Advances in Atmospheric Sciences

Impact Factor: 5.8

E-alert

ISSN 0256-1530

CN 11-1925/O4

Toward a Learnable Climate Model in the Artificial Intelligence Era

Gang HUANG, Ya WANG, Yoo–Geun HAM, Bin MU, Weichen TAO, Chaoyang XIE

Citation: Huang, G., Y. Wang, Y.–G. Ham, B. Mu, W. C. Tao, and C. Y. Xie 2024: Toward a Learnable Climate Model in the Artificial Intelligence Era, *Adv. Atmos. Sci.*, 41, 1281–1288. doi: 10.1007/s00376–024–3305–9.

View online: https://doi.org/10.1007/s00376–024–3305–9

ADVANCES IN ATMOSPHERIC SCIENCES, VOL. 41, JULY 2024, 1281–1288

• Perspectives •

Toward a Learnable Climate Model in the Artificial Intelligence Era※

Gang HUANG*[1,2,3], Ya WANG*[1], Yoo-Geun HAM[4], Bin MU[5], Weichen TAO[1], and Chaoyang XIE[1]

[1]*State key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics*, *Institute of Atmospheric Physics*, *Chinese Academy of Sciences*, *Beijing* 100029, *China*

[2]*Laboratory for Regional Oceanography and Numerical Modeling*, *Qingdao National Laboratory for Marine Science and Technology*, *Qingdao* 266237, *China*

[3]*University of Chinese Academy of Sciences*, *Beijing* 100049, *China*

[4]*Department of Environmental Planning*, *Graduate School of Environmental Studies*, *Seoul National University*, *Seoul* 08826, *South Korea*

[5]*School of Software Engineering*, *Tongji University*, *Shanghai* 200092, *China*

(Received 26 November 2023; revised 12 March 2024; accepted 18 March 2024)

ABSTRACT

Artificial intelligence (AI) models have significantly impacted various areas of the atmospheric sciences, reshaping our approach to climate-related challenges. Amid this AI-driven transformation, the foundational role of physics in climate science has occasionally been overlooked. Our perspective suggests that the future of climate modeling involves a synergistic partnership between AI and physics, rather than an "either/or" scenario. Scrutinizing controversies around current physical inconsistencies in large AI models, we stress the critical need for detailed dynamic diagnostics and physical constraints. Furthermore, we provide illustrative examples to guide future assessments and constraints for AI models. Regarding AI integration with numerical models, we argue that offline AI parameterization schemes may fall short of achieving global optimality, emphasizing the importance of constructing online schemes. Additionally, we highlight the significance of fostering a community culture and propose the OCR (Open, Comparable, Reproducible) principles. Through a better community culture and a deep integration of physics and AI, we contend that developing a learnable climate model, balancing AI and physics, is an achievable goal.

Key words: artificial intelligence, deep learning, learnable climate model

Citation: Huang, G., Y. Wang, Y.-G. Ham, B. Mu, W. C. Tao, and C. Y. Xie, 2024: Toward a learnable climate model in the artificial intelligence era. *Adv. Atmos. Sci.*, **41**(7), 1281−1288, https://doi.org/10.1007/s00376-024-3305-9.

1. Artificial intelligence's ascent in climate science

We are in a revolutionary era where artificial intelligence (AI), especially large AI models (Table 1), is redefining how we understand and address climate-related challenges. AI models have swept through most areas of the atmospheric sciences, including but not limited to weather forecasting (Table 1; Pathak et al., 2022; Bi et al., 2023; Chen et al., 2023a; Chen et al., 2023b; Lam et al., 2023), subseasonal and seasonal predictions (Ham et al., 2019, 2021; Pan et al., 2020; Kim et al., 2021; Mu et al., 2021; Ravuri et al., 2021; Ling et al., 2022; Fan et al., 2023; Hess et al., 2022; Zhang et al., 2023; Zhou and Zhang, 2023), extreme weather and climate prediction (e.g., Li et al., 2023), parameterization schemes (Rasp et al., 2018; Han et al., 2020), correction of model simulations (e.g., Pan et al., 2020), and climate detection and attribution (Labe and Barnes, 2021, 2022; Diffenbaugh and Barnes, 2023; Ham et al., 2023; Labe et al., 2023). With its ability to integrate and comprehend data from diverse sources such as satellites, weather stations, and ocean buoys, AI has delivered breakthrough advancements in numerous areas. Its strengths lie in enhancing the overall understanding of atmospheric systems, handling the complexity and nonlinearity of atmospheric dynamics, and offering rapid computational speeds.

While AI blazes forward, the role of physics, often described as the bedrock of climate science, has sometimes been overshadowed. Physical laws and climate theories provide the fundamental principles underpinning our understanding of climate systems. They form the basis of dynamical cli-

※ This paper is a contribution to the special issue on AI Applications in Atmospheric and Oceanic Science: Pioneering the Future.

* Corresponding authors: Gang HUANG, Ya WANG
Emails: hg@mail.iap.ac.cn, wangya@mail.iap.ac.cn

Table 1. Details of large AI weather models.

Model	Forecast Duration	Spatial Resolution	Accuracy (Z500 ACC>0.6)	Forecast Variables	Computational Requirements	Train. Data	Algorithm	Institutions	Initial Release Date	Journal Publication Date and Name	Number of Parameters
GraphCast	10 days	0.25°×0.25°, 6 h	9–10 days	5 surface variables (including precipitation) + 6 atmospheric variables	21 days; 32 TPU v4	ERA5 39yr	GNN	Google DeepMind	2022.12.24, arXiv	2023.11.14, Science	36.7 Million
Pangu-Weather	1 hour–7days	0.25°×0.25°, 1 h	7 days	4 surface variables (excluding precipitation) + 5 atmospheric variables	16 days; 192 V100 GPUs	ERA5 39yr	Transformer	Huawei Cloud	2022.11.3, arXiv	2023.7.5, Nature	256 Million
FourCastNet	3 days	0.25°×0.25°, 6 h	~7 days	5 surface variables (including precipitation) + 4 atmospheric variables	4 A100 GPUs	ERA5 40yr	Adaptive Fourier Neural Operators	Nvidia	2022.2.22, arXiv	N/A	N/A
FuXi	15 days	0.25°×0.25°, 6 h	10–11 days	5 surface variables (including precipitation) + 5 atmospheric variables	~30 hours(pre-training)+2 days(fine-tune); 8 A100 GPUs	ERA5 39yr	U-Transformer	Fudan University	2023.6.22, arXiv	2023.11.16, npj Climate and Atmospheric Science	4.5 Billion
FengWu	10.75 days	0.25°×0.25°, 6 h	10–11 days	4 surface variables (excluding precipitation) + 5 atmospheric variables	17 days; 32 A100 GPUs	ERA5 39yr	Transformer	Shanghai Artificial Intelligence Lab	2023.4.6, arXiv	N/A	N/A
SwinVRNN	5 days	5.625°×5.625°, 6 h	N/A	4 surface variables (including precipitation) + 5 atmospheric variables	8 Tesla V100 GPUs	ERA5 40yr	Swin Transformer	Alibaba Group	2022.5.26, arXiv	N/A	N/A

mate models, offering a comprehensible framework for climate predictions and climate simulations. These models are interpretable, and they have proved their mettle for nearly one hundred years (Bauer et al., 2015). However, in the AI era, interpretability often takes a back seat. AI models are often recognized as "black boxes", which generate predictions and simulations without transparently explaining their involved physical processes. This opaqueness poses huge challenges when it comes to validating AI models and building trust in their generated results. Additionally, AI models lacking physical constraints often face issues such as poor generalization capabilities. This can lead to results that are inconsistent with physical properties and highly unstable, particularly in the context of extreme events.

So, where do we go from here? The answer lies in the symbiotic relationship between AI and physics. AI brings unparalleled computational power and the ability to recognize complex patterns, while physics offers transparency, interpretability, and a solid foundation for scientific understanding. The future of climate modeling is not an "either/or" scenario, and is a partnership between these two realms, leveraging AI's computational prowess and the laws of physics' grounding to create a learnable climate model with good interpretability.

2. Physics for AI modeling: limitations and solutions

Recently emerged large AI weather models have begun to rival or surpass traditional numerical weather predictions. This naturally raises the question of whether these AI models adhere to the principles of meteorological and climate theories. Climate theories derived from various mathematical physics equations, like quasi-geostrophic theory and quasi-steady Rossby waves, serve as the foundational structure for climate research (Hoskins and Karoly, 1981; Held et al., 2002; Held, 2019). Some might envisage a scenario where, if an AI model is complex enough, with multiple layers and variables, it might closely replicate observational data, learning and following the implied conservation laws and climate theories present in the observations. (Hakim and Masanam (2023) conducted a crucial analysis on this issue, applying perturbations to Pangu-Weather, and observed that introducing a tropical heat source in the model reproduces the Matsuno–Gill response (Matsuno, 1966; Gill, 1980) in the tropics and the associated extratropical Rossby wave train. This qualitative finding suggests that AI models seem to possess the capability to learn fundamental physical laws and atmospheric dynamics. However, the actual situation is more intricate. Quantitative analyses indicate persistent challenges, such as issues in geostrophic wind balance and the rotational and divergent wind components, despite the abundance of data and high complexity in models like Pangu and Graphcast Weather (Bonavita, 2023; Kochkov et al., 2024).

These assessment conclusions may appear contradictory but are, in fact, reflective of different degrees of evaluation perspectives. Hakim and Masanam (2023) indicated that AI models can qualitatively reproduce certain dynamic modes based on climatic dynamics. Considering that these responses are inherently embedded in the data at the weather and climate scale, learning and qualitatively reproducing such modes are reasonable. On the other hand, quantitative evaluations by Bonavita (2023) and Kochkov et al. (2024) revealed significant disparities in aspects like geostrophic wind and divergent wind compared to traditional models. The absence of these factors in the inputs and outputs of the AI model, coupled with the relatively small magnitude of the divergent winds, makes it relatively natural for such errors to occur.

While these evaluations pertain to weather models, considering the lower complexity, data volume, and parameter quantity in climate models compared to weather models, similar issues need to be addressed in the evaluation and optimization processes of AI climate models. This emphasizes the importance of dynamic diagnostic and physical constraints to AI climate models. First and foremost, a meticulous examination of climate dynamics is essential for analyzing and comprehending the performance of AI models. Current evaluation methods primarily focus on first-order and second-order statistics, but good performance in these statistics does not mean a model is good at reproducing the crucial features of targeted dynamical climate variability, such as ENSO's seasonal phase-locking and asymmetry (Jin, 1996; Wallace et al., 1998; McPhaden et al., 2006). Consequently, there is an imperative need to reevaluate AI simulations, giving prominence not only to fundamental statistics but also to dynamic modes and physical processes. A case in point is the research conducted by Wang et al. (2024), who revealed that generative models can rectify not only climatological sea surface temperature (SST) biases but also markedly mitigate the prevalent excessive westward bias in ENSO SST anomalies, which is a common issue in climate models and has a considerable impact on the ENSO decay and the East Asian climate (Li and Xie, 2012, 2014; Tao et al., 2018, 2019; Jiang et al., 2021). Due to the complexity of the climate system, evaluating climate simulations could be a challenging task. Numerical climate model evaluations have largely relied on the Coupled Model Intercomparison Project (CMIP), allowing scientists from various research domains to thoroughly assess these models based on their climate physics knowledge. To enhance the evaluation of AI simulations, it is imperative to establish a similar comparative project, such as the AI Model Intercomparison Project (AI-MIP). This initiative should construct and release unified dynamic diagnostic tools of varying complexities and benchmarks to facilitate comprehensive assessments. AI-MIP mandates the construction of AI models on a unified and open dataset, and the evaluation and comparison of these models based on standardized metrics. Through AI-MIP, we expect to significantly reduce the costs of data acquisition, model comparison, and learning. Individuals will be able to easily compare their models with state-of-the-art models, accelerating progress in the field.

Furthermore, it is imperative to apply physical constraints informed by climate dynamics to refine AI models. Basic physical knowledge plays a crucial role in providing physical constraints to AI models (Mohan et al., 2019; Beucler et al., 2020). PINNs (Physics-Informed Neural Networks) are a groundbreaking approach where deep-learning (DL) models are informed by physical laws, integrating differential equations directly into the learning process (Raissi et al., 2019). Besides, conservation laws such as energy conservation, mass conservation, and momentum conservation, have always been essential for model development and climate diagnostics.

These constraints can be categorized as soft constraints and hard constraints. Soft constraints do not explicitly impose specific network changes but provide some representations of physical linkages, and physical equations or mathematical constraints. They may encompass connections between different regions, such as teleconnections, or interactions between different variables. One popular soft constraint uses existing physical equations to constrain the loss function (Beucler et al., 2021). Moreover, Mu et al. (2021) offered an example where a graph neural network (GNN) was designed to predict ENSO. They employed a graph structure to represent various physical variables, particularly the coupling relationships between SST and clouds. This method can be understood as a feature fusion, improving the nonlinear combination of different variables/features to better identify critical processes influencing ENSO development. Another example comes from Chen et al. (2024) who designed a GNN for precipitation prediction considering the constraints of the omega equation and moisture equation. By using this physics-constrained network, the prediction of heavy rainfall was significantly improved, and the whole prediction results were also better than those of the other DL models without constraints. Meanwhile, a hard constraint usually modifies the neural network, ensuring an expected constraint during learning and inference (González-Abad et al., 2023; Harder et al., 2024). For instance, when downscaling T_{min} (minimum temperature) and T_{mean} (mean temperature) in González-Abad et al. (2023), the activation function for T_{min} was configured as ReLU, whereas the activation function for $T_{\mathrm{mean}}/T_{\mathrm{min}}$ was set as (1 + ReLU). This specific configuration guarantees that the results consistently adhere to the *a priori* relationship where T_{mean} is greater than T_{min}. Besides, there are also studies that have extended convolutional networks to accommodate a wider variety of symmetries, such as rotations, reflections, and more extensive gauge symmetry transformations (Bronstein et al., 2017; Cohen et al., 2019). This enhancement enables convolutional networks to more effectively recognize and process data patterns exhibiting these symmetries, helping models more accurately capture complex patterns and dynamics in atmospheric phenomena, thereby improving the precision and reliability of weather forecasting and climate modeling.

3. AI for numerical modeling: limitations and solutions

In the previous discussion, we emphasize the importance of dynamic diagnostics and physical constraints for AI climate models. It is essential to understand how AI can affect the field of numerical climate models. AI applications within numerical models are already underway, with a prominent example being the use of AI to construct parameterization schemes, enhancing the performance of numerical models (e.g., Rasp et al., 2018; Han et al., 2020; Zhu et al., 2022; Wang and Tan, 2023). The Fortran Torch Adaptor (Mu et al., 2023) is an adapter designed to bridge Fortran and Torch, enabling researchers to load Torch model parameters and perform forward propagation within a Fortran environment, which facilitates the coupling of AI models with numerical models.

However, challenges persist in current AI parameterization schemes. Firstly, most AI parameterization schemes are currently offline. These offline parameterization schemes, when coupled with models for long-term integration, may lead to instability, resulting in model crashes, and potentially give rise to various issues such as climate drift. Secondly, a DL optimization of a specific parameter in isolation for a given parameterization scheme aims to achieve the optimum value for that parameter. In reality, climate models have numerous parameters, and optimizing one parameter may only represent a local optimum for the entire model.

A potential solution to address these issues could involve the development of online parameterization schemes. Such a system could potentially compute losses by comparing numerical model output to observations and optimizing AI weights that are coupled to numerical models using back-propagation. Is it possible to develop this AI-driven parameterization scheme that is updated online to correct the results of the numerical model (Chen et al., 2023c)? A similar approach has been adopted in the field of hydrology, known as "differentiable models", which couple DL parameterization schemes with hydrological models, optimizing DL parameterizations based on hydrological model outputs (Shen et al., 2023). Even though attempts have been made to determine a few model parameters to reduce the difference between the observed and simulated results through the online parameter estimation (Han et al., 2014; Wu et al., 2016), implementing such an approach, especially within complex climate or coupled models, is still considered challenging. Recently, a groundbreaking development occurred with Google's neural general circulation model (NGCM; Kochkov et al., 2024). It represents an online AI–physics hybrid GCM, exhibiting superior performance compared to current mainstream models and demonstrating excellent adherence to physical conservation laws. This pioneering work has the potential to catalyze the application of similar paradigms in climate modeling. Despite this advancement, constructing comparable climate models remains challenging. While the NGCM employs an embedding approach to create AMIP-like experiments, there are substantial differences from climate simulations. In contrast to weather models, finding corresponding observations for climate model simulations is challenging due to the presence of internal variability, making it difficult to calculate errors for iterative optimization as demonstrated by the NGCM.

4. Infrastructure for the climate community

Earlier, we addressed the primary challenges faced by the integration of AI and climate studies, along with proposed potential solutions. We emphasize the crucial role of detailed dynamic diagnostics and physical constraints for AI models, along with the importance of developing online parameterization schemes for climate models. However, resolving these issues heavily relies on community culture and robust leadership. Currently, the climate-AI community lacks unified benchmarks and baselines, resulting in significant redundant efforts and making model comparisons exceptionally challenging. The principles of unified benchmarks and baselines are crucial for the advancement of the climate AI community. Wisdom can be gleaned from distant sources. Relevant experiences can be drawn from the computer science community to inform our approach.

The rapid development of the AI era is propelled by the explosion of computational power and the open nature of DL communities. For the continued advancement of climate models in the AI era, a similar condition is required, especially within the climate community, that is an open and collaborative infrastructure. It is essential to foster progress in the climate field, creating an environment where innovative models and ideas can be shared and developed collectively. We can summarize this requirement as the "Open, Comparable, Reproducible" (OCR) principle (Fig. 1).

- "Open" emphasizes open-source code and open datasets of AI climate modeling.
- "Comparable" signifies the creation of unified evaluation metrics. It involves integrating climate model evaluation frameworks to conduct thorough assessments involving climate dynamic assessments and the establishment of uniform datasets, task sets, and baselines.
- "Reproducible" denotes that the results published in research papers should be reproducible.

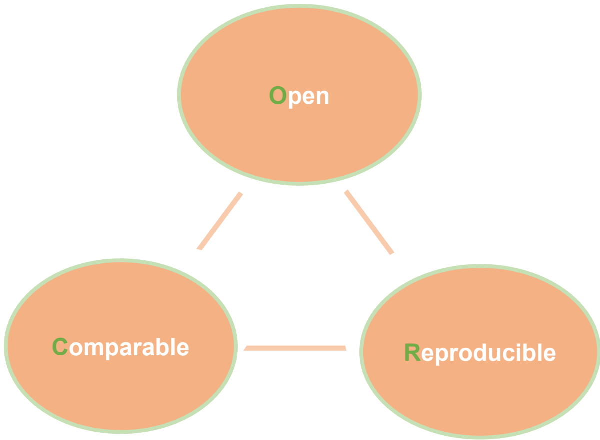

Fig. 1. The OCR principles.

The OCR principles have a nuanced interplay and mutual influence. "Open" serves as the foundational bedrock for the subsequent principles to materialize. "Comparable" and "Reproducible" complement each other symbiotically. The "Reproducible" part of AI model results is crucial for comparing different model outputs. This, in turn, increases scientists' awareness and emphasis on model reproducibility, creating a harmonious framework that propels scientific research toward more reliable and enduring progress.

The OCR principles hold significant potential for expediting the application of AI in climate science, particularly in cli-

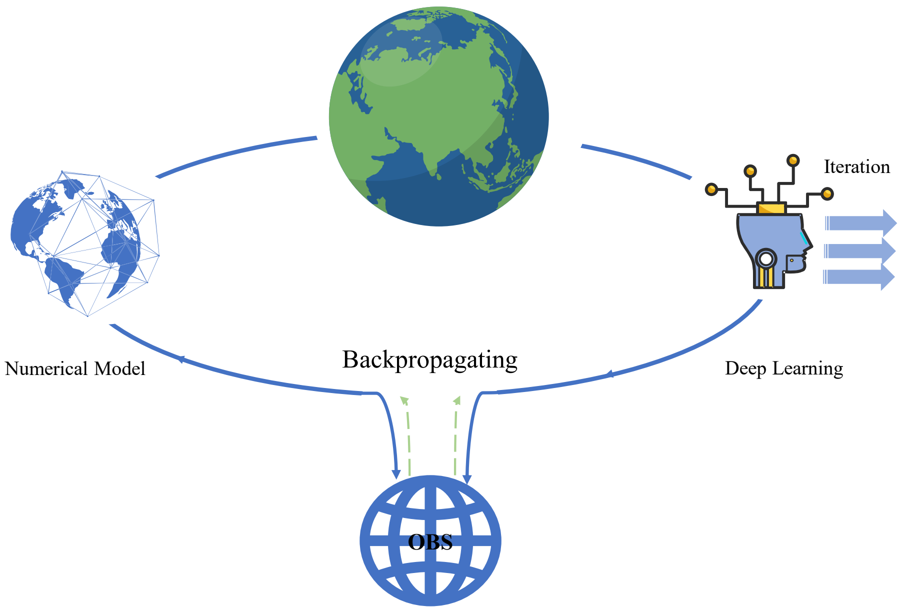

Fig. 2. Schematic of the balanced AI–physics model.

mate prediction and modeling. Regarding the latter, the success of CMIP has already demonstrated how an open-data comparison project can substantially advance the field. In accordance with the OCR principles, an AI-MIP can be established, which would greatly facilitate the evaluation, iteration, and optimization of AI models. Simultaneously, this initiative would promote the development of numerical modeling and AI–physics hybrid modeling.

5. Toward a balanced AI–physics climate model

Looking ahead, more comprehensive dynamic diagnostics and physical constraints can further enhance the performance of AI climate models. This progress will also stimulate the development of AI parameterization schemes. Considering the development of AI (especially unsupervised learning; physics-constrained AI) and numerical climate models, there is optimism that the challenge of discrepancies between climate model outputs and observations could be addressed. A balance between AI and numerical climate modeling may be attainable. This balance could involve using numerical methods to resolve solvable processes while employing online AI solutions to address critical parameterization schemes. As this hybrid approach advances, it becomes increasingly challenging to classify a model as purely numerical or AI-based. Ultimately, it evolves into a balanced AI–physics climate model (Fig. 2). Through the incorporation of AI-based parameterization schemes, such a model can be iteratively optimized, transforming into a learnable climate model. Furthermore, since it includes explicit computations of physical quantities and obeys physical equations, this model exhibits a high level of interpretability, guaranteeing direct climate diagnostic analyses.

Acknowledgements. This work was supported by the National Natural Science Foundation of China (Grant Nos. 42141019 and 42261144687) and STEP (Grant No. 2019QZKK0102). Yoo-Geun HAM was supported by the Korea Environmental Industry & Technology Institute (KEITI) through the "Project for developing an observation-based GHG emissions geospatial information map", funded by the Korea Ministry of Environment (MOE) (Grant No. RS-2023-00232066). Ya WANG would like to thank Prof. Baoxiang PAN for sharing related papers and useful discussions.

REFERENCES

Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525**(7567), 47−55, https://doi.org/10.1038/nature14956.

Beucler, T., M. Pritchard, P. Gentine, and S. Rasp, 2020: Towards physically-consistent, data-driven models of convection. *2020 IEEE International Geoscience and Remote Sensing Symposium*, Waikoloa, HI, USA, IEEE, 3987−3990, https://doi.org/10.1109/IGARSS39084.2020.9324569.

Beucler, T., M. Pritchard, S. Rasp, J. Ott, P. Baldi, and P. Gentine, 2021: Enforcing analytic constraints in neural networks emulating physical systems. *Physical Review Letters*, **126**(9), 098302, https://doi.org/10.1103/PhysRevLett.126.098302.

Bi, K. F., L. X. Xie, H. H. Zhang, X. Chen, X. T. Gu, and Q. Tian, 2023: Accurate medium-range global weather forecasting with 3D neural networks. *Nature*, **619**(7970), 533−538, https://doi.org/10.1038/s41586-023-06185-3.

Bonavita, M., 2023: On some limitations of data-driven weather forecasting models. arXiv preprint arXiv: 2309.08473, https://doi.org/10.48550/arXiv.2309.08473.

Bronstein, M. M., J. Bruna, Y. LeCun, A. Szlam, and P. Vandergheynst, 2017: Geometric deep learning: Going beyond Euclidean data. *IEEE Signal Processing Magazine*, **34**, 18−42, https://doi.org/10.1109/MSP.2017.2693418.

Chen, L., X. H. Zhong, F. Zhang, Y. Cheng, Y. H. Xu, Y. Qi, and H. Li, 2023a: FuXi: A cascade machine learning forecasting system for 15-day global weather forecast. arXiv preprint arXiv: 2306.12873, https://doi.org/10.48550/arXiv.2306.12873.

Chen, K. and Coauthors., 2023b: Fengwu: Pushing the skillful global medium-range weather forecast beyond 10 days lead. arXiv preprint arXiv:2304.02948, https://doi.org/10.48550/arXiv.2304.02948.

Chen, M., and Coauthors, 2023c: Iterative integration of deep learning in hybrid Earth surface system modelling. *Nature Reviews Earth & Environment*, **4**, 568−581, https://doi.org/10.1038/s43017-023-00452-7.

Chen, Y. T., Y. Wang, G. Huang, and Q. Tian, 2024: Coupling physical factors for precipitation forecast in China with graph neural network. *Geophys. Res. Lett.*, **51**, e2023GL106676, https://doi.org/10.1029/2023GL106676.

Cohen, T. S., M. Weiler, B. Kicanaoglu, and M. Welling, 2019: Gauge equivariant convolutional networks and the icosahedral CNN. *Proceedings of the 36th International Conference on Machine Learning*, https://doi.org/10.48550/arXiv.1902.04615.

Diffenbaugh, N. S., and E. A. Barnes, 2023: Data-driven predictions of the time remaining until critical global warming thresholds are reached. *Proceedings of the National Academy of Sciences of the United States of America*, **120**(6), e2207183120, https://doi.org/10.1073/pnas.2207183120.

Fan, P. Y., J. Yang, Z. P. Zhang, N. H. Zang, Y. F. Li, and G. L. Feng, 2023: Summer precipitation prediction in eastern China based on machine learning. *Climate Dyn.*, **60**(9−10), 2645−2663, https://doi.org/10.1007/s00382-022-06464-1.

Gill, A. E., 1980: Some simple solutions for heat-induced tropical circulation. *Quart. J. Roy. Meteor. Soc.*, **106**(449), 447−462, https://doi.org/10.1002/qj.49710644905.

González-Abad, J., Á. Hernández-García, P. Harder, D. Rolnick, and J. M. Gutiérrez, 2023: Multi-variable hard physical constraints for climate model downscaling. arXiv preprint arXiv: 2308.01868, https://doi.org/10.48550/arXiv.2308.01868.

Hakim, G. J., and S. Masanam, 2023: Dynamical tests of a deep-learning weather prediction model. arXiv preprint arXiv:2309.10867, https://doi.org/10.48550/arXiv.2309.10867.

Ham, Y. G., J. H. Kim, and J. J. Luo, 2019: Deep learning for multi-year ENSO forecasts. *Nature*, **573**(7775), 568−572, https://doi.org/10.1038/s41586-019-1559-7.

Ham, Y. G., J. H. Kim, E. S. Kim, and K. W. On, 2021: Unified deep learning model for El Niño/Southern Oscillation forecasts by incorporating seasonality in climate data. *Science Bulletin*, **66**(13), 1358–1366, https://doi.org/10.1016/j.scib.2021.03.009.

Ham, Y.-G., J.-H. Kim, S.-K. Min, D. Kim, T. M. Li, A. Timmermann, and M. F. Stuecker, 2023: Anthropogenic fingerprints in daily precipitation revealed by deep learning. *Nature*, **622**(7982), 301–307, https://doi.org/10.1038/s41586-023-06474-x.

Han, G. J., X. F. Zhang, S. Zhang, X. R. Wu, and Z. Liu, 2014: Mitigation of coupled model biases induced by dynamical core misfitting through parameter optimization: Simulation with a simple pycnocline prediction model. *Nonlinear Processes in Geophysics*, **21**, 357–366, https://doi.org/10.5194/npg-21-357-2014.

Han, Y. L., G. J. Zhang, X. M. Huang, and Y. Wang, 2020: A moist physics parameterization based on deep learning. *Journal of Advances in Modeling Earth Systems*, **12**(9), e2020MS002076, https://doi.org/10.1029/2020MS002076.

Harder, P., Q. D. Yang, V. Ramesh, P. Sattigeri, A. Hernandez-Garcia, C. Watson, D. Szwarcman, and D. Rolnick, 2024: Generating physically-consistent high-resolution climate data with hard-constrained neural networks. arXiv preprint arXiv:2208.05424, https://doi.org/10.48550/arXiv.2208.05424.

Held, I., 2019: 100 years of progress in understanding the general circulation of the atmosphere. *Meteor. Monogr.*, **59**, 6.1–6.23, https://doi.org/10.1175/AMSMONOGRAPHS-D-18-0017.1.

Held, I. M., M. F. Ting, and H. L. Wang, 2002: Northern winter stationary waves: Theory and modeling. *J. Climate*, **15**(16), 2125–2144, https://doi.org/10.1175/1520-0442(2002)015<2125:NWSWTA>2.0.CO;2.

Hess, P., M. Drüke, S. Petri, F. M. Strnad, and N. Boers, 2022: Physically constrained generative adversarial networks for improving precipitation fields from Earth system models. *Nature Machine Intelligence*, **4**(10), 828–839, https://doi.org/10.1038/s42256-022-00540-1.

Hoskins, B. J., and D. J. Karoly, 1981: The steady linear response of a spherical atmosphere to thermal and orographic forcing. *J. Atmos. Sci.*, **38**(6), 1179–1196, https://doi.org/10.1175/1520-0469(1981)038<1179:TSLROA>2.0.CO;2.

Jiang, W. P., P. Huang, G. Huang, and J. Ying, 2021: Origins of the excessive westward extension of ENSO SST simulated in CMIP5 and CMIP6 models. *J. Climate*, **34**(8), 2839–2851, https://doi.org/10.1175/JCLI-D-20-0551.1.

Jin, F. F., 1996: Tropical ocean-atmosphere interaction, the Pacific cold tongue, and the El Niño-Southern Oscillation. *Science*, **274**(5284), 76–78, https://doi.org/10.1126/science.274.5284.76.

Kim, H., Y. G. Ham, Y. S. Joo, and S. W. Son, 2021: Deep learning for bias correction of MJO prediction. *Nature Communications*, **12**, 3087, https://doi.org/10.1038/s41467-021-23406-3.

Kochkov, D., and Coauthors, 2024: Neural general circulation models for weather and climate. arXiv preprint arXiv:2311.07222, https://doi.org/10.48550/arXiv.2311.07222.

Labe, Z. M., and E. A. Barnes, 2021: Detecting climate signals using explainable AI with single-forcing large ensembles. *Journal of Advances in Modeling Earth Systems*, **13**(6), e2021MS002464, https://doi.org/10.1029/2021MS002464.

Labe, Z. M., and E. A. Barnes, 2022: Comparison of climate model large ensembles with observations in the arctic using simple neural networks. *Earth and Space Science*, **9**(7), e2022EA002348, https://doi.org/10.1029/2022EA002348.

Labe, Z. M., E. A. Barnes, and J. W. Hurrell, 2023: Identifying the regional emergence of climate patterns in the ARISE-SAI-1.5 simulations. *Environmental Research Letters*, **18**(4), 044031, https://doi.org/10.1088/1748-9326/acc81a.

Lam, R., and Coauthors, 2023: Learning skillful medium-range global weather forecasting. *Science*, **382**, 1416–1421, https://doi.org/10.1126/science.adi2336.

Li, G., and S. P. Xie, 2012: Origins of tropical-wide SST biases in CMIP multi-model ensembles. *Geophys. Res. Lett.*, **39**, L22703, https://doi.org/10.1029/2012GL053777.

Li, G., and S. P. Xie, 2014: Tropical biases in CMIP5 multimodel ensemble: The excessive equatorial pacific cold tongue and double ITCZ problems. *J. Climate*, **27**(4), 1765–1780, https://doi.org/10.1175/JCLI-D-13-00337.1.

Li, P. Y., Y. Yu, D. N. Huang, Z. H. Wang, and A. Sharma, 2023: Regional heatwave prediction using graph neural network and weather station data. *Geophys. Res. Lett.*, **50**(7), e2023GL103405, https://doi.org/10.1029/2023GL103405.

Ling, F. H., J. J. Luo, Y. Li, T. Tang, L. Bai, W. L. Ouyang, and T. Yamagata, 2022: Multi-task machine learning improves multi-seasonal prediction of the Indian Ocean Dipole. *Nature Communications*, **13**(1), 7681, https://doi.org/10.1038/s41467-022-35412-0.

Matsuno, T., 1966: Quasi-geostrophic motions in the equatorial area. *J. Meteor. Soc. Japan*, **44**(1), 25–43, https://doi.org/10.2151/jmsj1965.44.1_25.

McPhaden, M. J., S. E. Zebiak, and M. H. Glantz, 2006: ENSO as an integrating concept in earth science. *Science*, **314**(5806), 1740–1745, https://doi.org/10.1126/science.1132588.

Mohan, A. T., N. Lubbers, D. Livescu, and M. Chertkov, 2019: Embedding hard physical constraints in convolutional neural networks for 3D turbulence. *ICLR 2020 Workshop on Integration of Deep Neural Models and Differential Equations*.

Mu, B., B. Qin, and S. J. Yuan, 2021: ENSO-ASC 1.0.0: ENSO deep learning forecast model with a multivariate air-sea coupler. *Geoscientific Model Development*, **14**(11), 6977–6999, https://doi.org/10.5194/gmd-14-6977-2021.

Mu, B., L. Chen, S. J. Yuan, and B. Qin, 2023: A radiative transfer deep learning model coupled into WRF with a generic Fortran torch adaptor. *Frontiers in Earth Science*, **11**, 1149566, https://doi.org/10.3389/feart.2023.1149566.

Pan, B. X., G. J. Anderson, A. Goncalves, D. D. Lucas, C. J. W. Bonfils, and J. Lee, 2022: Improving seasonal forecast using probabilistic deep learning. mproving seasonal forecast using probabilistic deep learning. Journal of Advances in Modeling Earth Systems , 14, e2021MS002766. https://doi.org/10.1029/2021MS002766.

Pathak, Jaideep, and Coauthors, 2022: Fourcastnet: A global data-driven high-resolution weather model using adaptive Fourier neural operators. arXiv preprint arXiv:2202.11214 (2022), https://doi.org/10.48550/arXiv.2202.11214.

Raissi, M., P. Perdikaris, and G. E. Karniadakis, 2019: Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.*, **378**, 686–707, https://doi.org/10.1016/j.jcp.2018.10.045.

Rasp, S., M. S. Pritchard, and P. Gentine, 2018: Deep learning to represent subgrid processes in climate models. *Proceedings of the National Academy of Sciences of the United States of*

America, **115**(39), 9684−9689, https://doi.org/10.1073/pnas.1810286115.

Ravuri, S., and Coauthors, 2021: Skilful precipitation nowcasting using deep generative models of radar. *Nature*, **597**(7878), 672−677, https://doi.org/10.1038/s41586-021-03854-z.

Shen, C. P., and Coauthors, 2023: Differentiable modelling to unify machine learning and physical models for geosciences. *Nature Reviews Earth & Environment*, **4**(8), 552−567, https://doi.org/10.1038/s43017-023-00450-9.

Tao, W. C., G. Huang, R. G. Wu, K. M. Hu, P. F. Wang, and H. N. Gong, 2018: Origins of biases in CMIP5 models simulating northwest Pacific summertime atmospheric circulation anomalies during the decaying phase of ENSO. *J. Climate*, **31**(14), 5707−5729, https://doi.org/10.1175/JCLI-D-17-0289.1.

Tao, W. C., G. Huang, P. F. Wang, Y. Liu, G. H. Wen, and D. H. Dong, 2019: Dominant modes of CMIP3/5 models simulating northwest Pacific circulation anomalies during post-ENSO summer and their SST dependence. *Theor. Appl. Climatol.*, **138**(3−4), 1809−1820, https://doi.org/10.1007/s00704-019-02936-3.

Wallace, J. M., E. M. Rasmusson, T. P. Mitchell, V. E. Kousky, E. S. Sarachik, and H. von Storch, 1998: On the structure and evolution of ENSO-related climate variability in the tropical Pacific: Lessons from TOGA. *J. Geophys. Res.: Oceans*, **103**(C7), 14 241−14 259, https://doi.org/10.1029/97JC02905.

Wang, L.-Y., and Z.-M. Tan, 2023: Deep learning parameterization of the tropical cyclone boundary layer. *Journal of Advances in Modeling Earth Systems*, **15**, e2022MS003034, https://doi.org/10.1029/2022MS003034.

Wang, Y., and Coauthors, 2024: Correcting climate model sea surface temperature simulations with generative adversarial networks: Climatology, interannual variability, and extremes. *Adv. Atmos. Sci.*, https://doi.org/10.1007/s00376-024-3288-6. (in press)

Wu, X. R., G. J. Han, S. Q. Zhang, and Z. Y. Liu, 2016: A study of the impact of parameter optimization on ENSO predictability with an intermediate coupled model. *Climate Dyn.*, **46**, 711−727, https://doi.org/10.1007/s00382-015-2608-z.

Zhang, Y. C., M. S. Long, K. Y. Chen, L. X. Xing, R. H. Jin, M. I. Jordan, and J. M. Wang, 2023: Skilful nowcasting of extreme precipitation with NowcastNet. *Nature*, **619**(7970), 526−532, https://doi.org/10.1038/s41586-023-06184-4.

Zhou, L., and R. H. Zhang, 2023: A self-attention–based neural network for three-dimensional multivariate modeling and its skillful ENSO predictions. *Science Advances*, **9**(10), eadf2827, https://doi.org/10.1126/sciadv.adf2827.

Zhu, Y. C., R. H. Zhang, J. N. Moum, F. Wang, X. F. Li, and D. L. Li, 2022: Physics-informed deep-learning parameterization of ocean vertical mixing improves climate simulations. *National Science Review*, **9**(8), nwac044, https://doi.org/10.1093/nsr/nwac044.