State Planet, 5-opsr, 22, 2025
https://doi.org/10.5194/sp-5-opsr-22-2025
© Author(s) 2025. This work is distributed under
the Creative Commons Attribution 4.0 License.

Crafting the Future: Machine learning for ocean forecasting

Patrick Heimbach[1], Fearghal O'Donncha[2], Timothy A. Smith[3], Jose Maria Garcia-Valdecasas[4], Alain Arnaud[5], and Liying Wan[6]

[1]Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, TX, United States
[2]IBM Research, Dublin, Ireland
[3]NOAA Physical Sciences Laboratory, Boulder, CO, United States
[4]Nologin Oceanic Weather Systems, Santiago de Compostela, Spain
[5]Mercator Ocean International, Toulouse, France
[6]National Marine Environmental Forecasting Center Beijing, Beijing, China

Correspondence: Patrick Heimbach (heimbach@utexas.edu)

Received: 4 September 2024 – Discussion started: 20 September 2024
Revised: 9 December 2024 – Accepted: 18 December 2024 – Published: 2 June 2025

Abstract. Artificial intelligence and machine learning are accelerating research in Earth system science, with huge potential for impact and challenges in ocean prediction. Such algorithms are being deployed on different aspects of the forecasting workflow with the aim of improving its speed and skill. They include pattern classification and anomaly detection; regression and diagnostics; and state prediction from nowcasting to synoptic, sub-seasonal, and seasonal forecasting. This brief review emphasizes scientific machine learning methods that have the capacity to embed domain knowledge; to ensure interpretability through causal explanation, to be robust and reliable; to involve effectively high-dimensional statistical methods, supporting multi-scale and multi-physics simulations aimed at improving parameterization; and to drive intelligent automation, as well as decision support. An overview of recent numerical developments is discussed, highlighting the importance of fully data-driven ocean models for future expansion of ocean forecasting capabilities.

1 Introduction

Research into applications of artificial intelligence (AI) and machine learning (ML) in ocean, atmospheric, and climate sciences has accelerated at a breathtaking pace over the last 5 years or so (e.g., Schneider et al., 2023; Eyring et al., 2024). With essentially all of these applications concerned with ML, we will drop the more broadly defined "AI" term in most of the following, except when used by references cited. We will also take the perspective of scientific machine learning (SciML), defined in a 2019 U.S. Department of Energy report on "Basic Research Needs for Scientific Machine Learning" (Baker et al., 2019), which emphasizes six key elements of SciML algorithms: (i) ML approaches that incorporate domain knowledge, such as physical principles, symmetries, constraints, expert feedback, computational simulations, and formal uncertainties; (ii) ML approaches that are interpretable, such that a user's confidence in ML-based model predictions may be bolstered by causal explanations based on a user's domain knowledge; (iii) ML approaches that are robust and reliable as a prerequisite for making high-stakes and high-regret decisions; (iv) ML approaches that are data-intensive, i.e., that ingest high-dimensional, noisy, and uncertain input data which contain complex structures and which require statistical and probabilistic methods to deal with ill-conditioning, non-uniqueness, and over-fitting; (v) ML approaches that enhance modeling and simulation to support, e.g., multi-scale and multi-physics simulations in terms of improved model parameterization or model acceleration; and (vi) ML approaches to support intelligent automa-

tion and decision support, which can range from quality control to application-oriented post-processing workflows. Arguably, all of these criteria are fundamental to the uses of ML in ocean prediction.

Next, following the review by Reichstein et al. (2019), it is useful to distinguish different categories of ML applications, namely (A) classification and anomaly detection, which is concerned with, e.g., finding extreme event patterns or the classification of important structures or regimes; (B) regression, which is concerned with state reconstruction of important state variables, parameters, or diagnostics (metrics) from available data; and (C) state prediction, ranging from nowcasting to operational forecasting and sub-seasonal to seasonal prediction. A comprehensive collection of review articles on deep learning in Earth sciences is Camps-Valls et al. (2021), covering algorithmic foundations and examples of all three categories.

Because the subject of this document is ocean prediction, we will focus the following on the third category, namely state prediction or forecasting. To keep this review manageable, we will not review the interesting subjects of ML applications for state reconstruction, downscaling, or classification.

2 State prediction

The workflow of operational ocean prediction largely follows that of numerical weather prediction (NWP). Its core engine is a data assimilation (DA) framework, consisting of a physical model (i.e., a complex algorithm for solving a set of partial differential equations, PDEs), a workflow for quality-controlling and ingesting diverse observational data streams into the DA system (ideally in near-real time), and an optimal estimation algorithm that combines models and data in a formal manner that produces statistically optimal forecasts (e.g., Park and Zupanski, 2022). As pointed out by Stephen Penny in a 2022 U.S. National Academy of Sciences workshop on Machine Learning and Artificial Intelligence to Advance Earth System Science (NASEM, 2022), machine learning (ML) approaches hold the prospect for accelerating various elements of the DA workflow. We briefly summarize ML approaches targeting the physical model and the DA algorithm. Opportunities in the application of ML for partial differential equation (PDE)-based models fall into two main categories, where one is concerned with targeted insertion of ML within a physical model, and the other is concerned with the complete replacement of the physical model by a surrogate model. In the former, certain elements or subcomponents of a physical model are replaced by a surrogate model (e.g., a neural network), whereas in the latter, the entire model is emulated. Chantry et al. (2021) have used the terms "soft AI" versus "hard AI". We avoid the somewhat non-descriptive or ambiguous terminology to avoid giving a false sense of which of these approaches is "harder" to realize.

2.1 Hybrid physics–ML models: enhancing forecast models and data assimilation with ML algorithms

A major source of model uncertainty is the parameterization of subgrid-scale (SGS) processes in terms of structural errors (formulation of functional representations of parameterizations) and parametric uncertainties (calibrating empirical parameters in the functional representations). Exciting efforts are underway to apply machine learning to replace conventional functional representation subgrid-scale (SGS) turbulent oceanic processes with surrogate models that are based on machine learning and that have been trained either offline or online (Bolton and Zanna, 2019; Frezat et al., 2021, 2022; Zhang et al., 2023; Sane et al., 2023; Perezhogin et al., 2023). This follows on early ideas in the context of climate model parameterization (e.g., Schneider et al., 2017; Rasp et al., 2018). Similarly, equation discovery has proven successful to infer the functional form of such SGS ocean parameterization schemes (Zanna and Bolton, 2020, 2021; Perezhogin et al., 2024). A longer list of related efforts exists for numerical weather prediction and has been reviewed by Dueben et al. (2021) and Bouallègue et al. (2024). These surrogates, mostly some form of neural networks, have been trained on (i.e., fit to) what are considered simulations of much higher fidelity and where these processes are resolved (e.g., large-eddy simulations). Related efforts aim at learning improved parameterizations from online bias correction or analysis increments incurred in sequential data assimilation (e.g., Gregory et al., 2023, 2024; Storto et al., 2024). Rapid progress is expected on this front in the coming years.

A second important application of hybrid approaches is the desire to replace specific numerical algorithms within PDE-based models with surrogate models to accelerate the simulation's time to solution. Studies exist within the generic field of computational fluid dynamics (Kochkov et al., 2021) and atmospheric modeling (Arcomano et al., 2023; Kochkov et al., 2024), and there are ocean-specific applications currently underway. Most of these take advantage of the concept of differentiable programming (Gelbrecht et al., 2023; Shen et al., 2023; Zhang et al., 2023; Sapienza et al., 2024). The underlying idea is to eventually be able to generate code for the derivative of the physical model, in particular the adjoint model that enables efficient "online" (or "full model") learning of the model parameters (or neural network weights).

There is a strong conceptual correspondence between machine learning and data assimilation (e.g., Abarbanel et al., 2018). This provides various opportunities for embedding ML approaches within operational data assimilation workflows deployed in ocean prediction. Examples in ocean modeling so far are largely restricted to "toy problems" (such as the "Lorenz 96 model") or reduced-order versions of Earth system models that target eventual applications for

ocean prediction (Bocquet et al., 2020; Brajard et al., 2021; Penny et al., 2022; Irrgang et al., 2021). The use of hybrid DA/ML approaches, whether in the context of ensemble DA or adjoint-based methods (e.g., 4DVar), presents substantial algorithmic hurdles (e.g., availability of a differentiable dynamical core in the context of adjoint-based DA), which explains the relative paucity of such studies to date compared to purely data-driven methods.

2.2 Purely data-driven models: replacing numerical simulations with surrogate models

Over the last decade, with the acceleration in AI-based solutions in other fields, a number of approaches to model the atmosphere and ocean using different purely data-driven ML techniques have been developed. The overwhelming majority of these cases have so far been realized in weather prediction or computational fluid dynamics.

2.2.1 Deterministic applications in weather prediction

Arguably, the field of data-driven weather forecasting has seen the strongest advances over the last 5 years or so (Schneider et al., 2022). This is a strong incentive for providing a very brief review that is organized in terms of approaches as a function of underlying "blocks" of the ML architectures employed. In a number of cases, these architectural blocks are being combined. For example, the European Centre for Medium-Range Weather Forecast's AIFS system (Lang et al., 2024) uses an overall "encode–process–decode" architecture, with a graph-based encoder and decoder but a sliding window transformer as the processor.

- *Convolutional neural networks (CNNs)*. Perhaps among the first serious endeavors using ML for emulating weather forecast models have been the CNNs used by Weyn et al. (2019, 2020, 2021) and Karlbauer et al. (2024). CNNs use a mathematical operation called convolution to compress information, learning features, or patterns in the input. Most recently, CNNs have been used by Cresswell-Clay et al. (2024) to create a coupled atmosphere–ocean emulator which produces a stable climate for 1000-year periods and appears to be competitive with many CMIP6 models.
- *Graph neural networks*. Among the leading emulators for medium-range weather forecasts is the work by Lam et al. (2023). Based on graph neural networks, the GraphCast model was trained on atmospheric reanalysis data to produce autoregressive forecasts for up to 10 d.
- *Transformers*. These have been revolutionary in other ML/AI fields, such as natural language processing and image recognition/generation. They serve as the backbone of some of the leading atmospheric emulators, including Pangu-Weather (Bi et al., 2023), FuXi (L. Chen et al., 2023), and FengWu (K. Chen et al., 2023).
- *Fourier neural operators (FNOs)*. FNOs have been designed to move toward mesh-independent operators using Fourier bases (Li et al., 2020). FourCastNet (Pathak et al., 2022; Kurth et al., 2023) is based on a variant called the Adaptive FNO (AFNO). Another variant, the Spherical FNO (SFNO; Bonev et al., 2023; Watt-Meyer et al., 2023) seeks to take advantage of the spherical geometry (and underlying symmetries) in representing operator kernels for global-scale applications. Very recently, the use of SFNOs has been extended to coupled atmosphere–ocean modeling targeting seasonal prediction (C. Wang et al., 2024).
- *Recurrent neural networks (including long short-term memory, LSTM, and reservoir computing)*. Recurrent neural networks (RNNs) are well suited for sequential data processing, such as time series. Among special cases of RNNs, LSTM networks use a special type of neuron that keeps track of previous inputs (short-term memory) and are especially useful for predicting time series with memory, such as the case for the atmosphere and ocean. Reservoir computing (RC), another method based on RNNs with a pool of interconnected neurons forming the "reservoir", is particularly well adapted to the emulation of time series (e.g., Arcomano et al., 2020; Penny et al., 2022; Platt et al., 2023; Smith et al., 2023).

2.2.2 Probabilistic approaches – generative models

Most examples sketched in Sect. 2.2.1 describe emulators that are trained to be deterministic forecast models. Recent developments in ML have considered generative frameworks, i.e., models that are designed to be probabilistic. Such frameworks would include variational autoencoders, generative adversarial networks (GANs), and diffusion models. However, we note that GANs can suffer from a lack of sample diversity (Bayat, 2023), and they are notoriously challenging to train, requiring a careful setup to avoid training instabilities (e.g., Miyato et al., 2018). Moreover, in recent years, diffusion models have started to outperform GANs in image classification (Dhariwal and Nichol, 2021). For these reasons, diffusion models have become popular in generative modeling, despite their relatively high computational cost. Recent examples of diffusion models include GenCast (Price et al., 2024). Finally, we note a very recently developed technique, DYffusion (Cachay et al., 2023, 2024), which is a generative framework that aims to reduce the computational cost of diffusion modeling by encoding the temporal evolution expected in physical systems into the generative process.

2.2.3 Physics-informed machine learning

The results of purely data-driven solutions may potentially produce meaningless output as the training strategy of a neural network is to minimize a mathematical loss function, e.g.,

the mean squared error (i.e., L2 norm) between the prediction and the original target. Similar issues, e.g., producing overly blurred output, may arise with other choices of the loss function, such as an L1 norm. An evolution of this approach is to include some physical constraints in the loss function in order to force the ML algorithm to produce more consistent outputs, such as the Navier–Stokes equation (Ma et al., 2022; Daw et al., 2021). This class of methods is known as physics-informed neural networks (PINNs). However, the performance of PINNs for extrapolation remains subject to debate (e.g., Du et al., 2023, for a cautionary example). Recently, another approach, which tries to solve differential equations using neural networks, is under development. Although this method is mostly developed for other physics fields, the methodology and knowledge can be applied to ocean modeling (Zubov et al., 2021; Smets et al., 2023).

2.2.4 Applications in ocean surface state forecasting

With previous examples mostly limited to weather prediction and computational fluid dynamics (in a few cases), we turn our attention to applications in the context of predicting ocean surface properties. They include the use of multilayer perceptrons (James et al., 2018; Gracia et al., 2021) and LSTMs (Minuzzi and Farina, 2023; Lawal et al., 2024) for surface wave prediction, surface wave–current interaction forecasting, storm surge forecasting (Xie et al., 2023), and sea surface temperature prediction via deep learning (Wolff et al., 2020; Xu et al., 2023) and the use of neural networks for accelerating resonant nonlinear wave–wave interaction in an ocean surface wave model (Puscasu, 2014), regional to coastal sea level prediction (Nieves et al., 2021), ocean color mapping (Chen et al., 2019), and statistical downscaling (Accarino et al., 2021). Other applications include estimating ocean surface circulation (Sinha and Abernathey, 2021; Subel and Zanna, 2024) and predicting dissolved oxygen across scales (O'Donncha et al., 2022).

2.3 ML-based ocean circulation prediction

Among the challenges of fully realizing the opportunities of ML approaches in ocean circulation prediction is the fact that, in the absence of adequate and densely sampled observational data, most ML applications rely on the use of data obtained from high-fidelity model simulations as training data sets. These data sets are very expensive to generate, limited in the temporal ranges that they can represent, remain subject to unquantified structural and parametric model uncertainty, require vast amounts of storage (on the order of petabytes), and are thus challenging to query. Cloud-based solutions are the most promising approach for ubiquitous data access and analysis capabilities "close to the data" (Abernathey et al., 2020).

Within the realm of machine learning (ML) applications for ocean forecasting, progress has been somewhat limited. Recent developments have marked a shift in this landscape, particularly with the introduction of Fourier neural operators for modeling oceanic processes, as suggested by Bire et al. (2023), Chattopadhyay et al. (2024), and Sun et al. (2024). These studies present fully data-driven ocean models that match the capabilities of traditional numerical ocean models in predicting high-resolution sea surface height (SSH) fields. FNOs are attractive for their performance in learning complex and high-dimensional mappings and their ability to incorporate physical laws and constraints, which are prominently observable in the spectral domain. A drawback of FNOs applied to ocean (unlike atmospheric) modeling is the existence of land-covered portions of the domain, which renders the use of periodic basis functions challenging and may create artifacts near land–ocean boundaries.

Concurrently, X. Wang et al. (2024) introduced a transformer-based model tailored for oceanic applications, demonstrating performance that rivals that of leading operational global ocean forecasting systems. Similar advances are being made in the data-driven prediction of sea ice cover in the polar oceans (Andersson et al., 2021; see also Bertino et al., 2025, in this report). This body of work signifies the emergence of a promising research avenue in fully data-driven ocean modeling, despite it still lagging considerably behind the advancements seen in weather forecasting. We posit that the drive of fully data-driven solutions in NWP by private sector companies is related to the prospect of high-stakes and high-reward applications. Such applications for ocean predictions should be better articulated to attract similar research efforts. Careful evaluation of skill, such as that now being discussed more comprehensively in NWP (e.g., Bonavita, 2023; Charlton-Perez et al., 2024), will also be required for operational ocean prediction.

Another challenge presents the extension of ML applications to seasonal, inter-annual, and multi-decadal – i.e., climate – timescales (see, e.g., the discussion in Gentine et al., 2021; Beucler et al., 2024; Subel and Zanna, 2024). Here, the increased need for models or invariant operators (physics-based or surrogates) to conserve fundamental properties (mass, energy, momentum, and active tracers) puts severe demands on ML approaches. Arguably, as these approaches increasingly incorporate physical knowledge, they will converge to the realm of classical inverse methods (Willcox et al., 2021).

2.4 Benchmarking forecast models

Data-driven forecasting in meteorology – and to some extent in oceanography – is proceeding at a breathtaking pace. The use of different approaches, different training data, and different performance metrics complicates objective assessment of the different works at the present time. Recognizing the need for standardized evaluation has led to the proposition of common evaluation benchmarks that encompass both data-driven and "traditional" forecasting in weather predic-

tion (Dueben et al., 2022; Rasp et al., 2020, 2024), as well as climate model emulation (Yu et al., 2023). These benchmarks comprise common data sets, open-source evaluation workflows, and common evaluation metrics. Similar benchmarking efforts in ML-driven ocean circulation and surface wave forecasting will be equally important to advance the field and establish standardized evaluation metrics.

3 The role of surrogate models in digital twins

The concept of digital twins (DTs) is rapidly gaining traction within the ocean science community and Earth system science more broadly (e.g., Bauer et al., 2021a, b). Because of the differing view of what constitutes a DT in the recent literature, we here adopt and emphasize the definition from NASEM (2022) (see also Niederer et al., 2021; NASEM, 2023), which states that a DT is

> a set of virtual information constructs that mimics the structure, context and behavior of an individual/unique physical asset, or a group of physical assets, is dynamically updated with data from its physical twin throughout its life cycle and informs decisions that realize value. A digital twin is highly dynamical, mimicking the time evolution of its physical asset (PA) via advanced simulation and emulation capabilities; it is updated by ingesting vast amounts of observational data of diverse types; and it enables WHAT-IF queries and multiple realizations to support prediction of responses of the PA to hypothetical perturbations with quantified uncertainties.

Virtually all aspects of ocean forecasting – and ML opportunities therein – may be viewed through the DT lens from the need to generate high-fidelity simulations or digital representations, ingesting, i.e., assimilating, large and heterogeneous data streams, and the development of fast surrogates or emulators to either accelerate simulations or provide comprehensive uncertainty estimates, to the generation of diagnostic data that create value for (possibly rapid) decision support.

Data availability. No data sets were used in this article.

Author contributions. PH and FO'D conceptualized the paper. The original draft was prepared by PH, FO'D, and JMGV. All authors contributed to reviewing and editing the paper.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Acknowledgements. We would like to thank Enrique Alvarez Fanjul at the OceanPrediction Decade Collaborative Center for initiating this review.

Financial support. The authors have been supported in part by the U.S. National Science Foundation (grant no. 2103942), the U.S. Office of Naval Research (grant no. N00014-20-1-2772), the U.S. Department of Energy (grant no. DE-SC002317), the NOAA Physical Sciences Laboratory, and a JPL/Caltech subcontract.

Review statement. This paper was edited by Swadhin Behera and reviewed by two anonymous referees.

References

Abarbanel, H. D. I., Rozdeba, P. J., and Shirman, S.: Machine Learning: Deepest Learning as Statistical Data Assimilation Problems, Neural Comput., 30, 2025–2055, https://doi.org/10.1162/neco_a_01094, 2018.

Abernathey, R. P., Augspurger, T., Banihirwe, A., Blackmon-Luca, C. C., Crone, T. J., Gentemann, C. L., Hamman, J. J., Henderson, N., Lepore, C., McCaie, T. A., Robinson, N. H., and Signell, R. P.: Cloud-Native Repositories for Big Scientific Data, Comput. Sci. Eng., 23, 26–35, https://doi.org/10.1109/mcse.2021.3059437, 2020.

Accarino, G., Chiarelli, M., Immorlano, F., Aloisi, V., Gatto, A., and Aloisio, G.: MSG-GAN-SD: A Multi-Scale Gradients GAN for Statistical Downscaling of 2-Meter Temperature over the EURO-CORDEX Domain, AI, 2, 600–620, https://doi.org/10.3390/ai2040036, 2021.

Andersson, T. R., Hosking, J. S., Pérez-Ortiz, M., Paige, B., Elliott, A., Russell, C., Law, S., Jones, D. C., Wilkinson, J., Phillips, T., Byrne, J., Tietsche, S., Sarojini, B. B., Blanchard-Wrigglesworth, E., Aksenov, Y., Downie, R., and Shuckburgh, E.: Seasonal Arctic sea ice forecasting with probabilistic deep learning, Nat. Commun., 12, 5124, https://doi.org/10.1038/s41467-021-25257-4, 2021.

Arcomano, T., Szunyogh, I., Pathak, J., Wikner, A., Hunt, B. R., and Ott, E.: A Machine Learning-Based Global Atmospheric Forecast Model, Geophys. Res. Lett., 47, e2020GL087776, https://doi.org/10.1029/2020GL087776, 2020.

Arcomano, T., Szunyogh, I., Wikner, A., Hunt, B. R., and Ott, E.: A Hybrid Atmospheric Model Incorporating Machine Learning Can Capture Dynamical Processes Not Captured by Its Physics-Based Component, Geophys. Res. Lett., 50, e2022GL102649, https://doi.org/10.1029/2022GL102649, 2023.

Baker, N., Alexander, F., Bremer, T., Hagberg, A., Kevrekidis, Y., Najm, H., Parashar, M., Patra, A., Sethian, J., Wild, S., and Will-

cox, K.: U.S. Department of Energy Report on Basic Research Needs for Scientific Machine Learning: Core Technologies for Artificial Intelligence, 1–109, https://doi.org/10.2172/1478744, 2019.

Bauer, P., Stevens, B., and Hazeleger, W.: A digital twin of Earth for the green transition, Nat. Clim. Change, 5, 80–83, https://doi.org/10.1038/s41558-021-00986-y, 2021a.

Bauer, P., Dueben, P. D., Hoefler, T., Quintino, T., Schulthess, T. C., and Wedi, N. P.: The digital revolution of Earth-system science, Nature Computational Science, 1, 104–113, https://doi.org/10.1038/s43588-021-00023-0, 2021b.

Bayat, R.: A Study on Sample Diversity in Generative Models: GANs vs. Diffusion Models, The First Tiny Papers Track at ICLR 2023, Tiny Papers @ ICLR 2023, Kigali, Rwanda, May 5, https://openreview.net/forum?id=BQpCuJoMykZ (last access: 15 February 2025), 2023.

Bertino, L., Heimbach, P., Blockley, E., and Ólason, E.: Numerical Models for Monitoring and Forecasting Sea Ice: a short description of present status, in: Ocean prediction: present status and state of the art (OPSR), edited by: Álvarez Fanjul, E., Ciliberti, S. A., Pearlman, J., Wilmer-Becker, K., and Behera, S., Copernicus Publications, State Planet, 5-opsr, 14, https://doi.org/10.5194/sp-5-opsr-14-2025, 2025.

Beucler, T., Gentine, P., Yuval, J., Gupta, A., Peng, L., Lin, J., Yu, S., Rasp, S., Ahmed, F., O'Gorman, P. A., Neelin, J. D., Lutsko, N. J., and Pritchard, M.: Climate-invariant machine learning, Science Advances, 10, eadj7250, https://doi.org/10.1126/sciadv.adj7250, 2024.

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

Bire, S., Lütjens, B., Azizzadenesheli, K., Anandkumar, A., and Hill, C. N.: Ocean Emulation with Fourier Neural Operators: Double Gyre, ESS Open Archive [preprint], https://doi.org/10.22541/essoar.170110658.85641696/v1, 2023.

Bocquet, M., Brajard, J., Carrassi, A., and Bertino, L.: Bayesian inference of chaotic dynamics by merging data assimilation, machine learning and expectation-maximization. Foundations of Data Science, 2, 55–80, https://doi.org/10.3934/fods.2020004, 2020.

Bolton, T. and Zanna, L.: Applications of Deep Learning to Ocean Data Inference and Subgrid Parameterization, J. Adv. Model. Earth Sy., 11, 376–399, https://doi.org/10.1029/2018ms001472, 2019.

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

Bonev, B., Kurth, T., Hundt, C., Pathak, J., Baust, M., Kashinath, K., and Anandkumar, A.: Spherical Fourier Neural Operators: Learning Stable Dynamics on the Sphere, in: Proceedings of the 40th International Conference on Machine Learning, Honolulu, HI, USA, 23–29 July 2023, https://proceedings.mlr.press/v202/bonev23a.html (last access: 15 February 2025), 2023.

Bouallègue, Z. B., Weyn, J. A., Clare, M. C. A., Dramsch, J., Dueben, P., and Chantry, M.: Improving Medium-Range Ensemble Weather Forecasts with Hierarchical Ensemble Transformers, Artificial Intelligence for the Earth Systems, 3, e230027, https://doi.org/10.1175/aies-d-23-0027.1, 2024.

Brajard, J., Carrassi, A., Bocquet, M., and Bertino, L.: Combining data assimilation and machine learning to infer unresolved scale parametrization, Philos. T. R. Soc. A, 379, 20200086, https://doi.org/10.1098/rsta.2020.0086, 2021.

Cachay, S. R., Zhao, B., James, H., and Yu, R.: DYkusion: A Dynamicsinformed Dikusion Model for Spatiotemporal Forecasting, in: Advances in Neural Information Processing Systems 36 (NeurIPS 2023), Long Beach, California, USA, 4–9 December 2017. https://proceedings.neurips.cc/paper_files/paper/2023/hash/8df90a1440ce782d1f5607b7a38f2531-Abstract-Conference.html (last access: 15 February 2025), 2023.

Cachay, S. R., Henn, B., Watt-Meyer, O., Bretherton, C. S., and Yu, R.: Probabilistic Emulation of a Global Climate Model with Spherical DYffusion, arXiv [preprint], https://doi.org/10.48550/arxiv.2406.14798, 2024.

Camps-Valls, G., Tuia, D., Zhu, X. X., and Reichstein, M. (Eds.): Deep Learning for the Earth Sciences: A Comprehensive Approach to Remote Sensing, Climate Science, and Geosciences, John Wiley & Sons Ltd, https://doi.org/10.1002/9781119646181, 2021.

Chantry, M., Christensen, H., Dueben, P., and Palmer, T.: Opportunities and challenges for machine learning in weather and climate modelling: hard, medium and soft AI, Philos. T. R. Soc. A, 379, 20200083, https://doi.org/10.1098/rsta.2020.0083, 2021.

Charlton-Perez, A. J., Dacre, H. F., Driscoll, S., Gray, S. L., Harvey, B., Harvey, N. J., Hunt, K. M. R., Lee, R. W., Swaminathan, R., Vandaele, R., and Volonté, A.: Do AI models produce better weather forecasts than physics-based models? A quantitative evaluation case study of Storm Ciarán, Npj Climate and Atmospheric Science, 7, 93, https://doi.org/10.1038/s41612-024-00638-w, 2024.

Chattopadhyay, A., Gray, M., Wu, T., Lowe, A. B., and He, R.: OceanNet: a principled neural operator-based digital twin for regional oceans, Scientific Reports, 14, 21181, https://doi.org/10.1038/s41598-024-72145-0, 2024.

Chen, K., Han, T., Gong, J., Bai, L., Ling, F., Luo, J.-J., Chen, X., Ma, L., Zhang, T., Su, R., Ci, Y., Li, B., Yang, X., and Ouyang, W.: FengWu: Pushing the Skillful Global Medium-range Weather Forecast beyond 10 Days Lead, arXiv [preprint], https://doi.org/10.48550/arxiv.2304.02948, 2023.

Chen, L., Zhong, X., Zhang, F., Cheng, Y., Xu, Y., Qi, Y., & Li, H.: FuXi: a cascade machine learning forecasting system for 15-day global weather forecast, Npj Climate and Atmospheric Science, 6, 190, https://doi.org/10.1038/s41612-023-00512-1, 2023.

Chen, S., Hu, C., Barnes, B. B., Xie, Y., Lin, G., and Qiu, Z.: Improving ocean color data coverage through machine learning, Remote Sens. Environ., 222, 286–302, https://doi.org/10.1016/j.rse.2018.12.023, 2019.

Cresswell-Clay, N., Liu, B., Durran, D., Liu, A., Espinosa, Z. I., Moreno, R., abd Karlbauer, M.: A Deep Learning Earth System Model for Stable and Efficient Simulation of the Current Climate, arXiv [preprint], https://doi.org/10.48550/arXiv.2409.16247, 2024.

Daw, A., Maruf, M., and Karpatne, A.: PID-GAN: A GAN Framework based on a Physics-informed Discriminator for Uncertainty Quantification with Physics, in: KDD '21: Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery &

Data Mining, New York, NY, USA, 14–18 August 2021, 237–247, https://doi.org/10.1145/3447548.3467449, 2021.

Dhariwal, P. and Nichol, A.: Diffusion Models Beat GANs on Image Synthesis, Adv. Neur. In., 34, 8780–8794, https://proceedings.nips.cc/paper/2021/hash/49ad23d1ec9fa4bd8d77d02681df5cfa-Abstract.html (last access: 15 February 2025), 2021.

Du, Y., Wang, M., and Zaki, T. A.: State estimation in minimal turbulent channel flow: A comparative study of 4DVar and PINN. International Journal of Heat and Fluid Flow, 99, 109073, https://doi.org/10.1016/j.ijheatfluidflow.2022.109073, 2023.

Dueben, P. D., Bauer, P., and Adams, S.: Deep Learning to Improve Weather Predictions, in: In: Deep Learning for the Earth Sciences: A Comprehensive Approach to Remote Sensing, Climate Science, and Geosciences, edited by: Camps-Valls, G., Tuia, D., Zhu X. X., and Reichstein, M., Wiley & Sons, 204–217, https://doi.org/10.1002/9781119646181.ch14, 2021.

Dueben, P. D., Schultz, M. G., Chantry, M., Gagne, D. J., Hall, D. M., and McGovern, A.: Challenges and benchmark datasets for machine learning in the atmospheric sciences: Definition, status and outlook, Artif. Intell. Earth Syst., 1, e210002, https://doi.org/10.1175/aies-d-21-0002.1, 2022.

Eyring, V., Collins, W. D., Gentine, P., Barnes, E. A., Barreiro, M., Beucler, T., Bocquet, M., Bretherton, C. S., Christensen, H. M., Dagon, K., Gagne, D. J., Hall, D., Hammerling, D., Hoyer, S., Iglesias-Suarez, F., Lopez-Gomez, I., McGraw, M. C., Meehl, G. A., Molina, M. J., Monteleoni, C., Mueller, J., Pritchard, M. S., Rolnick, D., Runge, J., Stier, P., Watt-Meyer, O., Weigel, K., Yu, R., and Zanna, L.: Pushing the frontiers in climate modelling and analysis with machine learning, Nat. Clim. Change, 14, 916–928, https://doi.org/10.1038/s41558-024-02095-y, 2024.

Frezat, H., Balarac, G., Sommer, J. L., Fablet, R., and Lguensat, R.: Physical invariance in neural networks for subgrid-scale scalar flux modeling, Physical Review Fluids, 6, 024607, https://doi.org/10.1103/physrevfluids.6.024607, 2021.

Frezat, H., Sommer, J. L., Fablet, R., Balarac, G., and Lguensat, R.: A Posteriori Learning for Quasi-Geostrophic Turbulence Parametrization, J. Adv. Model. Earth Sy., 14, e2022MS003124, https://doi.org/10.1029/2022ms003124, 2022.

Gelbrecht, M., White, A., Bathiany, S., and Boers, N.: Differentiable programming for Earth system modeling, Geosci. Model Dev., 16, 3123–3135, https://doi.org/10.5194/gmd-16-3123-2023, 2023.

Gentine, P., Eyring, V., and Beucler, T.: Deep Learning for the Parametrization of Subgrid Processes in Climate Models, in: In: Deep Learning for the Earth Sciences: A Comprehensive Approach to Remote Sensing, Climate Science, and Geosciences, edited by: Camps-Valls, G., Tuia, D., Zhu, X. X., and Reichstein, M., Wiley & Sons, 307–314, https://doi.org/10.1002/9781119646181.ch21, 2021.

Gracia, S., Olivito, J., Resano, J., Martin-del-Brio, B., de Alfonso, M., and Álvarez, E.: Improving accuracy on wave height estimation through machine learning techniques, Ocean Eng., 236, 108699, https://doi.org/10.1016/j.oceaneng.2021.108699, 2021.

Gregory, W., Bushuk, M., Adcroft, A., Zhang, Y., and Zanna, L.: Deep Learning of Systematic Sea Ice Model Errors From Data Assimilation Increments, J. Adv. Model. Earth Sy., 15, e2023MS003757, https://doi.org/10.1029/2023ms003757, 2023.

Gregory, W., Bushuk, M., Zhang, Y., Adcroft, A., and Zanna, L.: Machine Learning for Online Sea Ice Bias Correction Within Global Ice-Ocean Simulations, Geophys. Res. Lett., 51, e2023GL106776, https://doi.org/10.1029/2023gl106776, 2024.

Irrgang, C., Boers, N., Sonnewald, M., Barnes, E. A., Kadow, C., Staneva, J., and Saynisch-Wagner, J.: Towards neural Earth system modelling by integrating artificial intelligence in Earth system science, Nature Machine Intelligence, 3, 667–674, https://doi.org/10.1038/s42256-021-00374-3, 2021.

James, S. C., Zhang, Y., and O'Donncha, F.: A machine learning framework to forecast wave conditions, Coast. Eng., 137, 1–10, https://doi.org/10.1016/j.coastaleng.2018.03.004, 2018.

Karlbauer, M., Cresswell-Clay, N., Durran, D. R., Moreno, R. A., Kurth, T., Bonev, B., Brenowitz, N., and Butz, M. V.: Advancing Parsimonious Deep Learning Weather Prediction Using the HEALPix Mesh, J. Adv. Model. Earth Sy., 16, e2023MS004021, https://doi.org/10.1029/2023ms004021, 2024.

Kochkov, D., Smith, J. A., Alieva, A., Wang, Q., Brenner, M. P., and Hoyer, S.: Machine learning–accelerated computational fluid dynamics, P. Natl. Acad. Sci. USA, 118, e2101784118, https://doi.org/10.1073/pnas.2101784118, 2021.

Kochkov, D., Yuval, J., Langmore, I., Norgaard, P., Smith, J., Mooers, G., Klöwer, M., Lottes, J., Rasp, S., Düben, P., Hatfield, S., Battaglia, P., Sanchez-Gonzalez, A., Willson, M., Brenner, M. P., and Hoyer, S.: Neural general circulation models for weather and climate, Nature, 632, 1060–1066, https://doi.org/10.1038/s41586-024-07744-y, 2024.

Kurth, T., Subramanian, S., Harrington, P., Pathak, J., Mardani, M., Hall, D., Miele, A., Kashinath, K., and Anandkumar, A.: FourCastNet: Accelerating Global High-Resolution Weather Forecasting Using Adaptive Fourier Neural Operators, in: PASC '23: Proceedings of the Platform for Advanced Scientific Computing Conference, Davos, Switzerland, 26–28 June 2023, https://doi.org/10.1145/3592979.3593412, 2023.

Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Pritzel, A., Ravuri, S., Ewalds, T., Alet, F., Eaton-Rosen, Z., Hu, W., Merose, A., Hoyer, S., Holland, G., Stott, J., Vinyals, O., Mohamed, S., and Battaglia, P.: GraphCast: Learning skillful medium-range global weather forecasting, Science, 382, 1416–1421, https://doi.org/10.1126/science.adi2336, 2023.

Lang, S., Alexe, M., Chantry, M., Dramsch, J., Pinault, F., Raoult, B., Clare, M. C. A., Lessig, C., Maier-Gerber, M., Magnusson, L., Bouallègue, Z. B., Nemesio, A. P., Dueben, P. D., Brown, A., Pappenberger, F., and Rabier, F.: AIFS – ECMWF's data-driven forecasting system, arXiv [preprint], https://doi.org/10.48550/arXiv.2406.01465, 2024.

Lawal, Z. K., Yassin, H., Teck Ching Lai, D., and Che Idris, A.: Understanding the Dynamics of Ocean Wave-Current Interactions Through Multivariate Multi-Step Time Series Forecasting, Appl. Artif. Intell., 38, https://doi.org/10.1080/08839514.2024.2393978, 2024.

Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A., and Anandkumar, A.: Fourier Neural Operator for Parametric Partial Differential Equations, arXiv [preprint], https://doi.org/10.48550/arXiv.2010.08895, 2020.

Ma, H., Zhang, Y., Thuerey, N., X. H., and Haidn, O. J.: Physics-Driven Learning of the Steady Navier-Stokes Equations using Deep Convolutional Neural Networks, Commun. Comput. Phys., 32, 715–736, https://doi.org/10.4208/cicp.OA-2021-0146, 2022.

Minuzzi, F. C. and Farina, L.: A deep learning approach to predict significant wave height using long short-term memory, Ocean Model., 181, 102151, https://doi.org/10.1016/j.ocemod.2022.102151, 2023.

Miyato, T., Kataoka, T., Koyama, M., and Yoshida, Y.: Spectral Normalization for Generative Adversarial Networks, arXiv [preprint], https://doi.org/10.48550/arXiv.1802.05957, 2018.

National Academies of Sciences, Engineering, and Medicine (NASEM): Machine Learning and Artificial Intelligence to Advance Earth System Science, in: Opportunities and Challenges: Proceedings of a Workshop, The National Academies Press, Washington, DC, https://doi.org/10.17226/26566, 2022.

National Academies of Sciences, Engineering, and Medicine (NASEM): Foundational Research Gaps and Future Directions for Digital Twins, The National Academies Press, https://doi.org/10.17226/26894, 2023.

Niederer, S. A., Sacks, M. S., Girolami, M., and Willcox, K.: Scaling digital twins from the artisanal to the industrial, Nature Computational Science, 1, 313–320, https://doi.org/10.1038/s43588-021-00072-5, 2021.

Nieves, V., Radin, C., and Camps-Valls, G.: Predicting regional coastal sea level changes with machine learning, Sci. Rep., 11, 7650, https://doi.org/10.1038/s41598-021-87460-z, 2021.

O'Donncha, F., Hu, Y., Palmes, P., Burke, M., Filgueira, R., and Grant, J.: A spatio-temporal LSTM model to forecast across multiple temporal and spatial scales, Ecol. Inform., 69, 101687, https://doi.org/10.1016/j.ecoinf.2022.101687, 2022.

Park, S. K. and Zupanski, M.: Principles of Data Assimilation, Cambridge University Press, ISBN 978-1-108-83176-5, https://doi.org/10.1017/9781108924238, 2022.

Pathak, J., Subramanian, S., Harrington, P., Raja, S., Chattopadhyay, A., Mardani, M., Kurth, T., Hall, D., Li, Z., Azizzadenesheli, K., Hassanzadeh, P., Kashinath, K., and Anandkumar, A.: FourCastNet: A Global Data-driven High-resolution Weather Model using Adaptive Fourier Neural Operators, arXiv [preprint], https://doi.org/10.48550/arxiv.2202.11214, 2022.

Penny, S. G., Smith, T. A., Chen, T.-C., Platt, J. A., Lin, H.-Y., Goodliff, M., and Abarbanel, H. D. I.: Integrating Recurrent Neural Networks With Data Assimilation for Scalable Data-Driven State Estimation, J. Adv. Model. Earth Sy., 14, e2021MS002843, https://doi.org/10.1029/2021ms002843, 2022.

Perezhogin, P., Zanna, L., and Fernandez-Granda, C.: Generative Data-Driven Approaches for Stochastic Subgrid Parameterizations in an Idealized Ocean Model, J. Adv. Model. Earth Sy., 15, e2023MS003681, https://doi.org/10.1029/2023MS003681, 2023.

Perezhogin, P., Zhang, C., Adcroft, A., Fernandez-Granda, C., and Zanna, L.: A Stable Implementation of a Data-Driven Scale-Aware Mesoscale Parameterization, J. Adv. Model. Earth Sy., 16, e2023MS004104, https://doi.org/10.1029/2023ms004104, 2024.

Platt, J. A., Penny, S. G., Smith, T. A., Chen, T.-C., and Abarbanel, H. D. I.: Constraining chaos: Enforcing dynamical invariants in the training of reservoir computers, Chaos: An Interdisciplinary Journal of Nonlinear Science, 33, 103107, https://doi.org/10.1063/5.0156999, 2023.

Price, I., Sanchez-Gonzalez, A., Alet, F., Andersson, T. R., El-Kadi, A., Masters, D., Ewalds, T., Stott, J., Mohamed, S., Battaglia, P., Lam, R., and Willson, M.: Probabilistic weather forecasting with machine learning, Nature, 637, 84–90, https://doi.org/10.1038/s41586-024-08252-9, 2024.

Puscasu, R. M.: Integration of artificial neural networks into operational ocean wave prediction models for fast and accurate emulation of exact nonlinear interactions, Procedia Comput. Sci., 29, 1156–1170, https://doi.org/10.1016/j.procs.2014.05.104, 2014.

Rasp, S., Pritchard, M. S., and Gentine, P.: Deep learning to represent subgrid processes in climate models, P. Natl. Acad. Sci. USA, 115, 9684–9689, https://doi.org/10.1073/pnas.1810286115, 2018.

Rasp, S., Dueben, P. D., Scher, S., Weyn, J. A., Mouatadid, S., and Thuerey, N.: WeatherBench: A Benchmark Data Set for Data-Driven Weather Forecasting, J. Adv. Model. Earth Sy., 12, e2020MS002203, https://doi.org/10.1029/2020ms002203, 2020.

Rasp, S., Hoyer, S., Merose, A., Langmore, I., Battaglia, P., Russell, T., Sanchez-Gonzalez, A., Yang, V., Carver, R., Agrawal, S., Chantry, M., Bouallegue, Z. B., Dueben, P., Bromberg, C., Sisk, J., Barrington, L., Bell, A., and Sha, F.: WeatherBench 2: A Benchmark for the Next Generation of Data-Driven Global Weather Models, J. Adv. Model. Earth Sy., 16, e2023MS004019, https://doi.org/10.1029/2023ms004019, 2024.

Reichstein, M., Camps-Valls, M, Stevens, G., Jung, B., Denzler, M., Carvalhais, J., and Prabhat, N.: Deep learning and process understanding for data-driven Earth system science, Nature, 566, 195–204, https://doi.org/10.1038/s41586-019-0912-1, 2019.

Sane, A., Reichl, B. G., Adcroft, A., and Zanna, L.: Parameterizing Vertical Mixing Coefficients in the Ocean Surface Boundary Layer Using Neural Networks, J. Adv. Model. Earth Sy., 15, e2023MS003890, https://doi.org/10.1029/2023ms003890, 2023.

Sapienza, F., Bolibar, J., Schäfer, F., Groenke, B., Pal, A., Boussange, V., Heimbach, P., Hooker, G., Pérez, F., Persson, P.-O., and Rackauckas, C.: Differentiable Programming for Differential Equations: A Review, arXiv [preprint], https://doi.org/10.48550/arxiv.2406.09699, 2024.

Schneider, R., Bonavita, M., Geer, A., Arcucci, R., Dueben, P., Vitolo, C., Saux, B. L., Demir, B., and Mathieu, P.-P.: ESA-ECMWF Report on recent progress and research directions in machine learning for Earth System observation and prediction, Npj Climate and Atmospheric Science, 5, 51, https://doi.org/10.1038/s41612-022-00269-z, 2022.

Schneider, T., Lan, S., Stuart, A., and Teixeira, J.: Earth System Modeling 2.0: A Blueprint for Models That Learn From Observations and Targeted High-Resolution Simulations, Geophys. Res. Lett., 44, 12396–12417, https://doi.org/10.1002/2017gl076101, 2017.

Schneider, T., Behera, S., Boccaletti, G., Deser, C., Emanuel, K., Ferrari, R., Leung, L. R., Lin, N., Müller, T., Navarra, A., Ndiaye, O., Stuart, A., Tribbia, J. and Yamagata, T.: Harnessing AI and computing to advance climate modelling and prediction, Nat. Clim. Change, 13, 887–889, https://doi.org/10.1038/s41558-023-01769-3, 2023.

Shen, C., Appling, A. P., Gentine, P., Bandai, T., Gupta, H., Tartakovsky, A., Baity-Jesi, M., Fenicia, F., Kifer, D., Li, L., Liu, X., Ren, W., Zheng, Y., Harman, C. J., Clark, M., Farthing, M., Feng, D., Kumar, P., Aboelyazeed, D., Rahmani, F., Song, Y., Beck, H. E., Bindas, T., Dwivedi, D., Fang, K., Höge, M., Rackauckas, C., Mohanty, B., Roy, T., Xu, C., and Lawson, K.: Dikerentiable modelling to unify machine learning and physical models for

geosciences, Nature Reviews Earth & Environment, 4, 552–567, https://doi.org/10.1038/s43017-023-00450-9, 2023.

Sinha, A. and Abernathey, R.: Estimating ocean surface currents with machine learning, Frontiers in Marine Science, 8, 672477, https://doi.org/10.3389/fmars.2021.672477, 2021.

Smets, B. M., Portegies, J., Bekkers, E. J., and Duits, R.: PDE-based group equivariant convolutional neural networks, J. Math. Imaging Vis., 65, 209–239, 2023.

Smith, T. A., Penny, S. G., Platt, J. A., and Chen, T.-C.: Temporal Subsampling Diminishes Small Spatial Scales in Recurrent Neural Network Emulators of Geophysical Turbulence, J. Adv. Model. Earth Sy., 15, e2023MS003792, https://doi.org/10.1029/2023MS003792, 2023.

Storto, A., Frolov, S., Slivinski, L., and Yang, C.: Correction of Air-Sea Heat Fluxes in the NEMO Ocean General Circulation Model Using Neural Networks, Geosci. Model Dev. Discuss. [preprint], https://doi.org/10.5194/gmd-2024-185, in review, 2024.

Subel, A. and Zanna, L.: Building Ocean Climate Emulators, arXiv [preprint], https://doi.org/10.48550/arxiv.2402.04342, 2024.

Sun, Y., Sowunmi, O., Egele, R., Narayanan, S. H. K., Roekel, L. V., and Balaprakash, P.: Streamlining Ocean Dynamics Modeling with Fourier Neural Operators: A Multiobjective Hyperparameter and Architecture Optimization Approach, Mathematics, 12, 1483, https://doi.org/10.3390/math12101483, 2024.

Wang, C., Pritchard, M. S., Brenowitz, N., Cohen, Y., Bonev, B., Kurth, T., Durran, D., and Pathak, J.: Coupled Ocean-Atmosphere Dynamics in a Machine Learning Earth System Model, arXiv [preprint], https://doi.org/10.48550/arXiv.2406.08632, 2024.

Wang, X., Wang, R., Hu, N., Wang, P., Huo, P., Wang, G., Wang, H., Wang, S., Zhu, J., Xu, J., Yin, J., Bao, S., Luo, C., Zu, Z., Han, Y., Zhang, W., Ren, K., Deng, K., and Song, J.: XiHe: A Data-Driven Model for Global Ocean Eddy-Resolving Forecasting, arXiv [preprint], https://doi.org/10.48550/arxiv.2402.02995, 2024.

Watt-Meyer, O., Dresdner, G., McGibbon, J., Clark, S. K., Henn, B., Duncan, J., Brenowitz, N. D., Kashinath, K., Pritchard, M. S., Bonev, B., Peters, M. E., and Bretherton, C. S.: ACE: A fast, skillful learned global atmospheric model for climate prediction, arXiv [preprint], https://doi.org/10.48550/arxiv.2310.02074, 2023.

Willcox, K. E., Ghattas, O., and Heimbach, P.: The imperative of physics-based modeling and inverse theory in computational science, Nature Computational Science, 1, 166–168, https://doi.org/10.1038/s43588-021-00040-z, 2021.

Weyn, J. A., Durran, D. R., and Caruana, R.: Can Machines Learn to Predict Weather? Using Deep Learning to Predict Gridded 500-hPa Geopotential Height From Historical Weather Data, J. Adv. Model. Earth Sy., 11, 2680–2693, https://doi.org/10.1029/2019MS001705, 2019.

Weyn, J. A., Durran, D. R., and Caruana, R.: Improving Data-Driven Global Weather Prediction Using Deep Convolutional Neural Networks on a Cubed Sphere, J. Adv. Model. Earth Sy., 12, e2020MS002109, https://doi.org/10.1029/2020MS002109, 2020.

Weyn, J. A., Durran, D. R., Caruana, R., and Cresswell-Clay, N.: Sub-Seasonal Forecasting With a Large Ensemble of Deep-Learning Weather Prediction Models, J. Adv. Model. Earth Sy., 13, e2021MS002502, https://doi.org/10.1029/2021MS002502, 2021.

Wolff, S., O'Donncha, F., and Chen, B.: Statistical and machine learning ensemble modelling to forecast sea surface temperature, J. Marine Syst., 208, 103347, https://doi.org/10.1016/j.jmarsys.2020.103347, 2020.

Xie, W., Xu, G., Zhang, H., and Dong, C.: Developing a deep learning-based storm surge forecasting model, Ocean Model., 182, 102179, https://doi.org/10.1016/j.ocemod.2023.102179, 2023.

Xu, S., Dai, D., Cui, X., Yin, X., Jiang, S., Pan, H., and Wang, G.: A deep learning approach to predict sea surface temperature based on multiple modes, Ocean Model., 181, 102158, https://doi.org/10.1016/j.ocemod.2022.102158, 2023.

Yu, S., Hannah, W., Peng, L., et al.: ClimSim: A large multi-scale dataset for hybrid physics-ML climate emulation, arXiv [preprint], https://doi.org/10.48550/arxiv.2306.08754, 2023.

Zanna, L. and Bolton, T.: Data-Driven Equation Discovery of Ocean Mesoscale Closures, Geophys. Res. Lett., 47, e2020GL088376-33, https://doi.org/10.1029/2020gl088376, 2020.

Zanna, L. and Bolton, T.: Deep Learning of Unresolved Turbulent Ocean Processes in Climate Models, in: Deep Learning for the Earth Sciences: A Comprehensive Approach to Remote Sensing, Climate Science, and Geosciences, edited by: Camps-Valls, G., Tuia, D., Zhu, X. X., and Reichstein, M., Wiley & Sons, 298–306, https://doi.org/10.1002/9781119646181.ch20, 2021.

Zhang, C., Perezhogin, P., Gultekin, C., Adcroft, A., Fernandez-Granda, C., and Zanna, L.: Implementation and Evaluation of a Machine Learned Mesoscale Eddy Parameterization Into a Numerical Ocean Circulation Model, J. Adv. Model. Earth Sy., 15, e2023MS003697, https://doi.org/10.1029/2023ms003697, 2023.

Zubov, K., McCarthy, Z., Ma, Y., Calisto, F., Pagliarino, V., Azeglio, S., Bottero, L., Luján, E., Sulzer, V., Bharambe, A., Vinchhi, N., Balakrishnan, K., Upadhyay, D., and Rackauckas, C.: Neuralpde: Automating physics-informed neural networks (PINNs) with error approximations, arXiv [preprint], https://doi.org/10.48550/arXiv.2107.09443, 2021.