ADVANCING EARTH AND SPACE SCIENCES

AGU Advances

COMMENTARY
10.1029/2025AV001716

Why Idealized Models Are More Important Than Ever in Earth System Science

Kevin A. Reed[1], Brian Medeiros[2], Christiane Jablonowski[3], Isla R. Simpson[2], Aiko Voigt[4], and Allison A. Wing[5]

[1]School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, NY, USA, [2]NSF National Center for Atmospheric Research, Boulder, CO, USA, [3]Department of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, MI, USA, [4]Department of Meteorology and Geophysics, University of Vienna, Vienna, Austria, [5]Department of Earth, Ocean and Atmospheric Science, Florida State University, Tallahassee, FL, USA

Peer Review The peer review history for this article is available as a PDF in the Supporting Information.

Key Points:

- Idealized models enable imaginative science that push researchers to think in new and creative ways making them useful educational tools
- Reduced complexity configurations remain important tools for model development in Earth system science
- We call on the scientific community to re-emphasize model hierarchies to aid in the continued understanding of the Earth system

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

K. A. Reed,
kevin.reed@stonybrook.edu

Citation:

Reed, K. A., Medeiros, B., Jablonowski, C., Simpson, I. R., Voigt, A., & Wing, A. A. (2025). Why idealized models are more important than ever in Earth system science. *AGU Advances*, *6*, e2025AV001716. https://doi.org/10.1029/2025AV001716

Received 5 MAR 2025
Accepted 14 JUN 2025

Author Contributions:

Conceptualization: Kevin A. Reed, Brian Medeiros, Christiane Jablonowski, Isla R. Simpson, Aiko Voigt, Allison A. Wing
Funding acquisition: Kevin A. Reed
Project administration: Kevin A. Reed
Visualization: Brian Medeiros
Writing – original draft: Kevin A. Reed, Brian Medeiros
Writing – review & editing: Kevin A. Reed, Brian Medeiros, Christiane Jablonowski, Isla R. Simpson, Aiko Voigt, Allison A. Wing

Abstract Simulating the Earth system is crucial for studying Earth's climate and how it changes. Modeling approaches that simplify the Earth system while retaining key characteristics are important tools to advance understanding. The simplicity and flexibility of idealized models enables imaginative science and makes them powerful educational tools. Evolving scientific community needs and increasing model complexity, however, makes it challenging to maintain and support idealized configurations in cutting-edge Earth system modeling frameworks. We call on the scientific community to re-emphasize model hierarchies within these frameworks to aid in understanding the Earth system, advancing model development, and developing the future workforce.

1. Growing Complexity

Numerical representation of the Earth system is a primary tool to advance understanding of Earth's climate and how it changes over time. Due to spatial and temporal sparsity, inhomogeneity of data records, and the inherent complexity of deciphering such measurements, scientists routinely rely on numerical models to probe atmospheric, oceanic and land processes, as well as the interactions among them across scales. General circulation models (GCMs) that represent the coupled atmosphere-ocean system are employed to advance our knowledge of past climate, as well as inform possible future climate states using plausible socioeconomic scenarios. In particular, modeling groups around the world make use of the Coupled Model Intercomparison Project (CMIP; Eyring et al., 2016) to standardize simulation protocols as well as apprize international assessments such as the United Nations Intergovernmental Panel on Climate Change (IPCC, 2021). Naturally, as understanding of the Earth system and high-performance computing infrastructures have advanced, so too has the complexity of GCMs. In recent years, these intercomparisons have expanded to include Earth System Models (ESMs) that incorporate many components of the Earth system including the cyrosphere, carbon cycle, and biogeochemical processes and continue the trend toward ever growing complexity. Steady progress has made climate projection a quasi-operational activity, leading to a tension between curiosity-driven research and actionable climate prediction (Stevens, 2024).

With increased complexity, however, it can become difficult to understand model solutions or perform targeted simulations with these models to gain additional understanding. This situation led past efforts to implement hierarchies of GCMs as tools for physical understanding and model development (Bony et al., 2013; Held, 2005; Jeevanjee et al., 2017; Maher et al., 2019; Mansfield et al., 2023; Shaw & Stevens, 2025). The hierarchy of the atmospheric component of GCMs, for example, can span from dry dynamical core tests, to idealized boundary forcing experiments such as aquaplanet frameworks, to comprehensive climate simulations typical of CMIP. A chief advantage of such idealizations is to allow numerical experiments targeting hypotheses about specific phenomena or processes while retaining salient properties of the climate system. Simplified settings can increase the signal-to-noise ratio and often require less computational resources than more realistic, CMIP-class, simulations.

As an example, consider GCMs configured as aquaplanets in which continents are removed and global sea surface temperatures are specified. Aquaplanet simulations have been used widely to understand various atmospheric processes from the structure of the Intertropical Convergence Zone (e.g., Möbis & Stevens, 2012), to storm track variability (e.g., Cash et al., 2002), to challenging classical views of monsoon dynamics (e.g., Bordoni & Schneider, 2008). Since aquaplanet simulations maintain critical aspects of the global circulation of the atmosphere they are also used to investigate climate feedbacks and sensitivity, helping to explain changes observed in more realistic model simulations (e.g., Medeiros et al., 2015; Ringer et al., 2014).

There are also many examples of reduced complexity configurations informing model development. The Dynamical Core Model Intercomparison Project (DCMIP, Ullrich et al., 2017) demonstrates the use of a suite of idealized deterministic tests used for model evaluation (sometimes with analytical solutions), comparison to reference simulations from previous model generations, and investigation of model conservation properties. The idealized tropical cyclone test of Reed and Jablonowski (2012) within DCMIP, for example, has been used to quantify the impact of the GCM atmospheric dynamical core selection on the simulated storm structure and intensity and to explain the differences observed in more realistic CMIP simulations with the same dynamical cores (Reed et al., 2015), for a fraction of the computational cost. Efforts like DCMIP have also had a profound impact on early career researchers through hands-on summer schools in partnership with model developers.

Hierarchies of complexity within components of GCMs and ESMs can be broadly classified by considering simplifications along two aspects of model complexity: physics and forcing. Physics complexity represents the number of processes and feedbacks represented within the model, and the sophistication of those representations. Dynamical core simulations, such as those in DCMIP, include only the dry equations of atmospheric motion. While complex in their own right, they are of low complexity compared to frameworks that attempt to represent the real world as closely as possible and include a myriad of physical processes such as convection, microphysics, radiation, and chemical reactions parameterizations. The Earth system, and atmosphere specifically, is driven by the flux of radiation from above and fluxes between the boundaries of the system's components (e.g., the atmosphere-ocean interface). This represents the forcing complexity component of the hierarchy. In reality, these boundary forcings are intricate: highly variable and non-symmetric. Within the hierarchy, the simplest boundary forcing for a component model is to prescribe a uniform surface condition and complexity can be increased through the inclusion of horizontal gradients and/or of different surface types. Hierarchical approaches afford students and researchers pathways to advance understanding and develop testable hypotheses.

The hierarchy is summarized in Figure 1 and emphasizes the physical climate system and representation of the spatiotemporal variations of climate. As mentioned, there have been significant recent advances that increase physics complexity to more faithfully represent the real world, moving toward a Digital Earth (in analogy to the digital twins concept used in engineering; Jones et al., 2020). These physics-based models provide a basis for reasoning about the Earth system. In addition to idealizations within component models, coupled models may be simplified in numerous ways, including through idealizations of the interactions between components (e.g., Weber, 2010). Conceptual and fundamental models of the system, such as energy balance or analytical models, are denoted at the bottom of Figure 1. The simplest models provide theoretical underpinnings for the findings in both nature and more complex models. The hierarchical approaches of idealizing, or even removing, physics and forcing provide connections between comprehensive system component models and the conceptual models. The hierarchy of models discussed here dovetails into those reduced complexity models through the continuum of complexity, and provides mechanisms for directly evaluating simpler models (Sarofim et al., 2021). It is also worth noting that different nodes of the hierarchy as shown in Figure 1 can have different strengths and applications, for example, some idealized models may not be as useful for predictions (as ESMs are) as they may be for advancing understanding.

In summary, idealized configurations are valuable tools for:

1. **understanding**: Allow for experiments that target specific features of interest, including physical processes and emergent phenomena, while retaining fundamental properties of the climate system.
2. **model development and validation**: Allow for benchmarking of model behavior and the elucidation of errors using test cases throughout the model development pipeline.
3. **training**: When the above are combined, the simplicity and flexibility of idealized models also make them powerful educational pathways for learning how to use global models as well as understanding physical characteristics of the Earth system.

2. A Modern Challenge

Climate model development resources are stressed by balancing multiple priorities. Faced with increasing demand for confident climate projections, including risk from extremes, modeling centers are exploring potential

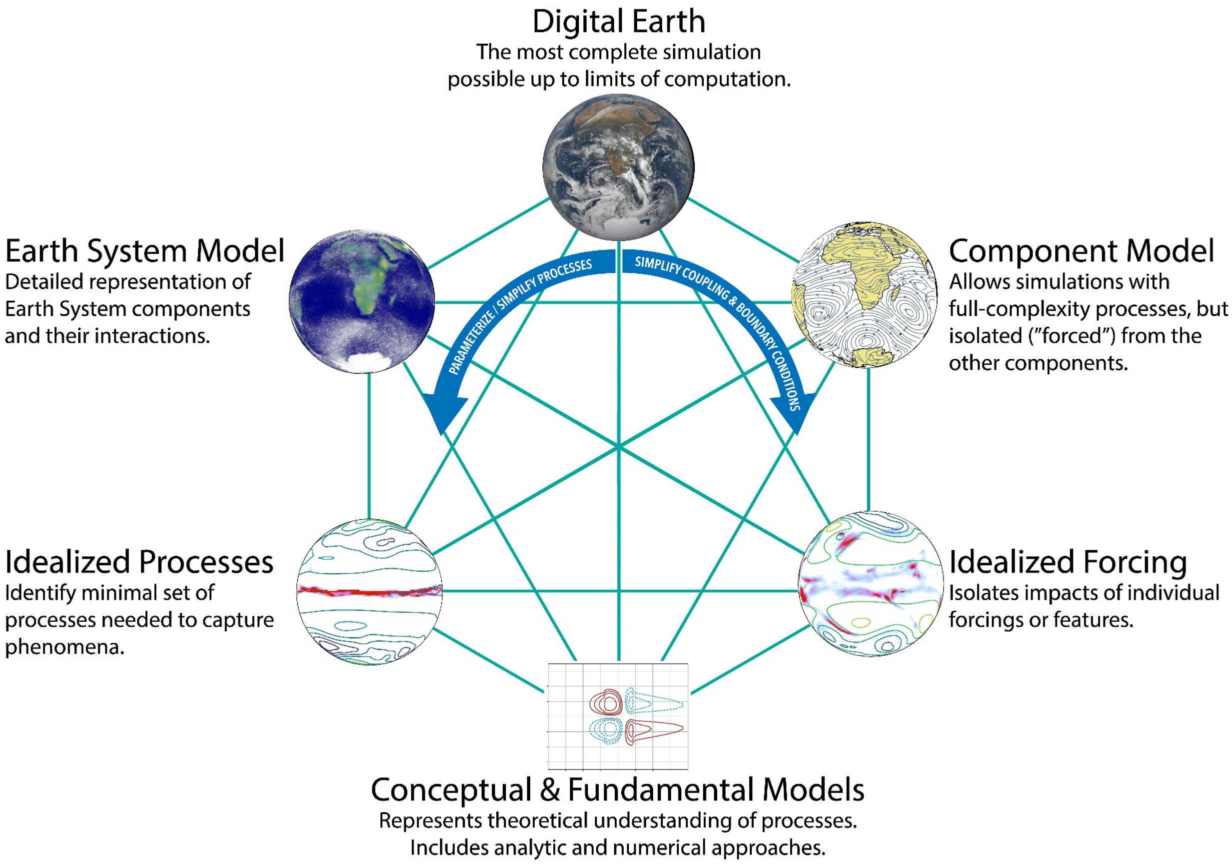

Figure 1. The Earth System Model hierarchy. Nodes depict model variants, and lines connecting hierarchy nodes indicate how ideas and advances propagate across the many facets of the model complexity. Approaches of simplifying processes (including parameterizations) or boundary forcing (including coupling) traverse the hierarchy roughly anticlockwise and clockwise, respectively (demonstrated by the arrows).

next-generation approaches. In the immediate future, CMIP continues to its seventh phase, pressuring centers to update existing models. At the same time, machine learning approaches are gaining traction with the anticipation of good fidelity at reduced computational cost. At the other end of the spectrum, a substantial effort is being devoted to pushing to very high resolution numerical models in the hopes of resolving small-scale weather features well enough to alleviate climate-scale biases. With limited resources and time, development efforts often neglect large portions of the hierarchy in order to demonstrate new capabilities, or the potential for future breakthroughs.

Recent efforts in global storm-resolving models or km-scale models, often referred to as global cloud-resolving models (GCRMs), provides an illustrative example. With horizontal grid spacings of less than 5 km, models participating in the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains (DYAMOND; Stevens et al., 2019) intercomparison push the boundaries of current computational resources. When compared to coarse CMIP models (grid spacing 50–250 km), the DYAMOND simulations require immense computational resources (30,000+ times). The simulations also pose potentially disruptive challenges to storing and analyzing the voluminous output (Schär et al., 2020). DYAMOND-class models shift complexity by reducing the empiricism of parameterizations in favor of explicit fluid dynamics while the massive increase in resolution permits more complexity in the model solution. The computational expense of GCRM simulations means that there are few of them and they are usually short. Envelope-pushing efforts are now stretching to multi-year simulations of the coupled system (e.g., Rackow et al., 2024; Segura et al., 2025). High resolution provides detailed circulation features, including tropical cyclones (Judt et al., 2021), organized convection (Bao et al., 2024), and the fractal nature of clouds (Christensen & Driver, 2021). Vexingly, however, many long-standing biases can remain, such as the spatial distribution of precipitation (Zhou et al., 2022) and continued effort will be needed to reduce them. It is apparent that additional experimentation and analysis of this emerging class of models is needed, as the increased resolution adds complexity in another aspect than physics or forcing as defined above. *Thus, the time is right to reemphasize using idealized settings with these models to advance our understanding of the physical processes and aid future model development toward incorporation of these higher-resolution models within ESMs.*

Another recent trend in climate model development is to employ machine learning methods that supplement or replace traditional schemes. This is a rapidly evolving area, and approaches range from emulating the entire atmosphere (Watt-Meyer et al., 2023) to hybrid approaches that couple a dynamical core to a physics emulator (Kochkov et al., 2024) to methods that learn individual processes to replace particular parameterizations (Gettelman et al., 2021). One approach is to train models using global km-scale models to develop hybrid GCMs that would allow for running ensembles of simulations that might better meet societal needs (e.g., Schneider et al., 2023; Eyring et al., 2024). Some of the efforts have used idealized configurations, especially aquaplanets, for training and evaluation (e.g., Brenowitz & Bretherton, 2018; Gentine et al., 2018), but a hierarchical approach has not yet emerged. *Experiments subjecting such models to idealized settings is an attractive way to objectively test these methods. They would provide out-of-sample evaluation of the models, and idealized settings are likely to easily expose some aspects of overfitting to training data*. For example, if an aquaplanet were to retain geographic features (e.g., orographic rainfall) in simulations with machine learning emulators, it could expose concerns with the approach.

At traditional GCM resolutions, development efforts tend to focus on increasing physics and component complexity using available computational resources. As such, CMIP models have expanded the scale of their code base, grid resolutions, and number of simulations all leading to increased computational requirements (Alexander & Easterbrook, 2015; Balaji et al., 2017). CMIP6 included over 20 specific MIPs with 190 different experiments producing approximately 40 PB of data (Acosta et al., 2024) and modeling centers are now using large-ensemble simulations to quantify, which can advance our understanding of, internal variability and climate risks (Deser et al., 2020). Without intentional software design practices in place, the increasing complexity poses a threat to the configurability of models and thereby introduces obstacles for scientific exploration, and limits the ability of individual scientists outside of modeling centers to run them. This challenge is made worse by the very fact that growth in computational resources has slowed in recent years and the conversion of existing code bases to take advantage of new architectures (i.e., GPUs) is slow, especially on current model development cycle timelines (Donahue et al., 2024; Porter & Heimbach, 2024). Thus, it is often challenging to support hierarchies in these modeling frameworks due to:

1. the ever expanding coupling and complexity of models,
2. limited software engineering resources, and
3. the compromises made to balance the demands for higher resolution and larger ensembles against the desire for configurability and hierarchical capabilities.

3. A Path Forward

The simplicity and flexibility of idealized models enables imaginative science that pushes researchers to think in new and creative ways and makes such models powerful educational tools. We call on the scientific community to re-emphasize model hierarchies within GCM and ESM frameworks, and expand them in GCRM frameworks, to aid in the continued understanding of the Earth system, as well as to advance model development and workforce development needs. We recommend the following:

3.1. Extend Idealized Frameworks Across Earth System Components to Include Coupled Processes

Many existing model hierarchies focus on the atmospheric component of GCMs with some expanding these frameworks to coupled atmosphere-ocean GCMs. These coupled approaches have provided fundamental insights into important processes of the Earth system. Examples include quantifying the relative contributions of the atmosphere and ocean to heat and moisture transport in the Earth system (Czaja & Marshall, 2006), the role of ocean coupling in the position of the Intertropical Convergence Zone (Fučkar et al., 2013), and interannual variability (Wu et al., 2021). There have also been efforts to develop reduced physics complexity and other ESM components, such as simplified land models (Laguë et al., 2019), but software interoperability needs to be made a priority.

3.2. Expand the Utility for Comparisons of Idealized Configurations Across Model Types

Idealizations can enable new pathways for GCMs and global km-scale models to be configured in consistent ways with other model types, such as limited domain models, that are challenging in more realistic settings. One example of this is the Radiative Convective Equilibrium (RCE) Model Intercomparison Project (RCEMIP; Wing et al., 2018, 2024). RCE, the statistical balance in the atmosphere between radiative cooling and latent heating, is a good approximation of Earth's climate, particularly tropical climate, making it an ideal framework to span different model types. RCEMIP has identified some important differences between the simulated RCE state in GCMs and in CRMs pointing to deficiencies in GCM convection parameterizations (Becker & Wing, 2020; Gasparini et al., 2024; O'Donnell et al., 2024; Wing et al., 2020).

3.3. Enable the Continued Use of These Frameworks to Help Train the Earth System Science Community

As models become more complex it remains as important as ever to ensure that the community has the resources to idealize models as classroom tools and for education-focused summer schools. These have direct impacts not only on early career professionals with a keen interest in model development, but the idealized configurations within a model hierarchy also provide an important tool for developing sound scientific hypotheses that can be tested as part of dissertations and research projects. In all cases the training benefits the wider scientific community through the continued improvement of models, the use of new scientific tools such as machine learning, and the advancement of knowledge. Recently, the Community ESM has expanded training resources to include simpler models within that GCM (https://www.cesm.ucar.edu/models/simple) and have developed tools to facilitate experiment configuration (https://github.com/ESMCI/visualCaseGen).

The community should develop more flexible, well-documented models, fund creative projects that make use of model hierarchies, and support training activities that engage the next generation of Earth system scientists in these efforts.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

Data were not used, nor created for this research.

Acknowledgments
KAR acknowledges support from the National Science Foundation (Grant AGS-2327958) with additional support from the U.S. Department of Energy (DOE) Office of Science Award Number DE-SC0016605 "A Framework for Improving Analysis and Modeling of Earth System and Intersectoral Dynamics at Regional Scales." BM additionally acknowledges support by the U.S. DOE, Office of Science, Office of Biological & Environmental Research, Regional and Global Model Analysis component of the Earth and Environmental System Modeling Program under Award Number DE-SC0022070, National Science Foundation IA 1947282 and NSF CSSI award 2311376. IRS was also supported by the National Science Foundation (Grant OAC-2311376). This work was also supported by the National Center for Atmospheric Research (NCAR) and acknowledges support from NSF CSSI awards 2004575 and 2311376, which is a major facility sponsored by the NSF under Cooperative Agreement No. 1852977. CJ is supported via the National Oceanic and Atmospheric Administration (Grant NA22OAR4320150) and National Science Foundation (Grant AGS-2332468). AV acknowledges support from the Austrian Research Promotion Agency (FFG) ACRP project "HighResLearn." AAW acknowledges support from the National Science Foundation (Grant AGS-2140419).

References

Acosta, M. C., Palomas, S., Paronuzzi Ticco, S. V., Utrera, G., Biercamp, J., Bretonniere, P.-A., et al. (2024). The computational and energy cost of simulation and storage for climate science: Lessons from CMIP6. *Geoscientific Model Development*, *17*(8), 3081–3098. https://doi.org/10.5194/gmd-17-3081-2024

Alexander, K., & Easterbrook, S. M. (2015). The software architecture of climate models: A graphical comparison of CMIP5 and EMICAR5 configurations. *Geoscientific Model Development*, *8*(4), 1221–1232. https://doi.org/10.5194/gmd-8-1221-2015

Balaji, V., Maisonnave, E., Zadeh, N., Lawrence, B. N., Biercamp, J., Fladrich, U., et al. (2017). CPMIP: Measurements of real computational performance of Earth system models in CMIP6. *Geoscientific Model Development*, *10*(1), 19–34. https://doi.org/10.5194/gmd-10-19-2017

Bao, J., Stevens, B., Kluft, L., & Muller, C. (2024). Intensification of daily tropical precipitation extremes from more organized convection. *Science Advances*, *10*(8), eadj6801. https://doi.org/10.1126/sciadv.adj6801

Becker, T., & Wing, A. A. (2020). Understanding the extreme spread in climate sensitivity within the radiative-convective equilibrium model intercomparison project. *Journal of Advances in Modeling Earth Systems*, *12*(10), e2020MS002165. https://doi.org/10.1029/2020MS002165

Bony, S., Stevens, B., Held, I. H., Mitchell, J. F., Dufresne, J.-L., Emanuel, K. A., et al. (2013). Carbon dioxide and climate: Perspectives on a scientific assessment. In G. R. Asrar & J. W. Hurrell (Eds.), *Climate Science for Serving Society* (pp. 391–413). Springer. https://doi.org/10.1007/978-94-007-6692-1_14

Bordoni, S., & Schneider, T. (2008). Monsoons as eddy-mediated regime transitions of the tropical overturning circulation. *Nature Geoscience*, *1*(8), 515–519. https://doi.org/10.1038/ngeo248

Brenowitz, N. D., & Bretherton, C. S. (2018). Prognostic validation of a neural network unified physics parameterization. *Geophysical Research Letters*, *45*(12), 6289–6298. https://doi.org/10.1029/2018GL078510

Cash, B. A., Kushner, P. J., & Vallis, G. K. (2002). The structure and composition of the annular modes in an aquaplanet general circulation model. *Journal of the Atmospheric Sciences*, *59*(23), 3399–3414. https://doi.org/10.1175/1520-0469(2002)059<3399:TSACOT>2.0.CO;2

Christensen, H. M., & Driver, O. G. A. (2021). The fractal nature of clouds in global storm-resolving models. *Geophysical Research Letters*, *48*(23), e2021GL095746. https://doi.org/10.1029/2021GL095746

Czaja, A., & Marshall, J. (2006). The partitioning of poleward heat transport between the atmosphere and ocean. *Journal of the Atmospheric Sciences*, *63*(5), 1498–1511. https://doi.org/10.1175/JAS3695.1

Deser, C., Lehner, F., Rodgers, K. B., Ault, T., Delworth, T. L., DiNezio, P. N., et al. (2020). Insights from Earth system model initial-condition large ensembles and future prospects. *Nature Climate Change*, *10*(4), 277–286. https://doi.org/10.1038/s41558-020-0731-2

Donahue, A. S., Caldwell, P. M., Bertagna, L., Beydoun, H., Bogenschutz, P. A., Bradley, A. M., et al. (2024). To exascale and beyond—The Simple Cloud-Resolving E3SM Atmosphere Model (SCREAM), a performance portable global atmosphere model for cloud-resolving scales. *Journal of Advances in Modeling Earth Systems*, *16*(7), e2024MS004314. https://doi.org/10.1029/2024ms004314

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Eyring, V., Gentine, P., Camps-Valls, G., Lawrence, D. M., & Reichstein, M. (2024). AI-empowered next-generation multiscale climate modelling for mitigation and adaptation. *Nature Geoscience*, *17*(10), 963–971. https://doi.org/10.1038/s41561-024-01527-w

Fučkar, N. S., Xie, S.-P., Farneti, R., Maroon, E. A., & Frierson, D. M. W. (2013). Influence of the extratropical ocean circulation on the intertropical convergence zone in an idealized coupled general circulation model. *Journal of Climate*, *26*(13), 4612–4629. https://doi.org/10.1175/JCLI-D-12-00294.1

Gasparini, B., Mandorli, G., Stubenrauch, C., & Voigt, A. (2024). Basic physics predicts stronger high cloud radiative heating with warming. *Geophysical Research Letters*, *51*(24), e2024GL111228. https://doi.org/10.1029/2024GL111228

Gentine, P., Pritchard, M., Rasp, S., Reinaudi, G., & Yacalis, G. (2018). Could machine learning break the convection parameterization deadlock? *Geophysical Research Letters*, *45*(11), 5742–5751. https://doi.org/10.1029/2018GL078202

Gettelman, A., Gagne, D. J., Chen, C.-C., Christensen, M. W., Lebo, Z. J., Morrison, H., & Gantos, G. (2021). Machine learning the warm rain process. *Journal of Advances in Modeling Earth Systems*, *13*(2), e2020MS002268. https://doi.org/10.1029/2020MS002268

Held, I. M. (2005). The gap between simulation and understanding in climate modeling. *Bulletin of the American Meteorological Society*, *86*(11), 1609–1614. https://doi.org/10.1175/BAMS-86-11-1609

IPCC. (2021). *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.

Jeevanjee, N., Hassanzadeh, P., Hill, S., & Sheshadri, A. (2017). A perspective on climate model hierarchies. *Journal of Advances in Modeling Earth Systems*, *9*(4), 1760–1771. https://doi.org/10.1002/2017MS001038

Jones, D., Snider, C., Nassehi, A., Yon, J., & Hicks, B. (2020). Characterising the Digital Twin: A systematic literature review. *CIRP Journal of Manufacturing Science and Technology*, *29*, 36–52. https://doi.org/10.1016/j.cirpj.2020.02.002

Judt, F., Klocke, D., Rios-Berrios, R., Vanniere, B., Ziemen, F., Auger, L., et al. (2021). Tropical cyclones in global storm-resolving models. *Journal of the Meteorological Society of Japan. Ser. II*, *99*(3), 579–602. https://doi.org/10.2151/jmsj.2021-029

Kochkov, D., Yuval, J., Langmore, I., Norgaard, P., Smith, J., Mooers, G., et al. (2024). Neural general circulation models for weather and climate. *Nature*, *632*(8027), 1060–1066. https://doi.org/10.1038/s41586-024-07744-y

Laguë, M. M., Bonan, G. B., & Swann, A. L. S. (2019). Separating the impact of individual land surface properties on the terrestrial surface energy budget in both the coupled and uncoupled land–atmosphere system. *Journal of Climate*, *32*(18), 5725–5744. https://doi.org/10.1175/JCLI-D-18-0812.1

Maher, P., Gerber, E. P., Medeiros, B., Merlis, T. M., Sherwood, S., Sheshadri, A., et al. (2019). Model hierarchies for understanding atmospheric circulation. *Reviews of Geophysics*, *57*(2), 250–280. https://doi.org/10.1029/2018RG000607

Mansfield, L. A., Gupta, A., Burnett, A. C., Green, B., Wilka, C., & Sheshadri, A. (2023). Updates on model hierarchies for understanding and simulating the climate system: A focus on data-informed methods and climate change impacts. *Journal of Advances in Modeling Earth Systems*, *15*(10), e2023MS003715. https://doi.org/10.1029/2023MS003715

Medeiros, B., Stevens, B., & Bony, S. (2015). Using aquaplanets to understand the robust responses of comprehensive climate models to forcing. *Climate Dynamics*, *44*(7–8), 1957–1977. https://doi.org/10.1007/s00382-014-2138-0

Möbis, B., & Stevens, B. (2012). Factors controlling the position of the Intertropical Convergence Zone on an aquaplanet. *Journal of Advances in Modeling Earth Systems*, *4*(4), 2012MS000199. https://doi.org/10.1029/2012MS000199

O'Donnell, K. L., Tomiczek, T., Higgins, A., Munoz, S., & Scyphers, S. (2024). Stakeholder driven sensor deployments to characterize chronic coastal flooding in key West Florida. *Earth's Future*, *12*(7), e2023EF003631. https://doi.org/10.1029/2023EF003631

Porter, A., & Heimbach, P. (2024). Unlocking the power of parallel computing: GPU technologies for ocean forecasting. *State of the Planet*. https://doi.org/10.5194/sp-2024-32

Rackow, T., Pedruzo-Bagazgoitia, X., Becker, T., Milinski, S., Sandu, I., Aguridan, R., et al. (2024). Multi-year simulations at kilometre scale with the Integrated Forecasting System coupled to FESOM2.5/NEMOv3.4. *EGUsphere*. https://doi.org/10.5194/egusphere-2024-913

Reed, K. A., Bacmeister, J. T., Rosenbloom, N. A., Wehner, M. F., Bates, S. C., Lauritzen, P. H., et al. (2015). Impact of the dynamical core on the direct simulation of tropical cyclones in a high-resolution global model. *Geophysical Research Letters*, *42*(9), 3603–3608. https://doi.org/10.1002/2015GL063974

Reed, K. A., & Jablonowski, C. (2012). Idealized tropical cyclone simulations of intermediate complexity: A test case for AGCMs. *Journal of Advances in Modeling Earth Systems*, *4*(2), 2011MS000099. https://doi.org/10.1029/2011MS000099

Ringer, M. A., Andrews, T., & Webb, M. J. (2014). Global-mean radiative feedbacks and forcing in atmosphere-only and coupled atmosphere-ocean climate change experiments. *Geophysical Research Letters*, *41*(11), 4035–4042. https://doi.org/10.1002/2014GL060347

Sarofim, M. C., Martinich, J., Neumann, J. E., Willwerth, J., Kerrich, Z., Kolian, M., et al. (2021). A temperature binning approach for multi-sector climate impact analysis. *Climatic Change*, *165*(1–2), 22. https://doi.org/10.1007/s10584-021-03048-6

Schär, C., Fuhrer, O., Arteaga, A., Ban, N., Charpilloz, C., Di Girolamo, S., et al. (2020). Kilometer-scale climate models: Prospects and challenges. *Bulletin of the American Meteorological Society*, *101*(5), E567–E587. https://doi.org/10.1175/BAMS-D-18-0167.1

Schneider, T., Behera, S., Boccaletti, G., Deser, C., Emanuel, K., Ferrari, R., et al. (2023). Harnessing AI and computing to advance climate modelling and prediction. *Nature Climate Change*, *13*(9), 887–889. https://doi.org/10.1038/s41558-023-01769-3

Segura, H., Pedruzo-Bagazgoitia, X., Weiss, P., Müller, S. K., Rackow, T., Lee, J., et al. (2025). nextGEMS: Entering the era of kilometer-scale Earth system modeling. *EGUsphere*. https://doi.org/10.5194/egusphere-2025-509

Shaw, T. A., & Stevens, B. (2025). The other climate crisis. *Nature*, *639*(8056), 877–887. https://doi.org/10.1038/s41586-025-08680-1

Stevens, B. (2024). A perspective on the future of CMIP. *AGU Advances*, *5*(1), e2023AV001086. https://doi.org/10.1029/2023AV001086

Stevens, B., Satoh, M., Auger, L., Biercamp, J., Bretherton, C. S., Chen, X., et al. (2019). DYAMOND: The DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains. *Progress in Earth and Planetary Science*, *6*(1), 61. https://doi.org/10.1186/s40645-019-0304-z

Ullrich, P. A., Jablonowski, C., Kent, J., Lauritzen, P. H., Nair, R., Reed, K. A., et al. (2017). DCMIP2016: A review of non-hydrostatic dynamical core design and intercomparison of participating models. *Geoscientific Model Development*, *10*(12), 4477–4509. https://doi.org/10.5194/gmd-10-4477-2017

Watt-Meyer, O., Dresdner, G., McGibbon, J., Clark, S. K., Henn, B., Duncan, J., et al. (2023). ACE: A fast, skillful learned global atmospheric model for climate prediction (Version 2). *arXiv*. https://doi.org/10.48550/ARXIV.2310.02074

Weber, S. L. (2010). The utility of Earth system Models of Intermediate Complexity (EMICs). *WIREs Climate Change*, *1*(2), 243–252. https://doi.org/10.1002/wcc.24

Wing, A. A., Reed, K. A., Satoh, M., Stevens, B., Bony, S., & Ohno, T. (2018). Radiative–convective equilibrium model intercomparison project. *Geoscientific Model Development*, *11*(2), 793–813. https://doi.org/10.5194/gmd-11-793-2018

Wing, A. A., Silvers, L. G., & Reed, K. A. (2024). RCEMIP-II: Mock-Walker simulations as phase II of the radiative–convective equilibrium model intercomparison project. *Geoscientific Model Development*, *17*(16), 6195–6225. https://doi.org/10.5194/gmd-17-6195-2024

Wing, A. A., Stauffer, C. L., Becker, T., Reed, K. A., Ahn, M., Arnold, N. P., et al. (2020). Clouds and convective self-aggregation in a multimodel ensemble of radiative-convective equilibrium simulations. *Journal of Advances in Modeling Earth Systems*, *12*(9), e2020MS002138. https://doi.org/10.1029/2020MS002138

Wu, X., Reed, K. A., Wolfe, C. L. P., Marques, G. M., Bachman, S. D., & Bryan, F. O. (2021). The dependence of tropical modes of variability on zonal asymmetry. *Geophysical Research Letters*, *48*(17), e2021GL093966. https://doi.org/10.1029/2021GL093966

Zhou, W., Leung, L. R., & Lu, J. (2022). Linking large-scale double-ITCZ bias to local-scale drizzling bias in climate models. *Journal of Climate*, *35*(24), 7965–7979. https://doi.org/10.1175/JCLI-D-22-0336.1