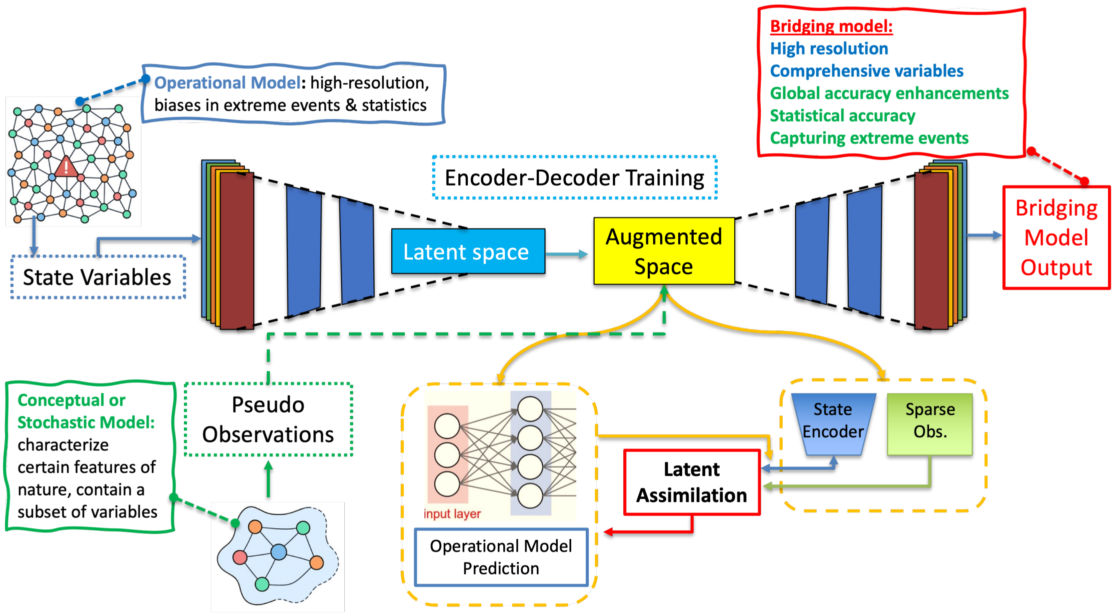

Figure 0.2: Overview of developing the bridging model. Image reproduced from [1].

physical representation but is computationally expensive and suffers from persistent biases, such as misrepresenting the spatial patterns of El Niño. All other operational models have similar or even more significant biases. In contrast, the idealized model, CF23, is a coarse-grained, intermediate-coupled model focused solely on fundamental atmosphere-ocean processes in the equatorial Pacific. It is reduced to one spatial dimension (the equator) and a lower resolution, making it computationally cheap and precisely tunable to accurately capture key large-scale statistics of ENSO, which it provides as sparse but targeted "pseudo-observations" for the bridging framework.

The resulting bridging model not only corrects statistical biases but also reproduces the realistic diversity of El Niño events. It successfully generates different types of El Niño events, including extreme eastern Pacific events and multi-year events that closely match observational records. See Figure 0.4 for a comparison of different models. A key advantage of the resulting bridging model is its computational efficiency. The bridging model can complete a 42-year simulation matching the observational record in just 8 minutes and 20 seconds on standard computing resources, compared to the days or weeks required for traditional operational model runs. This efficiency enables the generation of massive ensembles crucial for robust uncertainty quantification, particularly for assessing extreme events that are poorly sampled in short observational records.

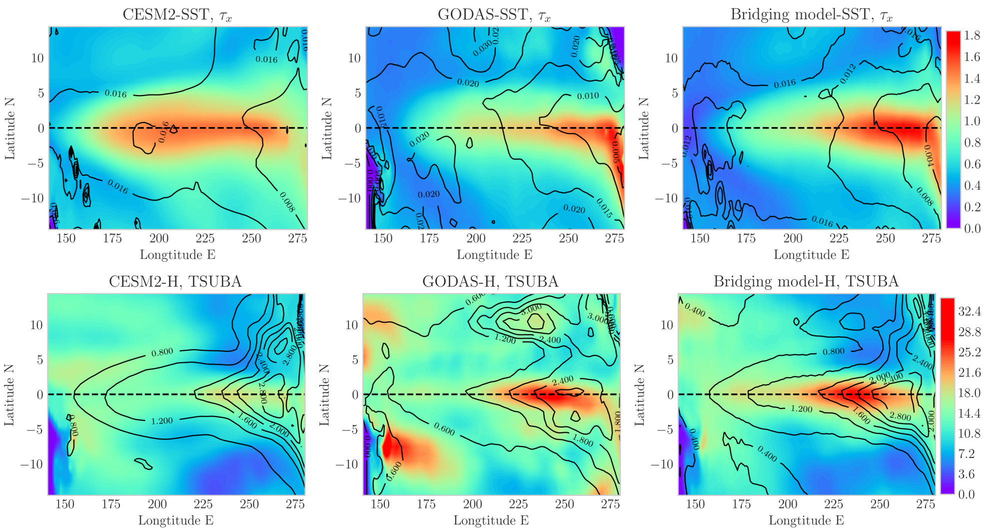

Figure 0.3: Comparison of different variability patterns, measured by the standard deviation, among CESM2, GODAS reanalysis, and the proposed bridging model. The left, middle, and right columns show results from CESM2, GODAS, and the bridging model, respectively. The top row displays sea surface temperature (SST, shading) and zonal wind stress (τ_x, contours). The bottom row shows thermocline depth (H, shading) and subsurface temperature (TSUBA, contours). Image reproduced from [1].

Examining "what-if" scenarios. The framework also functions as a digital twin for sensitivity analysis, allowing researchers to efficiently test "what-if" scenarios by modifying parameters in the cheap idealized model and propagating these changes through the full system. Figure 0.4 for the sensitivity of the bridging model to decadal variability.

Broader impact. The study highlights the importance of enhancing communication between modeling communities that have traditionally operated in relative isolation. By providing a concrete mechanism for translating idealized model insights into operational model corrections, the framework creates new incentives for collaboration.

The output generated by the bridging model is not merely a corrected version of operational simulation. It is a new hybrid data product that inherits high-resolution physical fields while being

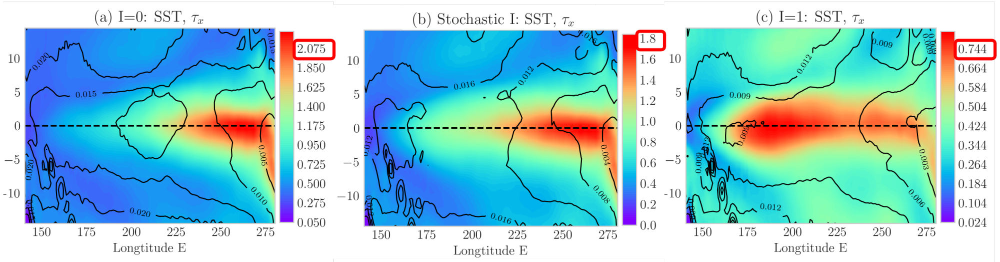

Figure 0.4: Sensitivity of the bridging model's variability to the decadal variability parameter I in the idealized model (CF23), a surrogate for Walker circulation strength. Image reproduced from [1].

globally informed by targeted statistical accuracy.

Notably, while demonstrated for climate science, the framework presents a generalizable strategy for integrating multi-fidelity models in any complex dynamical system where models of varying complexity provide complementary insights.