Old dog, new trick: Reservoir computing advances machine learning for climate modeling

Christopher S. Bretherton[1]

[1]Allen Institute for Artificial Intelligence

April 30, 2023

Abstract

Physics-informed machine learning (ML) applied to geophysical simulation is developing explosively. Recently, graph neural net and vision transformer architectures have shown 1-7 day global weather forecast skill superior to any conventional model with integration times over 1000 times faster, but longer simulations rapidly degrade. ML that achieves high skill in both weather and climate applications is a tougher goal. This Commentary was inspired by \citeA{ArcomanoEtAl2023}, who show impressive progress toward that goal using hybrid ML, combining reservoir computing to a coarse-grid climate model and coupling to a separate data-driven reservoir computing model that interactively predicts sea-surface temperature. This opens new horizons; where will the next ML breakthrough come from, and is conventional climate modeling about to be disrupted?

Old dog, new trick: Reservoir computing advances machine learning for climate modeling

Christopher S. Bretherton

[1]Allen Institute for Artificial Intelligence

Key Points:

- Arcomano et al. (2023) combined reservoir computing (RC) with a coarse-grid climate model for data-driven ocean-coupled simulations
- By building long-term memory into predictions, RC nearly removes climate bias
- Challenges remain with interpretability and scalability to fine-scale prediction that new machine learning approaches may soon surmount

Corresponding author: Christopher S. Bretherton, `christopherb@allenai.org`

Abstract

Physics-informed machine learning (ML) applied to geophysical simulation is developing explosively. Recently, graph neural net and vision transformer architectures have shown 1-7 day global weather forecast skill superior to any conventional model with integration times over 1000 times faster, but longer simulations rapidly degrade. ML that achieves high skill in both weather and climate applications is a tougher goal. This Commentary was inspired by Arcomano et al. (2023), who show impressive progress toward that goal using hybrid ML, combining reservoir computing to a coarse-grid climate model and coupling to a separate data-driven reservoir computing model that interactively predicts sea-surface temperature. This opens new horizons; where will the next ML breakthrough come from, and is conventional climate modeling about to be disrupted?

Plain Language Summary

Many new research groups are making rapid progress in applying diverse machine learning methodologies to weather forecasting and climate modeling. These new approaches could make simulations that are 1000x faster than conventional approaches for the same fidelity. One successful approach for weather forecasting has been replacing an entire conventional global atmospheric model with a machine learning emulator, but so far the climates generated by long simulations using this approach have had substantial biases in average temperature or precipitation. An alternate new approach, hybrid reservoir computing, combines the conventional model with a form of machine learning that remembers the recent atmospheric evolution. It produces much better climate simulations, including realistic El-Nino/La Nina variability, but on a much coarser spatial grid. This opens new horizons; where will the next ML breakthrough come from, and is conventional climate modeling about to be disrupted?

1 Introduction

Over the past five years, weather and climate modeling have become hot topics in physics-informed machine learning (ML), as the domain science and machine learning communities start to cross-fertilize. One central vision is global weather and climate emulators, which use machine learning to replace or supplement conventional global atmospheric prediction models.

The climate community has long used global weather simulators based on numerical discretizations to appropriate governing equations as part of climate models, as discussed in textbooks such as Drake (2014). Climate, after all, is comprised of the slowly-varying statistics of weather, including its means and extremes (AMS, 2023). The physical principles governing weather forecasting (including climate-relevant aspects such as clouds, aerosols and chemistry, surface exchange, and interactions with land and sea-ice) are mostly well understood and observationally tested by the instrumental record. 'Seamless' modeling of weather and climate (Rodwell & Palmer, 2007), i.e. insisting that a climate model accurately reproduce weather-induced variability and covariability of its predictands, also guards against overfitting of adjustable parameters to limited time-mean observational constraints. The same principles apply to ML emulators, whether trained on observational reanalyses in the present climate, or on fine-grid reference atmosphere simulations across a range of climates.

Some salient characteristics of a global atmosphere emulator for climate modeling are as follows:

- Stable
- Accurate (weather forecasts and climate means/extremes)
- Localizable (can make accurate predictions with fine spatial resolution)

- Interpretable (e.g. satisfies conservation principles and physically realistic bounds)
- Extensible (e.g. aerosols, chemistry)
- Naturally couplable (to ocean, land, ice)
- Time-efficient (vs. a conventional climate model of comparable skill)

These characteristics can be recalled by spelling out their first letters to get SALIENT.

Arcomano et al. (2023), hereafter A23, show impressive progress toward that goal using hybrid ML, combining reservoir computing with a coarse-grid climate model. Reservoir computing (Wikipedia, 2023) is an older type of long-short time memory ML that involves only linear calculations. Nevertheless, their global simulations are fast, stable, have reasonable weather skill, have remarkably little climate bias, and importantly can be coupled to a separate data-driven reservoir model with a longer memory time scale that interactively predicts sea-surface temperature (SST). As well as achieving low climate bias for both the atmosphere and coupled SST, the coupled ML system spontaneously simulates fairly realistic El Nino Southern Oscillation cycles, a first for this type of emulation.

This commentary discusses how well A23 have already achieved 'saliency', followed by some remaining challenges for their approach. We compare their work with some other promising emulation approaches, and consider some general new research horizons for ML atmospheric emulators. Given growing research interest in this area and the prospect for further rapid progress, we suggest that the climate projection community may be closer than widely appreciated to fully embracing ML into mainstream development.

2 A23's achievements and upcoming challenges

Let's consider how well A23's hybrid RC methodology meets the SALIENT characteristics of a good climate emulator. It satisfies S, A and T - it is stable, time-efficient, and accurate enough for climate modeling. It is well on the way to satisfing N, through its successful coupling with a RC sea-surface temperature model. It has not yet been coupled to externally developed model components such as conventional or ML-based ocean, land or sea-ice models, but this (as well as extensibility (E) to include other atmospheric components like chemistry and aerosols) could naturally be done mostly through its conventional AGCM component, SPEEDY (Molteni, 2003).

Its current incarnation is somewhat interpretable (I). The RC updates are applied in localized $7.5^\circ \times 10^\circ$ patches of grid columns, so tendency budgets of prognostic variables could be computed over such patches. A23's addition of precipitation as a novel diagnostic RC output is also a plus. However, the current version of the RC model does not automatically satisfy heat, moisture or momentum conservation equations in which a tendency can be ascribed to a flux convergence, nor does it currently correct SPEEDY's predictions of radiative fluxes at the surface or top of the atmosphere.

Perhaps the biggest shortcoming of the current hybrid RC is in localization (L). Unlike state-of-the-art full model emulators with 30 km horizontal grid resolutions, its coarse ($3.75^\circ \times 5^\circ$) horizontal grid doesn't resolve topographic details or intense storm systems such as tropical cyclones. Its 8 vertical grid levels also is a factor of 10 smaller than many current weather and climate models. The RC implementation is memory-intensive and might be difficult to scale to a grid 10-fold smaller in each direction. An ML-based super-resolution generator for each patch or grid column based on adversarial (Leinonen et al., 2021) or diffusion (Wang et al., 2020) modeling might help with this issue.

3 Hybrid RC vs. full-model emulation

What type of emulator is most promising for climate modeling? Numerous research groups have begun working on full model emulation (FME), in which an ML architecture such as a U-Net (Weyn et al., 2020, 2021), a vision transformer (Pathak et al., 2022), or a graph neural net (Keisler, 2022), is trained to forecast global weather. Recent FME papers have demonstrated forecast skill out to seven days that is superior to the world's best operational forecast model, made a thousand times faster using compact, energy-efficient purpose-built hardware rather than expensive supercomputers (Bi et al., 2022; Lam et al., 2022). A pre-trained FME has been proposed as a 'foundation' model for weather and climate simulation (Nguyen et al., 2023). Further rapid progress seems inevitable. However, these FME approaches are generally not yet accurate or even stable over the longer forecast periods needed for climate. In addition, fundamental unaddressed problems remain with FME, including coupling to other model components, extensibility to advection of trace species and hydrometeors, or even the physical interpretability of the resulting simulations.

In contrast, A23 and underlying previous work (Wikner et al., 2020; Arcomano et al., 2022) adopted a hybrid approach, in which ML elements are combined with a coarse-grid conventional climate model to improve its skill. The speed of the resulting simulations can be no faster than that of the coarse climate model, but even a twofold coarser horizontal and vertical grid spacing halves the number of time steps and reduces the overall computational effort ten-fold. A23 use the intermediate-complexity SPEEDY AGCM, which has very coarse $3.75^\circ \times 5^\circ$ grid resolution, 8 vertical levels, and simplified physical parameterizations. SPEEDY simulates a day per 2 seconds of execution time (Arcomano et al., 2022), comparable to current FME approaches (Pathak et al., 2022; Lam et al., 2022). The ML element (reservoir computing) with which SPEEDY is combined uses a much longer time step than SPEEDY, so it doesn't significantly slow down simulations, while it is surprisingly effective in counteracting SPEEDY's considerable systematic weather and climate biases. Thus A23 is computationally competitive with FME, though using a larger and less energy-efficient cluster of 1152 processors. An important trade-off is the lack of grid resolution, a factor of 15-20 smaller than the $0.25^\circ \times 0.25^\circ$ grid and up to 30 vertical grid levels used by recent FME approaches trained on the ERA5 reanalysis. A broader challenge with most hybrid ML is that the ML must be trained 'offline' with other model components given, but these other model components can react to the ML 'online' during ML-augmented simulations. Thus 'offline' optimization of ML weights doesn't guarantee improved online simulation accuracy or even stability (Brenowitz & Bretherton, 2019). Because RC weight optimization is linear, it partly sidesteps this problem and currently achieves much better climate stability and skill than FME. Another hybrid approach trained based on nudging coarse-grid simulations to reference reanalyses or fine-grid simulations also increases forecast skill and reduces climate biases of a coarse-grid target model (Watt-Meyer et al., 2021; Bretherton et al., 2022; Clark et al., 2022), but to a lesser extent than RC.

Neither SPEEDY or a pure reservoir computing approach based on the SPEEDY grid are nearly as skillful in making weather forecasts or simulating the observed mean state of the global atmosphere (Arcomano et al., 2020). Like FME, the hybrid RC method works by incorporating multiple spatial scales into the learning process. SPEEDY can be viewed as an efficient physics-based way to handle long-range spatial interactions affecting the atmospheric state in each patch, while the reservoir computing corrects local systematic errors associated with parameterizations and numerical discretization error. More so than FME approaches to date, the memory built into RC helps ensure that the ML also removes slowly-developing mean-state biases.

4 New horizons and prospects

Whether a hybrid method like A23's RC or a full model emulator proves most suitable for seamless ML weather and climate emulation, several issues will keep the ML development community busy, including:

- Reliable reference training data over a range of climates
- Building in conservation principles
- Out-of-sample extrapolation
- Strategies for natural model component coupling
- Achieving predictive locality
- Gaining the confidence of weather and climate model domain experts and users

Given recent results of A23 and others, none of these issues need block the weather and climate modeling community from starting to operationalize ML emulators within as little as a year or two, given their speed and affordability and consequent transformative potential for large ensemble simulation, data assimilation, etc. A23's RC model for coupling sea-surface temperature to a hybrid-RC atmosphere points the way toward emulators of full dynamical ocean, land and sea-ice model components, etc., needed to realize this vision. Hold onto your saddle for an exciting ride!

Open Research

No new data was used in writing this Commentary.

Acknowledgments

Thanks to the AI2 Climate Modeling team and numerous colleagues for discussions that informed this Commentary.

References

AMS. (2023). *Glossary of meteorology.* Retrieved from `https://glossary.ametsoc.org/wiki/Climate`

Arcomano, T., Szunyogh, I., Pathak, J., Wikner, A., Hunt, B. R., & Ott, E. (2020). A machine learning-based global atmospheric forecast model. *Geophysical Research Letters*, *47*, e2020GL087776. doi: 10.1029/2020GL087776

Arcomano, T., Szunyogh, I., Wikner, A., Hunt, B. R., & Ott, E. (2023). A hybrid atmospheric model incorporating machine learning can capture dynamical processes not captured by its physics-based component. *Geophysical Research Letters (accepted)*. doi: 10.1029/2022GL102649

Arcomano, T., Szunyogh, I., Wikner, A., Pathak, J., Hunt, B. R., & Ott, E. (2022). A hybrid approach to atmospheric modeling that combines machine learning with a physics-based numerical model. *Journal of Advances in Modeling Earth Systems*, *14*, e2021MS002712. doi: 10.1029/2021MS002712

Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., & Tian, Q. (2022). Pangu-weather: A 3d high-resolution model for fast and accurate global weather forecast. *arXiv*. doi: 10.48550/arXiv.2211.02556

Brenowitz, N. D., & Bretherton, C. S. (2019). Spatially extended tests of a neural network parametrization trained by coarse-graining. *J. Adv. Model. Earth Syst.*, *11*, 2728-2744. doi: 10.1029/2019MS001711

Bretherton, C. S., Henn, B., Kwa, A., Brenowitz, N. D., Watt-Meyer, O., McGibbon, J., ... Harris, L. (2022). Correcting coarse-grid weather and climate models by machine learning from global storm-resolving simulations. *Journal of Advances in Modeling Earth Systems*, *14*, e2021MS002794. doi:

10.1029/2021MS002794
Clark, S. K., Brenowitz, N. D., Henn, B., Kwa, A., McGibbon, J., Perkins, W. A., . . . Harris, L. M. (2022). Correcting a 200-km resolution climate model in multiple climates by machine learning from 25-km resolution simulations. *Journal of Advances in Modeling Earth Systems*, e2022MS003219. doi: 10.1029/2022MS003219
Drake, J. B. (2014). *Climate modeling for scientists and engineers* (1st ed.). SIAM.
Keisler, R. (2022). Forecasting global weather with graph neural networks. *arXiv*. doi: 10.48550/arXiv.2202.07575
Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Pritzel, A., . . . Battaglia, P. (2022). Graphcast: Learning skillful medium-range global weather forecasting. *arXiv*. doi: 10.48550/arXiv.2212.12794
Leinonen, J., Nerini, D., & Berne, A. (2021). Stochastic super-resolution for downscaling time-evolving atmospheric fields with a generative adversarial network. *IEEE Transactions on Geoscience and Remote Sensing*, *59*, 7211–7223. doi: 10.1109/tgrs.2020.3032790
Molteni, F. (2003). Atmospheric simulations using a gcm with simplified physical parametrizations. i: model climatology and variability in multi-decadal experiments. *Climate Dynamics*, *20*, 175–191. doi: 10.1007/s00382-002-0268-2
Nguyen, T., Brandstetter, J., Kapoor, A., Gupta, J. K., & Grover, A. (2023). Climax: A foundation model for weather and climate. *arXiv*. doi: 10.48550/arXiv.2301.10343
Pathak, J., Subramanian, S., Harrington, P., Raja, S., Chattopadhyay, A., Mardani, M., . . . Anandkumar, A. (2022). Fourcastnet: A global data-driven high-resolution weather model using adaptive fourier neural operators. *arXiv*. doi: 10.48550/arXiv.2202.11214
Rodwell, M. J., & Palmer, T. N. (2007). Using numerical weather prediction to assess climate models. *Quarterly Journal of the Royal Meteorological Society*, *133*, 129-146. doi: 10.1002/qj.23
Wang, C., Bentivegna, E., Zhou, W., Klein, L., & Elmegreen, B. (2020). Physics-informed neural network super resolution for advection-diffusion models. *arXiv*. doi: 10.48550/arXiv.2011.02519
Watt-Meyer, O., Brenowitz, N. D., Clark, S. K., Henn, B., Kwa, A., McGibbon, J., . . . Bretherton, C. S. (2021). Correcting weather and climate models by machine learning nudged historical simulations. *Geophys. Res. Lett.*, *48*, e2021GL092555. doi: 10.1029/2021GL092555
Weyn, J. A., Durran, D. R., & Caruana, R. (2020). Improving data-driven global weather prediction using deep convolutional neural networks on a cubed sphere. *Journal of Advances in Modeling Earth Systems*, *12*, e2020MS002109. doi: 10.1029/2020MS002109
Weyn, J. A., Durran, D. R., Caruana, R., & Cresswell-Clay, N. (2021). Sub-seasonal forecasting with a large ensemble of deep-learning weather prediction models. *Journal of Advances in Modeling Earth Systems*, *13*, e2021MS002502. doi: 10.1029/2021MS002502
Wikipedia. (2023). *Reservoir computing.* Retrieved from `https://en.wikipedia.org/wiki/Reservoir_computing`
Wikner, A., Pathak, J., Hunt, B., Girvan, M., Arcomano, T., Szunyogh, I., . . . Ott, E. (2020). Combining machine learning with knowledge-based modeling for scalable forecasting and subgrid-scale closure of large, complex, spatiotemporal systems. *Chaos*, *30*, 053111. doi: 10.1063/5.0005541

Old dog, new trick: Reservoir computing advances machine learning for climate modeling

Christopher S. Bretherton

[1]Allen Institute for Artificial Intelligence

Key Points:

- Arcomano et al. (2023) combined reservoir computing (RC) with a coarse-grid climate model for data-driven ocean-coupled simulations
- By building long-term memory into predictions, RC nearly removes climate bias
- Challenges remain with interpretability and scalability to fine-scale prediction that new machine learning approaches may soon surmount

Corresponding author: Christopher S. Bretherton, `christopherb@allenai.org`

Abstract

Physics-informed machine learning (ML) applied to geophysical simulation is developing explosively. Recently, graph neural net and vision transformer architectures have shown 1-7 day global weather forecast skill superior to any conventional model with integration times over 1000 times faster, but longer simulations rapidly degrade. ML that achieves high skill in both weather and climate applications is a tougher goal. This Commentary was inspired by Arcomano et al. (2023), who show impressive progress toward that goal using hybrid ML, combining reservoir computing to a coarse-grid climate model and coupling to a separate data-driven reservoir computing model that interactively predicts sea-surface temperature. This opens new horizons; where will the next ML breakthrough come from, and is conventional climate modeling about to be disrupted?

Plain Language Summary

Many new research groups are making rapid progress in applying diverse machine learning methodologies to weather forecasting and climate modeling. These new approaches could make simulations that are 1000x faster than conventional approaches for the same fidelity. One successful approach for weather forecasting has been replacing an entire conventional global atmospheric model with a machine learning emulator, but so far the climates generated by long simulations using this approach have had substantial biases in average temperature or precipitation. An alternate new approach, hybrid reservoir computing, combines the conventional model with a form of machine learning that remembers the recent atmospheric evolution. It produces much better climate simulations, including realistic El-Nino/La Nina variability, but on a much coarser spatial grid. This opens new horizons; where will the next ML breakthrough come from, and is conventional climate modeling about to be disrupted?

1 Introduction

Over the past five years, weather and climate modeling have become hot topics in physics-informed machine learning (ML), as the domain science and machine learning communities start to cross-fertilize. One central vision is global weather and climate emulators, which use machine learning to replace or supplement conventional global atmospheric prediction models.

The climate community has long used global weather simulators based on numerical discretizations to appropriate governing equations as part of climate models, as discussed in textbooks such as Drake (2014). Climate, after all, is comprised of the slowly-varying statistics of weather, including its means and extremes (AMS, 2023). The physical principles governing weather forecasting (including climate-relevant aspects such as clouds, aerosols and chemistry, surface exchange, and interactions with land and sea-ice) are mostly well understood and observationally tested by the instrumental record. 'Seamless' modeling of weather and climate (Rodwell & Palmer, 2007), i.e. insisting that a climate model accurately reproduce weather-induced variability and covariability of its predictands, also guards against overfitting of adjustable parameters to limited time-mean observational constraints. The same principles apply to ML emulators, whether trained on observational reanalyses in the present climate, or on fine-grid reference atmosphere simulations across a range of climates.

Some salient characteristics of a global atmosphere emulator for climate modeling are as follows:

- Stable
- Accurate (weather forecasts and climate means/extremes)
- Localizable (can make accurate predictions with fine spatial resolution)

- Interpretable (e.g. satisfies conservation principles and physically realistic bounds)
- Extensible (e.g. aerosols, chemistry)
- Naturally couplable (to ocean, land, ice)
- Time-efficient (vs. a conventional climate model of comparable skill)

These characteristics can be recalled by spelling out their first letters to get SALIENT.

Arcomano et al. (2023), hereafter A23, show impressive progress toward that goal using hybrid ML, combining reservoir computing with a coarse-grid climate model. Reservoir computing (Wikipedia, 2023) is an older type of long-short time memory ML that involves only linear calculations. Nevertheless, their global simulations are fast, stable, have reasonable weather skill, have remarkably little climate bias, and importantly can be coupled to a separate data-driven reservoir model with a longer memory time scale that interactively predicts sea-surface temperature (SST). As well as achieving low climate bias for both the atmosphere and coupled SST, the coupled ML system spontaneously simulates fairly realistic El Nino Southern Oscillation cycles, a first for this type of emulation.

This commentary discusses how well A23 have already achieved ‘saliency’, followed by some remaining challenges for their approach. We compare their work with some other promising emulation approaches, and consider some general new research horizons for ML atmospheric emulators. Given growing research interest in this area and the prospect for further rapid progress, we suggest that the climate projection community may be closer than widely appreciated to fully embracing ML into mainstream development.

2 A23’s achievements and upcoming challenges

Let’s consider how well A23’s hybrid RC methodology meets the SALIENT characteristics of a good climate emulator. It satisfies S, A and T - it is stable, time-efficient, and accurate enough for climate modeling. It is well on the way to satisfing N, through its successful coupling with a RC sea-surface temperature model. It has not yet been coupled to externally developed model components such as conventional or ML-based ocean, land or sea-ice models, but this (as well as extensibility (E) to include other atmospheric components like chemistry and aerosols) could naturally be done mostly through its conventional AGCM component, SPEEDY (Molteni, 2003).

Its current incarnation is somewhat interpretable (I). The RC updates are applied in localized $7.5^\circ \times 10^\circ$ patches of grid columns, so tendency budgets of prognostic variables could be computed over such patches. A23’s addition of precipitation as a novel diagnostic RC output is also a plus. However, the current version of the RC model does not automatically satisfy heat, moisture or momentum conservation equations in which a tendency can be ascribed to a flux convergence, nor does it currently correct SPEEDY’s predictions of radiative fluxes at the surface or top of the atmosphere.

Perhaps the biggest shortcoming of the current hybrid RC is in localization (L). Unlike state-of-the-art full model emulators with 30 km horizontal grid resolutions, its coarse ($3.75^\circ \times 5^\circ$) horizontal grid doesn’t resolve topographic details or intense storm systems such as tropical cyclones. Its 8 vertical grid levels also is a factor of 10 smaller than many current weather and climate models. The RC implementation is memory-intensive and might be difficult to scale to a grid 10-fold smaller in each direction. An ML-based super-resolution generator for each patch or grid column based on adversarial (Leinonen et al., 2021) or diffusion (Wang et al., 2020) modeling might help with this issue.

3 Hybrid RC vs. full-model emulation

What type of emulator is most promising for climate modeling? Numerous research groups have begun working on full model emulation (FME), in which an ML architecture such as a U-Net (Weyn et al., 2020, 2021), a vision transformer (Pathak et al., 2022), or a graph neural net (Keisler, 2022), is trained to forecast global weather. Recent FME papers have demonstrated forecast skill out to seven days that is superior to the world's best operational forecast model, made a thousand times faster using compact, energy-efficient purpose-built hardware rather than expensive supercomputers (Bi et al., 2022; Lam et al., 2022). A pre-trained FME has been proposed as a 'foundation' model for weather and climate simulation (Nguyen et al., 2023). Further rapid progress seems inevitable. However, these FME approaches are generally not yet accurate or even stable over the longer forecast periods needed for climate. In addition, fundamental unaddressed problems remain with FME, including coupling to other model components, extensibility to advection of trace species and hydrometeors, or even the physical interpretability of the resulting simulations.

In contrast, A23 and underlying previous work (Wikner et al., 2020; Arcomano et al., 2022) adopted a hybrid approach, in which ML elements are combined with a coarse-grid conventional climate model to improve its skill. The speed of the resulting simulations can be no faster than that of the coarse climate model, but even a twofold coarser horizontal and vertical grid spacing halves the number of time steps and reduces the overall computational effort ten-fold. A23 use the intermediate-complexity SPEEDY AGCM, which has very coarse $3.75^\circ \times 5^\circ$ grid resolution, 8 vertical levels, and simplified physical parameterizations. SPEEDY simulates a day per 2 seconds of execution time (Arcomano et al., 2022), comparable to current FME approaches (Pathak et al., 2022; Lam et al., 2022). The ML element (reservoir computing) with which SPEEDY is combined uses a much longer time step than SPEEDY, so it doesn't significantly slow down simulations, while it is surprisingly effective in counteracting SPEEDY's considerable systematic weather and climate biases. Thus A23 is computationally competitive with FME, though using a larger and less energy-efficient cluster of 1152 processors. An important trade-off is the lack of grid resolution, a factor of 15-20 smaller than the $0.25^\circ \times 0.25^\circ$ grid and up to 30 vertical grid levels used by recent FME approaches trained on the ERA5 reanalysis. A broader challenge with most hybrid ML is that the ML must be trained 'offline' with other model components given, but these other model components can react to the ML 'online' during ML-augmented simulations. Thus 'offline' optimization of ML weights doesn't guarantee improved online simulation accuracy or even stability (Brenowitz & Bretherton, 2019). Because RC weight optimization is linear, it partly sidesteps this problem and currently achieves much better climate stability and skill than FME. Another hybrid approach trained based on nudging coarse-grid simulations to reference reanalyses or fine-grid simulations also increases forecast skill and reduces climate biases of a coarse-grid target model (Watt-Meyer et al., 2021; Bretherton et al., 2022; Clark et al., 2022), but to a lesser extent than RC.

Neither SPEEDY or a pure reservoir computing approach based on the SPEEDY grid are nearly as skillful in making weather forecasts or simulating the observed mean state of the global atmosphere (Arcomano et al., 2020). Like FME, the hybrid RC method works by incorporating multiple spatial scales into the learning process. SPEEDY can be viewed as an efficient physics-based way to handle long-range spatial interactions affecting the atmospheric state in each patch, while the reservoir computing corrects local systematic errors associated with parameterizations and numerical discretization error. More so than FME approaches to date, the memory built into RC helps ensure that the ML also removes slowly-developing mean-state biases.

4 New horizons and prospects

Whether a hybrid method like A23's RC or a full model emulator proves most suitable for seamless ML weather and climate emulation, several issues will keep the ML development community busy, including:

- Reliable reference training data over a range of climates
- Building in conservation principles
- Out-of-sample extrapolation
- Strategies for natural model component coupling
- Achieving predictive locality
- Gaining the confidence of weather and climate model domain experts and users

Given recent results of A23 and others, none of these issues need block the weather and climate modeling community from starting to operationalize ML emulators within as little as a year or two, given their speed and affordability and consequent transformative potential for large ensemble simulation, data assimilation, etc. A23's RC model for coupling sea-surface temperature to a hybrid-RC atmosphere points the way toward emulators of full dynamical ocean, land and sea-ice model components, etc., needed to realize this vision. Hold onto your saddle for an exciting ride!

Open Research

No new data was used in writing this Commentary.

Acknowledgments
Thanks to the AI2 Climate Modeling team and numerous colleagues for discussions that informed this Commentary.

References

AMS. (2023). *Glossary of meteorology.* Retrieved from `https://glossary.ametsoc.org/wiki/Climate`

Arcomano, T., Szunyogh, I., Pathak, J., Wikner, A., Hunt, B. R., & Ott, E. (2020). A machine learning-based global atmospheric forecast model. *Geophysical Research Letters*, *47*, e2020GL087776. doi: 10.1029/2020GL087776

Arcomano, T., Szunyogh, I., Wikner, A., Hunt, B. R., & Ott, E. (2023). A hybrid atmospheric model incorporating machine learning can capture dynamical processes not captured by its physics-based component. *Geophysical Research Letters (accepted)*. doi: 10.1029/2022GL102649

Arcomano, T., Szunyogh, I., Wikner, A., Pathak, J., Hunt, B. R., & Ott, E. (2022). A hybrid approach to atmospheric modeling that combines machine learning with a physics-based numerical model. *Journal of Advances in Modeling Earth Systems*, *14*, e2021MS002712. doi: 10.1029/2021MS002712

Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., & Tian, Q. (2022). Pangu-weather: A 3d high-resolution model for fast and accurate global weather forecast. *arXiv*. doi: 10.48550/arXiv.2211.02556

Brenowitz, N. D., & Bretherton, C. S. (2019). Spatially extended tests of a neural network parametrization trained by coarse-graining. *J. Adv. Model. Earth Syst.*, *11*, 2728-2744. doi: 10.1029/2019MS001711

Bretherton, C. S., Henn, B., Kwa, A., Brenowitz, N. D., Watt-Meyer, O., McGibbon, J., ... Harris, L. (2022). Correcting coarse-grid weather and climate models by machine learning from global storm-resolving simulations. *Journal of Advances in Modeling Earth Systems*, *14*, e2021MS002794. doi:

10.1029/2021MS002794

Clark, S. K., Brenowitz, N. D., Henn, B., Kwa, A., McGibbon, J., Perkins, W. A., ... Harris, L. M. (2022). Correcting a 200-km resolution climate model in multiple climates by machine learning from 25-km resolution simulations. *Journal of Advances in Modeling Earth Systems*, e2022MS003219. doi: 10.1029/2022MS003219

Drake, J. B. (2014). *Climate modeling for scientists and engineers* (1st ed.). SIAM.

Keisler, R. (2022). Forecasting global weather with graph neural networks. *arXiv*. doi: 10.48550/arXiv.2202.07575

Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Pritzel, A., ... Battaglia, P. (2022). Graphcast: Learning skillful medium-range global weather forecasting. *arXiv*. doi: 10.48550/arXiv.2212.12794

Leinonen, J., Nerini, D., & Berne, A. (2021). Stochastic super-resolution for downscaling time-evolving atmospheric fields with a generative adversarial network. *IEEE Transactions on Geoscience and Remote Sensing*, *59*, 7211–7223. doi: 10.1109/tgrs.2020.3032790

Molteni, F. (2003). Atmospheric simulations using a gcm with simplified physical parametrizations. i: model climatology and variability in multi-decadal experiments. *Climate Dynamics*, *20*, 175–191. doi: 10.1007/s00382-002-0268-2

Nguyen, T., Brandstetter, J., Kapoor, A., Gupta, J. K., & Grover, A. (2023). Climax: A foundation model for weather and climate. *arXiv*. doi: 10.48550/arXiv.2301.10343

Pathak, J., Subramanian, S., Harrington, P., Raja, S., Chattopadhyay, A., Mardani, M., ... Anandkumar, A. (2022). Fourcastnet: A global data-driven high-resolution weather model using adaptive fourier neural operators. *arXiv*. doi: 10.48550/arXiv.2202.11214

Rodwell, M. J., & Palmer, T. N. (2007). Using numerical weather prediction to assess climate models. *Quarterly Journal of the Royal Meteorological Society*, *133*, 129-146. doi: 10.1002/qj.23

Wang, C., Bentivegna, E., Zhou, W., Klein, L., & Elmegreen, B. (2020). Physics-informed neural network super resolution for advection-diffusion models. *arXiv*. doi: 10.48550/arXiv.2011.02519

Watt-Meyer, O., Brenowitz, N. D., Clark, S. K., Henn, B., Kwa, A., McGibbon, J., ... Bretherton, C. S. (2021). Correcting weather and climate models by machine learning nudged historical simulations. *Geophys. Res. Lett.*, *48*, e2021GL092555. doi: 10.1029/2021GL092555

Weyn, J. A., Durran, D. R., & Caruana, R. (2020). Improving data-driven global weather prediction using deep convolutional neural networks on a cubed sphere. *Journal of Advances in Modeling Earth Systems*, *12*, e2020MS002109. doi: 10.1029/2020MS002109

Weyn, J. A., Durran, D. R., Caruana, R., & Cresswell-Clay, N. (2021). Sub-seasonal forecasting with a large ensemble of deep-learning weather prediction models. *Journal of Advances in Modeling Earth Systems*, *13*, e2021MS002502. doi: 10.1029/2021MS002502

Wikipedia. (2023). *Reservoir computing.* Retrieved from `https://en.wikipedia.org/wiki/Reservoir_computing`

Wikner, A., Pathak, J., Hunt, B., Girvan, M., Arcomano, T., Szunyogh, I., ... Ott, E. (2020). Combining machine learning with knowledge-based modeling for scalable forecasting and subgrid-scale closure of large, complex, spatiotemporal systems. *Chaos*, *30*, 053111. doi: 10.1063/5.0005541