Geophysical Research Letters

RESEARCH LETTER
10.1029/2023GL106060

Internal Variability Increased Arctic Amplification During 1980–2022

Aodhan J. Sweeney[1], Qiang Fu[1], Stephen Po-Chedley[2], Hailong Wang[3], and Muyin Wang[4,5]

[1]Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA, [2]Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, Livermore, CA, USA, [3]Atmospheric Sciences and Global Change Division, PNNL, Richland, WA, USA, [4]Cooperative Institute for Climate, Ocean, and Ecosystem Studies, University of Washington, Seattle, WA, USA, [5]Pacific Marine Environmental Laboratory, Oceanic and Atmospheric Research, NOAA, Silver Spring, MD, USA

Key Points:

- Internally generated and externally forced temperature trends over the Arctic and globe can be partitioned using machine learning methods
- Internal variability has enhanced Arctic warming while damping global warming over 1980-2022
- Accounting for internal variability in observations reconciles discrepancies between simulated and observed Arctic Amplification

Supporting Information:

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

Correspondence to:

A. J. Sweeney and Q. Fu,
aodhan@uw.edu;
qfu@uw.edu

Citation:

Sweeney, A. J., Fu, Q., Po-Chedley, S., Wang, H., & Wang, M. (2023). Internal variability increased Arctic amplification during 1980–2022. *Geophysical Research Letters*, *50*, e2023GL106060. https://doi.org/10.1029/2023GL106060

Received 26 AUG 2023
Accepted 2 DEC 2023
Corrected 8 JAN 2024

This article was corrected on 8 JAN 2024. See the end of the full text for details.

Author Contributions:

Conceptualization: Qiang Fu
Formal analysis: Aodhan J. Sweeney
Funding acquisition: Qiang Fu, Hailong Wang
Methodology: Aodhan J. Sweeney, Stephen Po-Chedley
Project Administration: Qiang Fu
Software: Aodhan J. Sweeney
Supervision: Qiang Fu
Visualization: Aodhan J. Sweeney

Abstract Since 1980, the Arctic surface has warmed four times faster than the global mean. Enhanced Arctic warming relative to the global average warming is referred to as Arctic Amplification (AA). While AA is a robust feature in climate change simulations, models rarely reproduce the observed magnitude of AA, leading to concerns that models may not accurately capture the response of the Arctic to greenhouse gas emissions. Here, we use CMIP6 data to train a machine learning algorithm to quantify the influence of internal variability in surface air temperature trends over both the Arctic and global domains. Application of this machine learning algorithm to observations reveals that internal variability increases the Arctic warming but slows global warming in recent decades, inflating AA since 1980 by 38% relative to the externally forced AA. Accounting for the role of internal variability reconciles the discrepancy between simulated and observed AA.

Plain Language Summary The Arctic has been warming four times as quickly as the global mean since 1980. This so-called Arctic Amplification (AA) has unprecedented impacts on Arctic environments and livelihoods. AA is robustly simulated by climate models, but simulations rarely reproduce the observed levels of AA for 1980–2022. This may be due to a model misrepresentation of the Arctic's sensitivity to increasing greenhouse gases. Another possibility is that the large, observed value of AA is inflated by natural fluctuations in the climate system. Here, we use machine learning to quantify the contribution of natural fluctuations to observed AA. We show that natural fluctuations have inflated AA by 38%, and thus reconcile model-observation differences and suggest that the observed large AA over 1980 to present would not persist into the future.

1. Introduction

Manabe and Wetherald (1975) first found that the "warming in higher latitudes is magnified two to three times the overall amount" in response to the CO_2 increase. This phenomenon was later termed as Arctic Amplification (AA), and has been consistently seen in both model simulations and observations (e.g., Rantanen et al., 2022). From 1980 to 2022 observed surface temperatures in the Arctic (defined here as the region poleward of 70°N) have warmed about four times faster than the global mean (Chylek et al., 2022; Rantanen et al., 2022). Climate models reliably simulate an amplified Arctic warming, but the magnitude of simulated AA is consistently lower than in observations (e.g., England et al., 2021; Hahn et al., 2021; Holland & Landrum, 2021). Many physical processes have been proposed to explain the observed and simulated AA, including both local feedbacks (England et al., 2021; Feldl et al., 2020; Goosse et al., 2018; Hahn et al., 2021; Holland & Bitz, 2003; Manabe & Wetherald, 1975; Zhang et al., 2018, 2020, 2021) and remote teleconnections (Baxter et al., 2019), yet the relative contribution of each of these processes is not well known (Previdi et al., 2021). Differences between the observed and simulated AA suggests that current climate models may not correctly capture the response of the Arctic and/or global climate to external forcings (Chylek et al., 2022; Rantanen et al., 2022).

The observed and simulated AA differences might also be partly caused by natural, internal climate variability (Chylek et al., 2022, 2023; Rantanen et al., 2022), given that certain components of the Arctic (e.g., sea ice) exhibit substantial decadal variations due to internal climate variability (Bonan et al., 2021; Deser et al., 2020; Ding et al., 2019; Kay et al., 2011; Olonscheck et al., 2019; Stroeve et al., 2012; Swart et al., 2015; Topál et al., 2020; Wu et al., 2021). Arctic sea ice cover trends are tightly coupled to surface temperature, due to strong

Writing – review & editing: Aodhan J. Sweeney, Qiang Fu, Stephen Po-Chedley, Hailong Wang, Muyin Wang

impacts on albedo and surface heat fluxes (Serreze & Barry, 2011; IPCC Chapter 3; Feldl et al., 2020; Deng & Dai, 2022). Due to this coupling, the large internal variability in sea ice likely manifests as changes in Arctic surface temperature. Decadal atmospheric and oceanic internal variability may also contribute to recent Arctic warming (Kim & Kim, 2017; Proshutinsky et al., 2015). Internal variability has also been implicated in the recent slowdown of global warming in the early 21st century (Guan et al., 2015; Huber & Knutti, 2014; Kosaka & Xie, 2013; Zhang, 2016). However, it is still an open question whether the large differences in AA between model simulations and observations are mainly caused by climate model deficiencies, internal variability, or both (Chylek et al., 2022; Rantanen et al., 2022).

When comparing the model simulations with observations, it is important to account for the effects of internal variability (Deser et al., 2020). In single-model large ensembles, the same model is run with small perturbations in the initial conditions leading to unique realizations of internal variability in each ensemble member. The externally forced signal can be estimated using the ensemble mean and the internal variability associated with each ensemble member can be obtained as the deviations from this mean (Kay et al., 2015). However, this technique cannot be applied to observations because there is only one observational record. To disentangle the effects of external forcing and internal variability on observed changes in climate, previous work has used various spatiotemporal analysis methods (e.g., Deser et al., 2014; Deser et al., 2016; Gong et al., 2019; Guo et al., 2019; Po-Chedley et al., 2021; Po-Chedley et al., 2022; Räisänen, 2021; Smoliak et al., 2010; Smoliak et al., 2015; Wallace et al., 2012; Wills et al., 2020). Here, we build upon previous methods using a machine learning (ML) approach, which is trained to separate the contribution of external forcing and internal variability to surface warming using climate model large ensembles (see Table S1 in Supporting Information S1 for information about the model large ensembles). The model-trained ML algorithm is then applied to observations to estimate the relative influence of external forcing and internal variability on recent (1980–2022) Arctic and global surface temperature changes. We find that internal variability enhanced Arctic warming but damped global warming, resulting in amplified AA in the observed record. We show that accounting for the effects of internal variability on Arctic and global surface warming reconciles differences between observed and model-simulated AA.

2. Data and Methods

The magnitude of AA depends on the southern boundary used to define the Arctic (Davy et al., 2018). In this study, AA is defined as the surface air temperature trend for the region poleward of 70°N divided by the global mean trend from 1980 to 2022. AA is derived from four different observational temperature data sets including the Met Office Hadley Centre/Climatic Research Unit's global surface temperature data set version 5 (HadCRUTv5), Berkeley Earth Land/Ocean Temperature Record (BerkeleyEarth), GISS Surface Temperature Analysis version 4 (GISTv4), and the NOAA Merged Land Ocean Global Surface Temperature Analysis version 5 (NOAAv5) (Hersbach et al., 2020; Lenssen et al., 2019; Morice et al., 2021; Rohde & Hausfather, 2020; Zhang et al., 2019). Figure S1 in Supporting Information S1 indicates that warming is amplified north of 70°N in all four observational data sets.

Following recent work showing that ML methods can effectively isolate internally generated and externally forced trends (Barnes et al., 2019; Connolly et al., 2023; Gordon & Barnes, 2022; Po-Chedley et al., 2022), we create ML algorithms to isolate these trend contributions in observed surface air temperature during the 43-year period from 1980 to 2022 over both the Arctic and globe. To do this, we create a training data set based on 10 CMIP6 models, of which each have at least 10 ensemble members (Table S1 in Supporting Information S1). Aside from the CESM2 large ensemble from the CMIP6 archive, we also include the 50 member CESM2 large ensemble with updated biomass burning aerosol emissions that better represents the historical radiative forcings in the high latitude northern hemisphere (referred to here as CESM2_SBMB) (Fasullo et al., 2022; Rodgers et al., 2021). The target data in our training are the externally forced and internally generated surface air temperature trends averaged over a given region (either Arctic or globe), which are derived as the mean trend and deviation from the mean in each ensemble. These trends are calculated using 43-year periods separated by five years spanning 1900–2047 (i.e., 1900–1942, 1905–1947, …, 1980–2022, …, 2005–2047). CMIP6 and CESM2_SBMB historical runs end in 2014, so we extend these simulations using either SSP3-7.0 or SSP5-8.5 (SSP5-8.5 is used when both are available) (O'Neill et al., 2016) until 2047 for seven of the models in our training data. The remaining four models only have sufficient ensemble members until 2014, and thus only periods from 1900 to 2012 are used to train our ML algorithm for these models. The ML algorithm is trained using 10 models with large ensembles

but with one model leftout (see more details below). We test the results from 1980 to 2022 using the leftout model that is one of seven with extensions beyond 2014 (Table S1 in Supporting Information S1). The observationally derived AAs are compared with the seven large ensembles for 1980–2022, and with all other CMIP6 models with data available over 1980–2022 (OthersAllEM), even though each of them does not have enough ensemble members to properly derive the externally forced AA.

The predictor data (i.e., the input used to estimate the targets) are maps of surface air temperature (SAT) and sea level pressure (SLP) trends. While SAT trend patterns can capture signals related to both internally and externally generated trends (Po-Chedley et al., 2022), SLP trend patterns provide information on the dynamically induced variability, especially over the high latitudes (Deser et al., 2016; Wallace et al., 2012), thus complementing the information provided by SAT trend patterns. Our ML pipeline is thus designed to accept 43- year trend maps of both SAT and SLP and returns the components of the trend averaged over the Arctic or globe due to internal variability and external forcing. All maps of SAT and SLP trends are regridded to a common 2.5° × 2.5° grid. The ML algorithms are trained for the predictions of the Arctic and global cases separately. For the global case, input data are global trend maps of SLP and SAT. For the Arctic case, we only use trend maps poleward of 20°N (Smoliak et al., 2015; Wallace et al., 2012). Patterns of surface temperature changes can impact both regional and global scale warming and can provide information about the relative role of internal variability (Dong et al., 2019, 2020). Outside the tropics SLP can be used as a proxy for the atmospheric circulation (Deser et al., 2014; Smoliak et al., 2010) and has been used to isolate dynamically induced changes in surface temperature in the northern hemisphere (Guan et al., 2015; Wallace et al., 2012). Further, using more than one geophysical variable may help in identifying signals of external forcing (Rader et al., 2022).

We use the convolutional neural networks (CNNs) that are trained separately for the Arctic and global-mean temperature trends (see Text S1 in Supporting Information S1). We validate the skill of the CNN using a leave-one-out cross validation, where the CNN is trained on data from all models except the model we test on (which is one of seven models covering 1900–2047) (see Text S1 in Supporting Information S1). This prohibits the CNN from learning model specific biases. When applying the CNN to the out-of-sample large ensemble, we also apply it to observed SAT and SLP trend patterns to derive the externally and internally generated trends. SAT trends are those of the four observational data sets from 1980 to 2022. SLP trends are from the ERA5, MERRA-2, and JRA-55 reanalysis data sets over the same time period (see Figure S2 in Supporting Information S1 for a comparison of SLP trends between reanalyzes used and the 20th century reanalysis for 1980–2015, showing a good agreement). Because we have four SAT data sets and three SLP data sets, in total we have 12 sets of SAT and SLP trend maps. For each of the seven models that we test on during the cross validation, we get estimates of internally generated and externally forced trends from each of the 12 observational SAT and SLP sets, providing 84 estimates of the internally generated and externally forced trends. The central value is then the mean over all 84 observational predictions and the uncertainty is quantified by taking into account both observational and ML prediction uncertainties (Text S2 in Supporting Information S1).

3. Arctic Amplification in Observations and CMIP6

Figure 1a shows the patterns of local amplification over the northern hemisphere high latitudes from the observational mean and multi-model mean (MMM). Observations show maximum amplification poleward of 70°N and that large extents of the Arctic Ocean have warmed at least four times as quickly as the global mean. Local amplification ratios exceed six in the Barents Sea, consistent with strong reductions in sea ice concentration in the same region (Isaksen et al., 2022; Parkinson, 2022; Screen & Simmonds, 2010). Although the MMM exhibits a similar pattern of local amplification, it substantially underestimates the magnitude as compared to observations (Rantanen et al., 2022; Ye & Messori, 2021).

Figure 1b shows the AA from the four observational data sets and seven large ensembles and OthersAllEM (see Section 2) for 1980–2022. While the forced component of AA ranges from 2.13 to 3.58 across models, individual ensemble members span a much larger range (Figure S3 in Supporting Information S1 shows each large ensembles AA distribution indivisually). Since the forcing is the same for all members of each model large ensemble, the deviations from the forced AA for a given ensemble member is entirely due to internal variability (OthersAllEM is an exception for which the deviations could also be partly due to differences in forced trends). While the magnitude of AA varies across the observations, all show extreme AA compared to the distribution of model simulations (Figure 1b). All observational AA estimates sit outside the range of forced AA predicted by the large

10.1029/2023GL106060

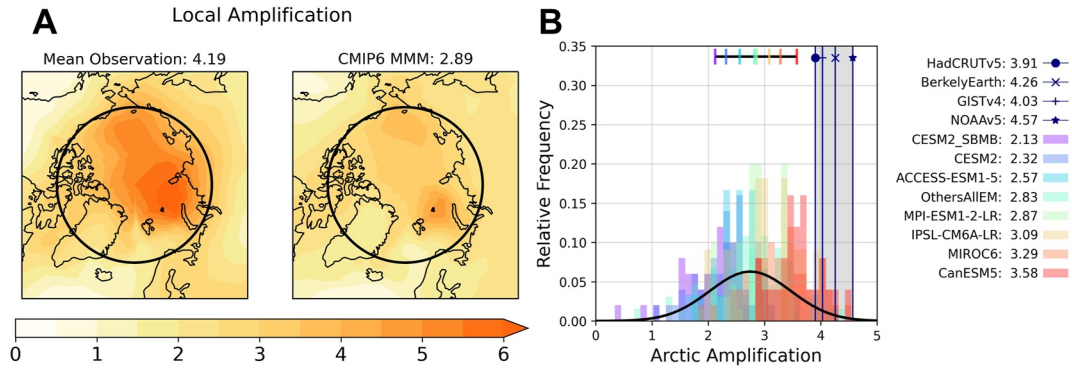

Figure 1. (a) Local amplification (i.e., local surface air temperature trend divided by global mean temperature trend): over the northern high latitudes from the average of observational data sets and the multi-model mean during 1980–2022. The Arctic is the region poleward of 70°N (black circle), and the corresponding Arctic Amplification (AA) (i.e., the Arctic mean temperature trend divided by global mean trend) is provided at the top of each plot. (b) Comparisons of AA in observations and CMIP6 models. Observations are shown using vertical lines, and gray shading shows their range. Histograms show the relative frequency distribution of AA over 1980–2022 for each model, which is normalized by its number of ensemble members. The black curve shows a normal distribution fitted to all model AA values. The black horizontal line shows the range of forced AAs and the vertical tick marks represent the ensemble-mean AA for each model. The values of AA from each observation and forced AA from each model are provided in the legend.

ensembles (i.e., outside the range of the horizontal black bar), and AA from NOAA Global Temperature v5 (4.57) exceeds AA from all model ensemble members. Observationally derived AA is weakest in HadCRUTv5 (3.91), which still exceeds 94% of the simulated AAs in Figure 1b.

The exceptionally high AA in observations compared to model simulations could be due to systematic model biases in the representation of internal variability, biases in the simulated response to external forcing, biases in the prescribed model forcing, or the observed AA being an extremely unlikely event (Rantanen et al., 2022). Because AA is defined as the trend poleward of 70°N divided by the global mean, biases in either the Arctic or global warming would impact the comparison of AA. To investigate how well models simulate global and Arctic warming individually, Figure 2 shows the distribution of Arctic and global warming in observations and model simulations (Figures S4 and S5 in Supporting Information S1 show each large ensemble trend distribution individually over the Arctic and globe, respectively). Simulations of Arctic warming exhibit a very large range of trends from nearly 0 up to 2 K/decade for 1980–2022. Although a significant amount of this spread is due to differences between models (e.g., compare forced trends from CanESM5 to those from all other models), even individual models have Arctic warming trends that vary by ±0.5 K/decade due to internal variability (see, e.g., ACCESS-ESM1-5 in Figure S4 in Supporting Information S1). The observed Arctic warming ranges from 0.742 to 0.835 K/decade from the four observational

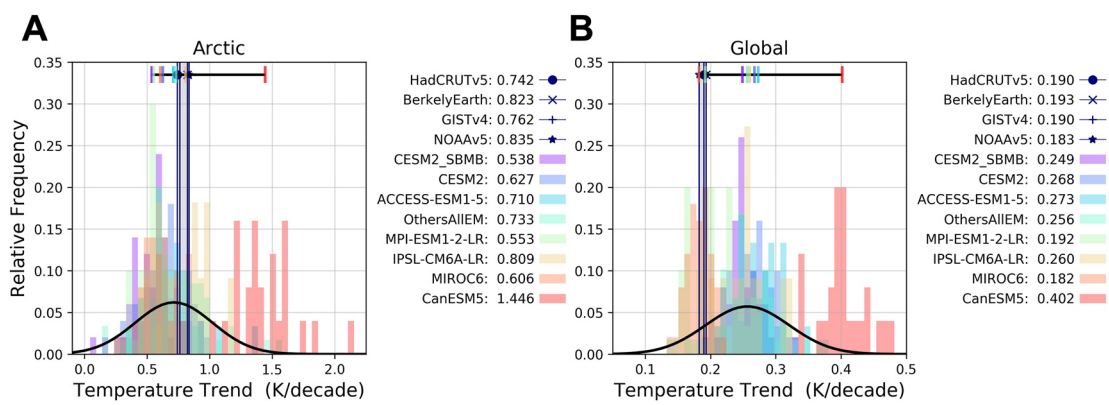

Figure 2. Surface air temperature trends over 1980–2022 for the (a) Arctic and (b) global mean. The histograms show the distributions from model simulations, and the vertical lines represent the observations where gray shading shows their range. The black curve shows a normal distribution fitted to all the simulated temperature trends. The horizontal black line shows the range of externally forced trends with ticks showing individual models' forced trends. The trend values from individual observational data sets and forced trend values from individual models are provided in the legend.

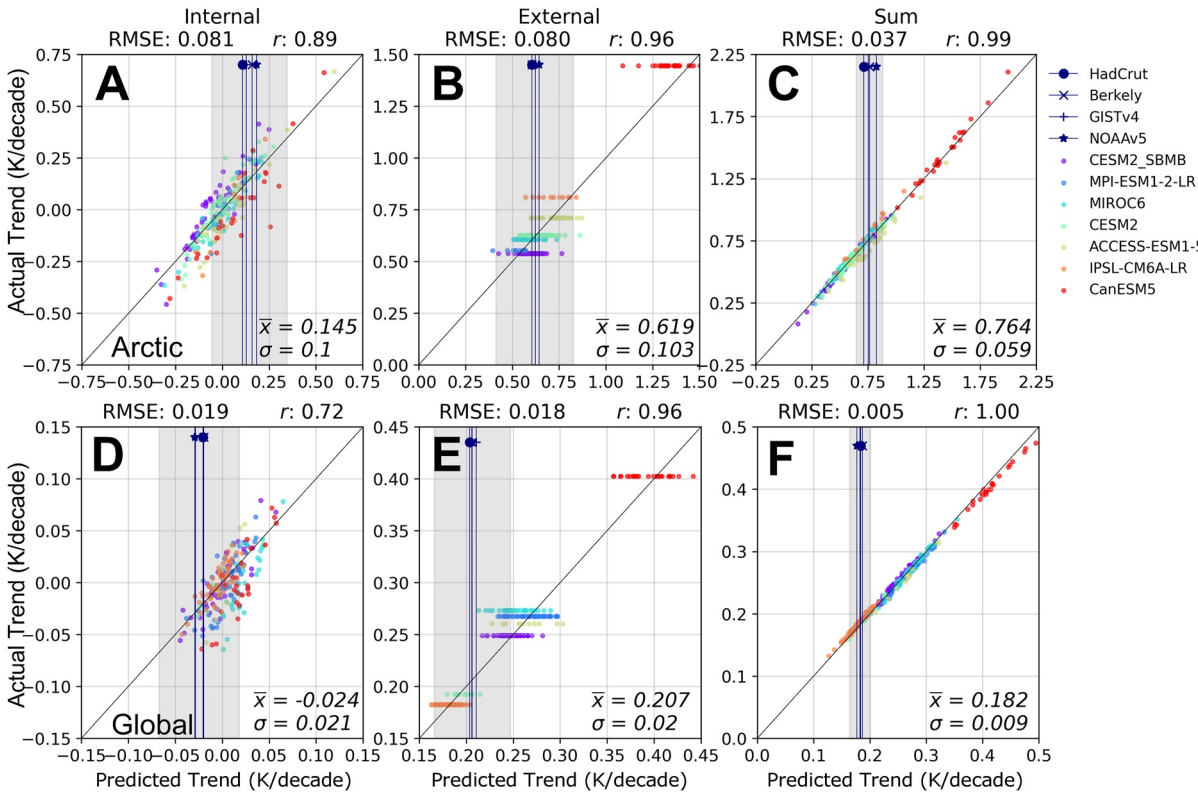

Figure 3. (a)–(c) Arctic and (d)–(f) global surface air temperature trends predicted from the CNN (*x*-axis) versus corresponding actual trends (*y*-axis) over 1980–2022. The root mean squared error (RMSE) and correlation coefficient (r) are shown at the top of each plot. Panels (a), (d) shows results for internally generated trends, (b), (e) shows the externally forced trends, and (c), (f) shows the sum of the internally generated and externally forced trends. The vertical lines show the mean observational estimate for each temperature record, and the gray shading shows the ±2σ uncertainty of mean prediction. The mean ($\bar{x}$) and its standard deviation (σ) based on observations are provided in the bottom right of each plot. The black diagonal line in (a)–(f) is the 1:1 line.

data sets, with a mean of 0.791 K/decade, which are all well within the range of Arctic warming predicted by models (Figure 2a). Thus, the observed Arctic warming is not as extreme as AA when compared to model simulations.

The global warming trend from the four observational data sets ranges from 0.183 to 0.193 K/decade with a mean of 0.189 K/decade, which is on the lower side of the simulated range of externally forced trends (Figure 2b). However, some models have ensemble members that simulate global warming trends below what is observed, suggesting that internal variability may damp the rate of global warming (Kosaka & Xie, 2013; Watanabe et al., 2014; Wu et al., 2019; Xie & Kosaka, 2017; Zhang, 2016). Next, we attempt to partition the observed Arctic and global warming trends into their externally and internally generated components.

4. Separating Internal Generated and Externally Forced Trends in Observations

The test of the CNN algorithms on each of the seven models for 1980–2022 are shown as scatter points in Figure 3, which suggest that when presented with a set of SAT and SLP trend maps from a model ensemble not used during training, the CNN can reliably separate the internal and external contributions to the trends averaged over the Arctic and globe. This is despite the wide range of internally generated and externally forced trends simulated by models (see Figure 2). The skill of the CNN results from its ability to learn the patterns (in SAT and SLP) that correspond with the internally generated and external forced trends in both the Arctic and global domains. The CNN also generalizes well to simulations with forced trends far from the MMM (e.g., red dots showing results for CanESM5 in Figures 3b and 3e). Although the CNN predicts the internal and external trends separately, their sum accurately reproduces the total trend (see Figures 3c and 3f). This conservation of the total trend is not explicitly targeted during training but arises from learning this closure in the training data.

Having shown that the CNNs can reliably predict the internal and external trends in models, we apply the CNNs to observations from 1980 to 2022 using the four SAT data sets and three SLP data sets. The mean results for each

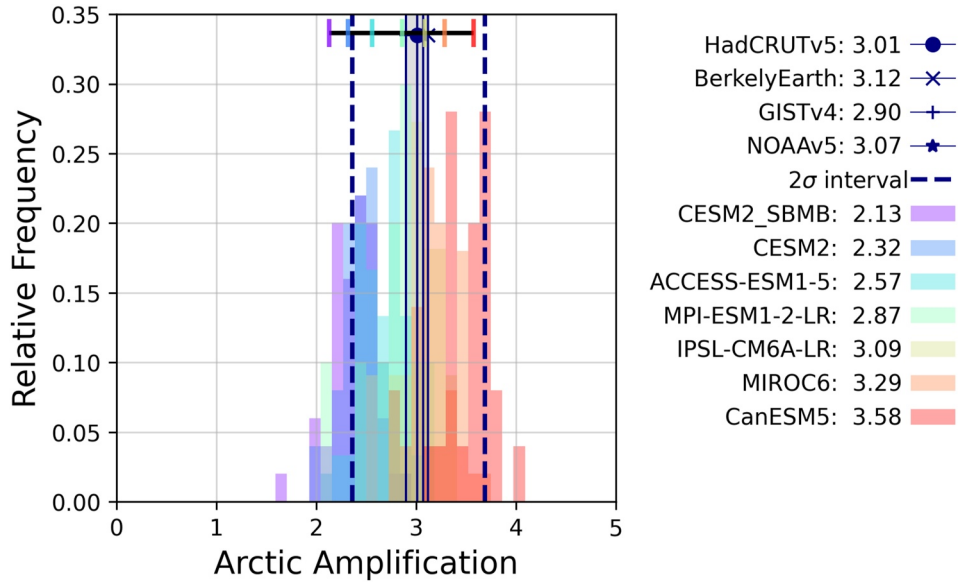

Figure 4. Same as Figure 1(b) but showing the AA estimates after subtracting the contributions of internal variability, as derived from the ML algorithms, from Arctic and global warming. This was done for both observations and each ensemble member of the seven large ensembles. The vertical blue dashed lines show the 2σ range of the estimated forced AA based on observations.

SAT data set are shown by the vertical lines in Figure 3 with the 2σ confidence interval (Text S2 in Supporting Information S1) for the mean of all observational data sets. Predictions based on observational data sets indicate that internal variability has enhanced Arctic surface warming over 1980–2022 by 0.145 K/decade (Figure 3a). The CNN predicts that the externally generated Arctic surface temperature trend is 0.619 K/decade. This suggests that internal variability has accelerated the pace of Arctic warming by ∼23% relative to the forced trend. Using all ensembles from the seven models for all 43-year periods separated by 5-year increment over 1900–2047/2012, the 2σ spread of Arctic internal variability is ±0.324 K/decade. Many studies have shown that surface temperature trends in the Arctic are strongly coupled to sea ice trends, and that recent declines in sea ice cover have been enhanced by multidecadal variability (Deng & Dai, 2022; Ding et al., 2019; Kay et al., 2011; Screen & Simmonds, 2010; Serreze et al., 2009). Our results agree with previous studies showing that internal variability is an important contribution to recent trends in Arctic climate change (Bonan & Blanchard-Wrigglesworth, 2020; Chylek et al., 2022; Ding et al., 2019).

Application of the CNN to the global case suggests that internal variability dampens the observed temperature trend, which is also consistent with previous studies (Kosaka & Xie, 2013; Po-Chedley et al., 2022; Xie & Kosaka, 2017). All observational estimates show that internal variability reduces global surface warming over 1980–2022, with a central estimate of −0.024 K/decade (Figure 3d). The global CNN predicts the externally generated trend to be 0.207 K/decade. This suggests that internal variability has damped the global warming by ∼12% relative to the forced trend since 1980. Although this internal variability is substantial, the 2σ spread of internal variability from all large ensembles over the 1900–2047/2012 period is ±0.051 K/decade.

5. Implications for Arctic Amplification and Discussions

Internal variability can impact AA through its effect on Arctic warming, global warming, or both. ML algorithms applied here can partition the contribution of externally forced and internally generated trends both over the Arctic and over the globe. Application of these algorithms to observations suggests that internal variability has enhanced Arctic surface warming (+0.145 K/decade) while simultaneously dampening global mean surface warming (−0.024 K/decade) over 1980–2022 (Figures 3a and 3d). Because AA is the surface temperature trend in the Arctic divided by the global mean trend, the opposing role of internal variability in the Arctic and global average inflates observed AA. Figure 4 is the same as Figure 1b but with AA estimates after we first subtract the contribution of internal variability derived from ML algorithms from both the Arctic and global mean trends and

then recalculate their ratio. This was done for both observations and each ensemble member of the seven large ensembles (Figure S6 in Supporting Information S1 shows each large ensembles distribution of AA after removing internal variability indivisually). Upon removing the estimated effect of internal variability from the Arctic and global mean surface air temperature trend, AA from climate model simulations and observational data sets exhibit excellent agreement (cf., Figures 4–1b).

After subtracting the internally generated trend from the mean observational trend over the Arctic (0.791 K/decade) and globe (0.189 K/decade), the externally generated trend is estimated as 0.646 K/decade and 0.213 K/decade, respectively, meaning that the externally forced AA is 3.03. A similar result is obtained by using the externally forced Arctic to global warming trends directly estimated by the CNN, which are 0.619 K/decade and 0.207 K/decade (Figures 3b and 3e), and the resulting externally forced AA is 2.99. Our results shown here suggest that internal variability plays a substantial role in inflating recent AA and increased the 1980–2022 AA by 38%. Key to this result, is recognizing that internal variability has enhanced Arctic warming while simultaneously damping global warming. Vertical blue dashed lines show the 2σ spread of externally forced AA based on observations (see Text S2 in Supporting Information S1). Figure 4 shows that the estimates of observed, externally forced AA is still within the range of forced AA based on model simulations even when this uncertainty is included. Although here we present results using a definition of the Arctic as poleward of 70°N, repeating the analysis by defining the Arctic as poleward of 60°N produces similar results (see Figure S7 in Supporting Information S1). We also compare the results of the CNNs used in this study (Text S1 in Supporting Information S1) to Partial Least Squares regression (PLS), a linear pattern matching algorithm that has been shown to have skill in a similar context (Po-Chedley et al., 2022). We chose to use CNNs because they better minimize the RMSE (Figure S8 in Supporting Information S1), but results are similar using either CNNs or PLS methods (see Figures S9 and S10 in Supporting Information S1). The mean AA ratio after removing internal variability contributions to observed trends based on PLS regression and the CNN is 2.98 and 3.03, respectively.

Although we stress internal variability's role in inflating recent AA, these results do not discount the possible influence of forcing on the simulated-versus-observed differences in AA. Systematic biases in the forcing prescription can have a significant impact on simulated AA during 1980–2022. For example, changes in the amount of biomass burning prescribed in CESM2_SBMB compared to CESM2 lead to decreased surface warming in the Northern Hemisphere high latitudes and thus a smaller AA ratio in CESM2_SBMB (Figure 1b). Because AA is defined as the ratio of the total Arctic and global warming, a forcing bias in either region will impact the magnitude of AA even if internally generated trends match observations. Given that the externally forced and internally generated trends estimated from observations are within the bounds of the simulated externally forced and internally generated trends in the large ensembles, a pertinent question is why don't more models simulate the observed levels of AA? Crucial to reproducing the observed AA is simulating internal variability that enhances Arctic warming while simultaneously dampening global warming. The fact that model simulations generally do not reproduce the observed levels of AA may suggest that while models during the 1980–2022 period can simulate the observed amplitude of internal variability in the Arctic and over the globe separately, they struggle to simulate the combined manifestation of internal variability that enhances Arctic warming while suppressing global warming (Rantanen et al., 2022; Rosenblum & Eisenman, 2017). While prior research links multidecadal Arctic and Pacific variability (Bonan & Blanchard-Wrigglesworth, 2020; Screen & Deser, 2019), the causal link between internal variability in these regions and its portrayal in climate models remains an ongoing research area (Baxter et al., 2019). Our results show that considering internal variability can reconcile the discrepancy between observed and simulated AA but also calls for the need to better understand this unusual manifestation of internal variability.

Data Availability Statement

The processed data on which this article is based are available in Sweeney et al. (2023). Software required to recreate the machine learning results is provided in Sweeney (2023).

Acknowledgments
This research was supported by the U.S. Department of Energy (DOE), Office of Science, Office of Biological and Environmental Research, Regional and Global Model Analysis (RGMA) program area, as part of the HiLAT-RASM project. This research was also supported by the NASA FINESST Grant 80NSSC22K1438 and NSF Grant AGS-2202812. Additional funding was provided by the Calvin Professorship in Atmospheric Sciences. S.P.-C was supported through the PCMDI Project, which is funded by the RGMA program area of the Office of Science at DOE. M. Wang is funded with support of the Arctic Research Program of the NOAA Global Ocean Monitoring and Observing (GOMO) office through the Cooperative Institute for Climate, Ocean, and Ecosystem Studies (CICOES) under NOAA Cooperative Agreement NA20OAR4320271, Contribution No 2023-1295, and Pacific Marine Environmental Laboratory Contribution No 5531. Research at Lawrence Livermore National Laboratory was performed under the auspices of U.S. DOE Contract DE-AC52-07NA27344. The Pacific Northwest National Laboratory (PNNL) is operated for DOE by Battelle Memorial Institute under contract DE-AC05-76RLO1830. We would like to acknowledge high-performance computing support from Cheyenne (https://doi.org/10.5065/D6RX99HX) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation, for the analyses presented in this study and for data management, storage, and preservation.

References

Barnes, E. A., Hurrell, J. W., Ebert-Uphoff, I., Anderson, C., & Anderson, D. (2019). Viewing forced climate patterns through an AI lens. *Geophysical Research Letters*, *46*(22), 13389–13398. https://doi.org/10.1029/2019GL084944

Baxter, I., Ding, Q., Schweiger, A., L'Heureux, M., Baxter, S., Wang, T., et al. (2019). How tropical pacific surface cooling contributed to accelerated sea ice melt from 2007 to 2012 as ice is thinned by anthropogenic forcing. *Journal of Climate*, *32*(24), 8583–8602. https://doi.org/10.1175/JCLI-D-18-0783.1

Bonan, D. B., & Blanchard-Wrigglesworth, E. (2020). Nonstationary teleconnection between the Pacific Ocean and Arctic sea ice. *Geophysical Research Letters*, *47*(2), e2019GL085666. https://doi.org/10.1029/2019GL085666

Bonan, D. B., Lehner, F., & Holland, M. M. (2021). Partitioning uncertainty in projections of Arctic sea ice. *Environmental Research Letters*, *16*(4), 044002. https://doi.org/10.1088/1748-9326/abe0ec

Chylek, P., Folland, C., Klett, J. D., Wang, M., Hengartner, N., Lesins, G., & Dubey, M. K. (2022). Annual mean Arctic amplification 1970–2020: Observed and simulated by CMIP6 climate models. *Geophysical Research Letters*, *49*(13), e2022GL099371. https://doi.org/10.1029/2022GL099371

Chylek, P., Folland, C. K., Klett, J. D., Wang, M., Lesins, G., & Dubey, M. K. (2023). High values of the Arctic Amplification in the early decades of the 21st century: Causes of discrepancy by CMIP6 models between observation and simulation. *Journal of Geophysical Research: Atmospheres*, *128*(23), e2023JD039269. https://doi.org/10.1029/2023JD039269

Connolly, C., Barnes, E. A., Hassanzadeh, P., & Pritchard, M. (2023). Using neural networks to learn the jet stream forced response from natural variability. *Artificial Intelligence for the Earth Systems*, *2*(2). https://doi.org/10.1175/AIES-D-22-0094.1

Davy, R., Chen, L., & Hanna, E. (2018). Arctic amplification metrics. *International Journal of Climatology*, *38*(12), 4384–4394. https://doi.org/10.1002/joc.5675

Deng, J., & Dai, A. (2022). Sea ice–air interactions amplify multidecadal variability in the North Atlantic and Arctic region. *Nature Communications*, *13*(1), 2100. https://doi.org/10.1038/s41467-022-29810-7

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

Deser, C., Phillips, A. S., Alexander, M. A., & Smoliak, B. V. (2014). Projecting North American climate over the next 50 years: Uncertainty due to internal variability. *Journal of Climate*, *27*(6), 2271–2296. https://doi.org/10.1175/JCLI-D-13-00451.1

Deser, C., Terray, L., & Phillips, A. S. (2016). Forced and internal components of winter air temperature trends over North America during the past 50 Years: Mechanisms and implications. *Journal of Climate*, *29*(6), 2237–2258. https://doi.org/10.1175/JCLI-D-15-0304.1

Ding, Q., Schweiger, A., L'Heureux, M., Steig, E. J., Battisti, D. S., Johnson, N. C., et al. (2019). Fingerprints of internal drivers of Arctic sea ice loss in observations and model simulations. *Nature Geoscience*, *12*(1), 28–33. https://doi.org/10.1038/s41561-018-0256-8

Dong, Y., Armour, K. C., Zelinka, M. D., Proistosescu, C., Battisti, D. S., Zhou, C., & Andrews, T. (2020). Intermodel spread in the pattern effect and its contribution to climate sensitivity in CMIP5 and CMIP6 models. *Journal of Climate*, *33*(18), 7755–7775. https://doi.org/10.1175/JCLI-D-19-1011.1

Dong, Y., Proistosescu, C., Armour, K. C., & Battisti, D. S. (2019). Attributing historical and future evolution of radiative feedbacks to regional warming patterns using a Green's function approach: The preeminence of the western Pacific. *Journal of Climate*, *32*(17), 5471–5491. https://doi.org/10.1175/JCLI-D-18-0843.1

England, M. R., Eisenman, I., Lutsko, N. J., & Wagner, T. J. W. (2021). The recent emergence of arctic amplification. *Geophysical Research Letters*, *48*(15), e2021GL094086. https://doi.org/10.1029/2021GL094086

Fasullo, J. T., Lamarque, J.-F., Hannay, C., Rosenbloom, N., Tilmes, S., DeRepentigny, P., et al. (2022). Spurious late historical-era warming in CESM2 driven by prescribed biomass burning emissions. *Geophysical Research Letters*, *49*(2), e2021GL097420. https://doi.org/10.1029/2021GL097420

Feldl, N., Po-Chedley, S., Singh, H. K. A., Hay, S., & Kushner, P. J. (2020). Sea ice and atmospheric circulation shape the high-latitude lapse rate feedback. *Npj Climate and Atmospheric Science*, *3*(1), 1–9. https://doi.org/10.1038/s41612-020-00146-7

Gong, H., Wang, L., Chen, W., & Wu, R. (2019). Attribution of the east Asian winter temperature trends during 1979–2018: Role of external forcing and internal variability. *Geophysical Research Letters*, *46*(19), 10874–10881. https://doi.org/10.1029/2019GL084154

Goosse, H., Kay, J. E., Armour, K. C., Bodas-Salcedo, A., Chepfer, H., Docquier, D., et al. (2018). Quantifying climate feedbacks in polar regions. *Nature Communications*, *9*(1), 1919. https://doi.org/10.1038/s41467-018-04173-0

Gordon, E. M., & Barnes, E. A. (2022). Incorporating uncertainty into a regression neural network enables identification of decadal state-dependent predictability in CESM2. *Geophysical Research Letters*, *49*(15), e2022GL098635. https://doi.org/10.1029/2022GL098635

Guan, X., Huang, J., Guo, R., & Lin, P. (2015). The role of dynamically induced variability in the recent warming trend slowdown over the Northern Hemisphere. *Scientific Reports*, *5*(1), 12669. https://doi.org/10.1038/srep12669

Guo, R., Deser, C., Terray, L., & Lehner, F. (2019). Human influence on winter precipitation trends (1921–2015) over North America and Eurasia revealed by dynamical adjustment. *Geophysical Research Letters*, *46*(6), 3426–3434. https://doi.org/10.1029/2018GL081316

Hahn, L. C., Armour, K. C., Zelinka, M. D., Bitz, C. M., & Donohoe, A. (2021). Contributions to polar amplification in CMIP5 and CMIP6 Models. *Frontiers in Earth Science*, *9*. https://doi.org/10.3389/feart.2021.710036

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Holland, M. M., & Bitz, C. M. (2003). Polar amplification of climate change in coupled models. *Climate Dynamics*, *21*(3), 221–232. https://doi.org/10.1007/s00382-003-0332-6

Holland, M. M., & Landrum, L. (2021). The emergence and transient nature of Arctic amplification in coupled climate models. *Frontiers in Earth Science*, *9*. https://doi.org/10.3389/feart.2021.719024

Huber, M., & Knutti, R. (2014). Natural variability, radiative forcing and climate response in the recent hiatus reconciled. *Nature Geoscience*, *7*(9), 651–656. https://doi.org/10.1038/ngeo2228

Isaksen, K., Nordli, Ø., Ivanov, B., Køltzow, M. A. Ø., Aaboe, S., Gjelten, H. M., et al. (2022). Exceptional warming over the Barents area. *Scientific Reports*, *12*(1), 9371. https://doi.org/10.1038/s41598-022-13568-5

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., et al. (2015). The community Earth system model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability. *Bulletin of the American Meteorological Society*, *96*(8), 1333–1349. https://doi.org/10.1175/BAMS-D-13-00255.1

Kay, J. E., Holland, M. M., & Jahn, A. (2011). Inter-annual to multi-decadal Arctic sea ice extent trends in a warming world. *Geophysical Research Letters*, *38*(15), L15708. https://doi.org/10.1029/2011GL048008

Kim, H.-M., & Kim, B.-M. (2017). Relative contributions of atmospheric energy transport and sea ice loss to the recent warm arctic winter. *Journal of Climate*, *30*(18), 7441–7450. https://doi.org/10.1175/JCLI-D-17-0157.1

Kosaka, Y., & Xie, S.-P. (2013). Recent global-warming hiatus tied to equatorial Pacific surface cooling. *Nature*, *501*(7467), 403–407. https://doi.org/10.1038/nature12534

Lenssen, N. J. L., Schmidt, G. A., Hansen, J. E., Menne, M. J., Persin, A., Ruedy, R., & Zyss, D. (2019). Improvements in the GISTEMP uncertainty model. *Journal of Geophysical Research: Atmospheres*, *124*(12), 6307–6326. https://doi.org/10.1029/2018JD029522

Manabe, S., & Wetherald, R. T. (1975). The effects of doubling the CO_2 concentration on the climate of a general circulation model. *Journal of the Atmospheric Sciences*, *32*(1), 3–15. https://doi.org/10.1175/1520-0469(1975)032<0003:TEODTC>2.0.CO;2

Morice, C. P., Kennedy, J. J., Rayner, N. A., Winn, J. P., Hogan, E., Killick, R. E., et al. (2021). An updated assessment of near-surface temperature change from 1850: The HadCRUT5 data set. *Journal of Geophysical Research: Atmospheres*, *126*(3), e2019JD032361. https://doi.org/10.1029/2019JD032361

Olonscheck, D., Mauritsen, T., & Notz, D. (2019). Arctic sea-ice variability is primarily driven by atmospheric temperature fluctuations. *Nature Geoscience*, *12*(6), 430–434. https://doi.org/10.1038/s41561-019-0363-1

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., et al. (2016). The scenario model intercomparison project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, *9*(9), 3461–3482. https://doi.org/10.5194/gmd-9-3461-2016

Parkinson, C. L. (2022). Arctic sea ice coverage from 43 years of satellite passive-microwave observations. *Frontiers in Remote Sensing*, *3*. https://doi.org/10.3389/frsen.2022.1021781

Po-Chedley, S., Fasullo, J. T., Siler, N., Labe, Z. M., Barnes, E. A., Bonfils, C. J. W., & Santer, B. D. (2022). Internal variability and forcing influence model–satellite differences in the rate of tropical tropospheric warming. *Proceedings of the National Academy of Sciences*, *119*(47), e2209431119. https://doi.org/10.1073/pnas.2209431119

Po-Chedley, S., Santer, B. D., Fueglistaler, S., Zelinka, M. D., Cameron-Smith, P. J., Painter, J. F., & Fu, Q. (2021). Natural variability contributes to model–satellite differences in tropical tropospheric warming. *Proceedings of the National Academy of Sciences*, *118*(13), e2020962118. https://doi.org/10.1073/pnas.2020962118

Previdi, M., Smith, K. L., & Polvani, L. M. (2021). Arctic amplification of climate change: A review of underlying mechanisms. *Environmental Research Letters*, *16*(9), 093003. https://doi.org/10.1088/1748-9326/ac1c29

Proshutinsky, A., Dukhovskoy, D., Timmermans, M.-L., Krishfield, R., & Bamber, J. L. (2015). Arctic circulation regimes. *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*, *373*(2052), 20140160. https://doi.org/10.1098/rsta.2014.0160

Rader, J. K., Barnes, E. A., Ebert-Uphoff, I., & Anderson, C. (2022). Detection of forced change within combined climate fields using explainable neural networks. *Journal of Advances in Modeling Earth Systems*, *14*(7), e2021MS002941. https://doi.org/10.1029/2021MS002941

Räisänen, J. (2021). Effect of atmospheric circulation on surface air temperature trends in years 1979–2018. *Climate Dynamics*, *56*(7), 2303–2320. https://doi.org/10.1007/s00382-020-05590-y

Rantanen, M., Karpechko, A. Y., Lipponen, A., Nordling, K., Hyvärinen, O., Ruosteenoja, K., et al. (2022). The Arctic has warmed nearly four times faster than the globe since 1979. *Communications Earth & Environment*, *3*(1), 1–10. https://doi.org/10.1038/s43247-022-00498-3

Rodgers, K. B., Lee, S.-S., Rosenbloom, N., Timmermann, A., Danabasoglu, G., Deser, C., et al. (2021). Ubiquity of human-induced changes in climate variability. *Earth System Dynamics*, *12*(4), 1393–1411. https://doi.org/10.5194/esd-12-1393-2021

Rohde, R. A., & Hausfather, Z. (2020). The Berkeley Earth land/ocean temperature record. *Earth System Science Data*, *12*(4), 3469–3479. https://doi.org/10.5194/essd-12-3469-2020

Rosenblum, E., & Eisenman, I. (2017). Sea ice trends in climate models only accurate in runs with biased global warming. *Journal of Climate*, *30*(16), 6265–6278. https://doi.org/10.1175/JCLI-D-16-0455.1

Screen, J. A., & Deser, C. (2019). Pacific Ocean variability influences the time of emergence of a seasonally ice-free Arctic Ocean. *Geophysical Research Letters*, *46*(4), 2222–2231. https://doi.org/10.1029/2018GL081393

Screen, J. A., & Simmonds, I. (2010). The central role of diminishing sea ice in recent Arctic temperature amplification. *Nature*, *464*(7293), 1334–1337. https://doi.org/10.1038/nature09051

Serreze, M. C., Barrett, A. P., Stroeve, J. C., Kindig, D. N., & Holland, M. M. (2009). The emergence of surface-based Arctic amplification. *The Cryosphere*, *3*(1), 11–19. https://doi.org/10.5194/tc-3-11-2009

Serreze, M. C., & Barry, R. G. (2011). Processes and impacts of Arctic amplification: A research synthesis. *Global and Planetary Change*, *77*(1), 85–96. https://doi.org/10.1016/j.gloplacha.2011.03.004

Smoliak, B. V., Wallace, J. M., Lin, P., & Fu, Q. (2015). Dynamical adjustment of the Northern Hemisphere surface air temperature field: Methodology and application to observations. *Journal of Climate*, *28*(4), 1613–1629. https://doi.org/10.1175/JCLI-D-14-00111.1

Smoliak, B. V., Wallace, J. M., Stoelinga, M. T., & Mitchell, T. P. (2010). Application of partial least squares regression to the diagnosis of year-to-year variations in Pacific Northwest snowpack and Atlantic hurricanes. *Geophysical Research Letters*, *37*(3), L03801. https://doi.org/10.1029/2009GL041478

Stroeve, J. C., Kattsov, V., Barrett, A., Serreze, M., Pavlova, T., Holland, M., & Meier, W. N. (2012). Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations. *Geophysical Research Letters*, *39*(16), L16502. https://doi.org/10.1029/2012GL052676

Swart, N. C., Fyfe, J. C., Hawkins, E., Kay, J. E., & Jahn, A. (2015). Influence of internal variability on Arctic sea-ice trends. *Nature Climate Change*, *5*(2), 86–89. https://doi.org/10.1038/nclimate2483

Sweeney, A. (2023). 2023 release version (1.0.0) AodhanSweeney/AA_InternalExternalPartitioning [Software]. GitHub. https://github.com/AodhanSweeney/AA_InternalExternalPartitioning

Sweeney, A., Fu, Q., Po-Chedley, S., Wang, H., & Wang, M. (2023). Internal variability increased Arctic amplification during 1980-2022 [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.8286589

Topál, D., Ding, Q., Mitchell, J., Baxter, I., Herein, M., Haszpra, T., et al. (2020). An internal atmospheric process determining summertime Arctic Sea Ice melting in the next three decades: Lessons learned from five large ensembles and multiple CMIP5 climate simulations. *Journal of Climate*, *33*(17), 7431–7454. https://doi.org/10.1175/JCLI-D-19-0803.1

Wallace, J. M., Fu, Q., Smoliak, B. V., Lin, P., & Johanson, C. M. (2012). Simulated versus observed patterns of warming over the extratropical Northern Hemisphere continents during the cold season. *Proceedings of the National Academy of Sciences*, *109*(36), 14337–14342. https://doi.org/10.1073/pnas.1204875109

Watanabe, M., Shiogama, H., Tatebe, H., Hayashi, M., Ishii, M., & Kimoto, M. (2014). Contribution of natural decadal variability to global warming acceleration and hiatus. *Nature Climate Change*, *4*(10), 893–897. https://doi.org/10.1038/nclimate2355

Wills, R. C. J., Battisti, D. S., Armour, K. C., Schneider, T., & Deser, C. (2020). Pattern recognition methods to separate forced responses from internal variability in climate model ensembles and observations. *Journal of Climate*, *33*(20), 8693–8719. https://doi.org/10.1175/JCLI-D-19-0855.1

Wu, F., Li, W., Zhang, P., & Li, W. (2021). Relative contributions of internal atmospheric variability and surface processes to the interannual variations in wintertime arctic surface air temperatures. *Journal of Climate*, *34*(17), 7131–7148. https://doi.org/10.1175/JCLI-D-20-0779.1

Wu, T., Hu, A., Gao, F., Zhang, J., & Meehl, G. A. (2019). New insights into natural variability and anthropogenic forcing of global/regional climate evolution. *Npj Climate and Atmospheric Science*, *2*(1), 1–13. https://doi.org/10.1038/s41612-019-0075-7

Xie, S.-P., & Kosaka, Y. (2017). What caused the global surface warming hiatus of 1998–2013? *Current Climate Change Reports*, *3*(2), 128–140. https://doi.org/10.1007/s40641-017-0063-0

Ye, K., & Messori, G. (2021). Inter-model spread in the wintertime Arctic amplification in the CMIP6 models and the important role of internal climate variability. *Global and Planetary Change*, *204*, 103543. https://doi.org/10.1016/j.gloplacha.2021.103543

Zhang, J. H. L., Huang, B., Menne, M. J., Yin, X., Sánchez-Lugo, A., Gleason, B. E., et al. (2019). Updated temperature data give a sharper view of climate trends. Retrieved from http://eos.org/science-updates/updated-temperature-data-give-a-sharper-view-of-climate-trends

Zhang, L. (2016). The roles of external forcing and natural variability in global warming hiatuses. *Climate Dynamics*, *47*(9), 3157–3169. https://doi.org/10.1007/s00382-016-3018-6

Zhang, R., Wang, H., Fu, Q., Pendergrass, A. G., Wang, M., Yang, Y., et al. (2018). Local radiative feedbacks over the arctic based on observed short-term climate variations. *Geophysical Research Letters*, *45*(11), 5761–5770. https://doi.org/10.1029/2018GL077852

Zhang, R., Wang, H., Fu, Q., & Rasch, P. J. (2020). Assessing global and local radiative feedbacks based on AGCM simulations for 1980–2014/2017. *Geophysical Research Letters*, *47*(12), e2020GL088063. https://doi.org/10.1029/2020GL088063

Zhang, R., Wang, H., Fu, Q., Rasch, P. J., Wu, M., & Maslowski, W. (2021). Understanding the cold season arctic surface warming trend in recent decades. *Geophysical Research Letters*, *48*(19), e2021GL094878. https://doi.org/10.1029/2021GL094878

References From the Supporting Information

Chattopadhyay, A., Hassanzadeh, P., & Pasha, S. (2020). Predicting clustered weather patterns: A test case for applications of convolutional neural networks to spatio-temporal climate data. *Scientific Reports*, *10*(1), 1317. https://doi.org/10.1038/s41598-020-57897-9

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*(7), e2021MS002502. https://doi.org/10.1029/2021ms002502

Yu, X., Millet, D. B., Henze, D. K., Turner, A. J., Delgado, A. L., Bloom, A. A., & Sheng, J. (2023). A high-resolution satellite-based map of global methane emissions reveals missing wetland, fossil fuel, and monsoon sources. *Atmospheric Chemistry and Physics*, *23*(5), 3325–3346. https://doi.org/10.5194/acp-23-3325-2023

Erratum

In the originally published version of this article, the Author Contributions list contained typographical errors. Formal analysis, methodology, software, visualization and writing – review & editing should be attributed to Aodhan Sweeney. Conceptualization, funding acquisition, project administration, supervision and writing – review & editing should be attributed to Qiang Fu. Methodology and writing – review & editing should be attributed to Stephen Po-Chedley. Funding acquisition and writing – review & editing should be attributed to Hailong Wang. Writing – review & editing should be attributed to Muyin Wang. The errors have been corrected, and this may be considered the authoritative version of record.

Geophysical Research Letters

Supporting Information for

Internal Variability Increased Arctic Amplification during 1980-2022

Aodhan J. Sweeney[1], Qiang Fu[1], Stephen Po-Chedley[2], Hailong Wang[3], Muyin Wang[4,5]

[1]University of Washington, Department of Atmospheric Sciences
[2]Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, Livermore, CA, USA
[3]Atmospheric Sciences and Global Change Division, PNNL
[4]Pacific Marine Environmental Laboratory, Oceanic and Atmospheric Research, NOAA
[5]University of Washington, Cooperative Institute for Climate, Ocean, and Ecosystem Studies

Contents of this file

Introduction

Table S1 is referenced in the main text, and Table S2 is referenced in the supplementary text. The supplementary text includes a more detailed description of machine learning methods, including the process for choosing model architecture, the approach for validation of the model, and the quantification of uncertainty. Supplementary figures (Figures S1 to S10) are referenced throughout the main text as well as in Text S1 and S2.

Table S1: Models with large ensembles with the number of ensemble members and time periods used in this study. The * indicates a combination of all other available models. The ML algorithm is not trained on this * data, but it is used for histograms in Figs 1, 2, and 4 in the main text.

Model Name	N Ensemble Members	Time Period Used
CESM2_SBMB	50	1900-2047
CESM2	50	1900-2047
ACCESS-ESM1-5	30	1900-2047
MPI-ESM1-2-LR	10	1900-2047
IPSL-CM6A-LR	11	1900-2047
MIROC6	50	1900-2047
CanESM5	25	1900-2047
GISS-E2-1-G	12	1900-2012
GISS-E2-1-H	10	1900-2012
MPI-ESM1-2-HR	10	1900-2012
NorCPM1	29	1900-2012
*OtherAllEM	59	1980-2022

Table S2: Hyperparameters of chosen architectures for the Arctic and Global CNNs.

Hyperparameter	Arctic CNN	Global CNN
Convolutional Layers	1	1
Activation Function	ReLU	ReLU
N filters	64	48
Kernel size	3x7	3x7
Dropout layer	50%	50%
Epochs	10	5

Text S1.

CNNs are widely used in image classification and have shown success when applied to climate science data (e.g., Chattopadhyay et al., 2020; Weyn et al., 2021; Yu et al., 2023; Connolly et al., 2023). CNNs work by using filters that are convolved over input data, which in our case are maps of SAT and SLP trends. CNNs can also handle different aspects of input data, like red, green, and blue color channels in an image. Here, our input data consists of two channels (the SAT and SLP trends). All input maps are regridded to a common 2.5°x2.5° latitude longitude grid before training. Input vectors for the global training are then Nx2x72x144 (N-samples, 2 channels, 72 latitudes, and 144 longitudes). For the Arctic case we use trend maps poleward of 20°N only, meaning the inputs are Nx2x28x144. The outputs for both CNNs are a vector of size Nx2 which represent the internally generated and externally forced trends averaged over the regions considered.

The architecture of our CNNs is chosen by trying a variety of model configurations and selecting the one which minimizes the mean squared error (MSE) over all models during 1980-2022 (see Fig. S8), where MSE here is the average MSE for either the internal or externally driven trends. We try filter sizes of 32, 48, and 64, kernel sizes of (1,1), (1,3), and (3,7) with training of either 5 or 10 epochs (i.e., 18 different architectures for both the Arctic and global CNN, tested by iterating through filter sizes, then kernel sizes, then epoch lengths, which correspond to x-axis labels in Fig. S8). CNNs used here are shallow (i.e., only have one convolutional layer) as increasing the number of convolutional layers significantly reduced the ability of the CNN to accurately separate internal and externally generated trends. To avoid overfitting, we opt for using a drop-out rate of 50% for the convolutional layer as opposed to using a regularization. While we note that this drop-out rate is high, previous studies have also shown that these large dropout rates cause the CNN to learn more slowly and reduce the risk of overfitting (Gordon and Barnes, 2022). The optimal Arctic and Global CNN hyperparameters are given in Table S2.

In testing our CNNs, we use a cross-validation approach whereby we predict the external and internal trends for each of the large ensembles by training on data from all other large ensembles. This step ensures that predictions for any set of simulations are not reliant on training data coming from the same large ensemble, and thus cannot learn model specific biases to predict the trend contributions. Because of the varied amounts of members for each of the different large ensembles, we only train our algorithm using 10 ensemble members from each large ensemble chosen at random with replacement. This process is repeated 50 times resulting in 50 different CNNs for each iteration of the cross-validation. Our prediction of the different trend components of the out-of-sample large ensemble is then the average prediction from these 50 CNNs. The standard deviation of estimates for the Arctic (global) case are 0.036 and 0.043 K/decade (0.011 and 0.023 K/decade) for the internal and externally generated trends among the 50 CNNs, thus averaging over 50 CNNs for each cross validation is critical to reduce variability associated with the CNN's random initialization of weights.

Text S2.

Uncertainty shown in Figs. 3 and 4 of the main text (and Figs. S7, S9, and S10 of the supplementary materials) are derived by considering both observational uncertainty (σ_{obs}) and ML prediction uncertainty (σ_{ML}). The σ_{obs} represents the uncertainty in the estimates of the forced and internal trends using different observational datasets, which is computed as the standard deviation of the 12 predictions (4 SAT trend maps X 3 SLP trend maps) and each prediction here is an average of the results using the 7 cross-validation ML algorithms.

The σ_{ML} represents the error associated with the ML predictions of the forced and internal trends compared to their actual values computed from the large ensemble. This σ_{ML} is the standard deviation over the differences between predictions and actual values. This standard deviation is weighted by the number of members in each large ensemble as to not bias the σ_{ML} towards large ensembles with more members. The final uncertainty is then computed as $\sigma = \sqrt{{\sigma_{obs}}^2 + {\sigma_{ML}}^2}$. This final uncertainty is then used to create the 2σ confidence intervals shown in Figs. 3 and 4 of the main text, and Figs. S7, S9, and S10 in the supplementary materials.

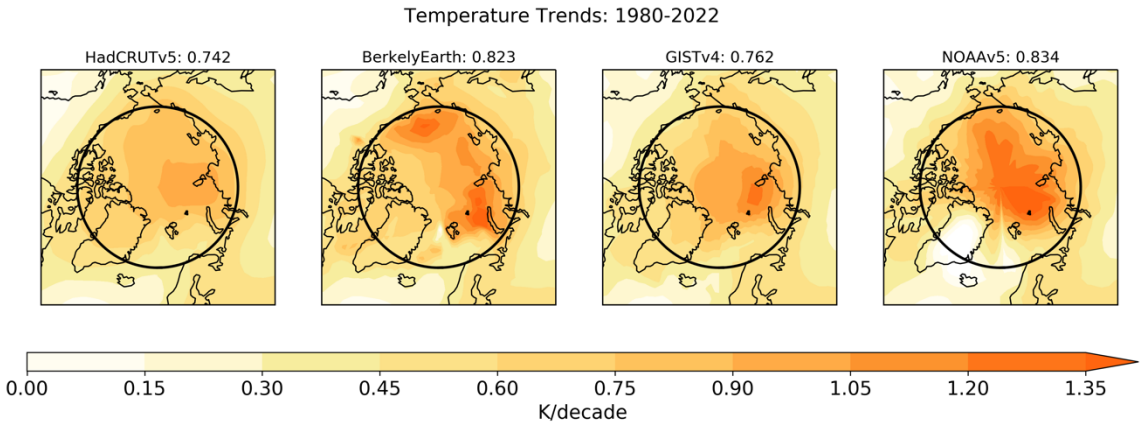

Fig. S1: Surface air temperature trends over the northern high latitudes from 1980-2022 from four different observational datasets. The Arctic is the region poleward of 70°N (black circle), and the temperature trend averaged over the Arctic in K/dec is provided next to the name of the observational dataset.

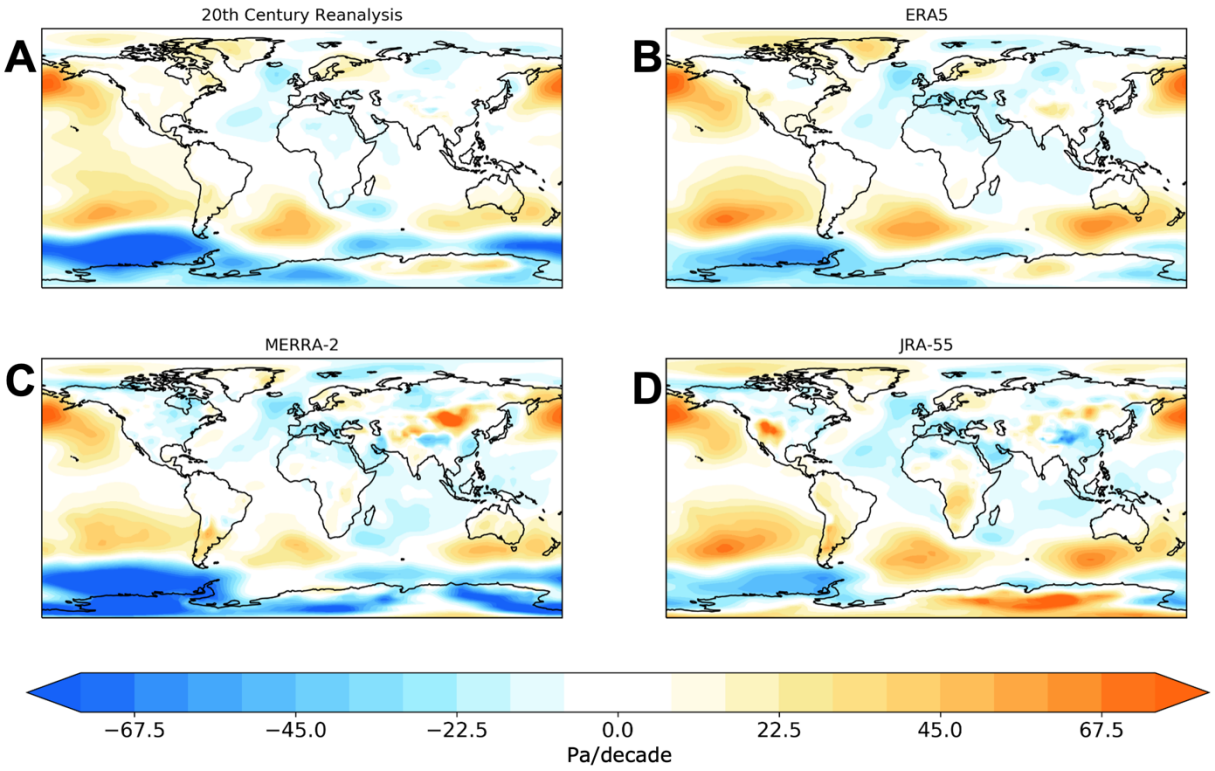

Fig. S2: Comparison of SLP trends from reanalysis datasets from 1980-2015. Panel A shows the trends computed from the 20th Century Reanalysis which was widely used in previous studies but is only available until 2015. Panels B-D show SLP trends over the same period from ERA5, MERRA-2, JRA-55, which are used in the present study for 1980-2022.

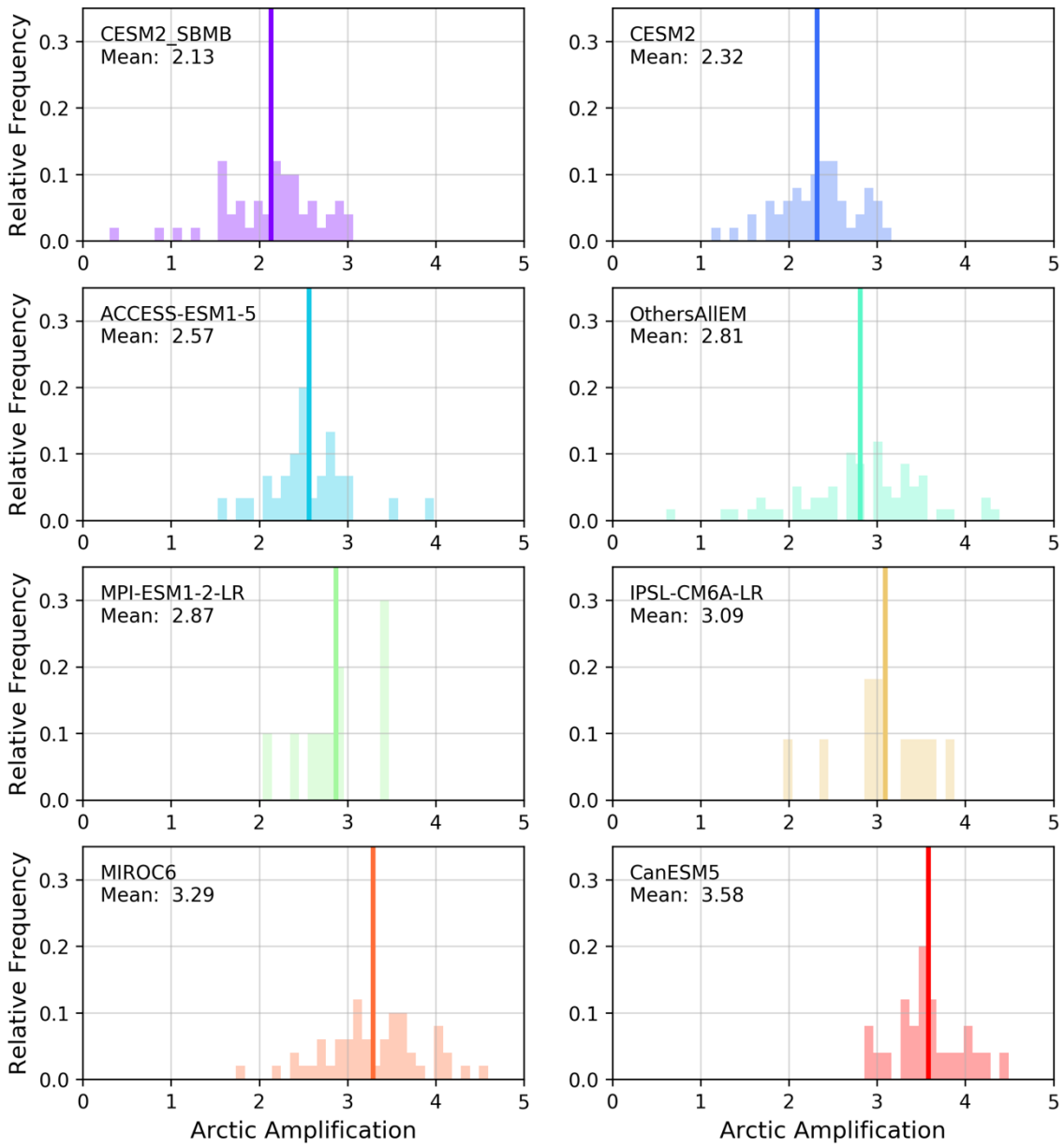

Fig. S3: Arctic amplification from 1980-2022, simulated from each large ensemble shown individually (instead of overlaid on top of one another in Fig. 1B). The mean Arctic amplification is shown as the vertical line.

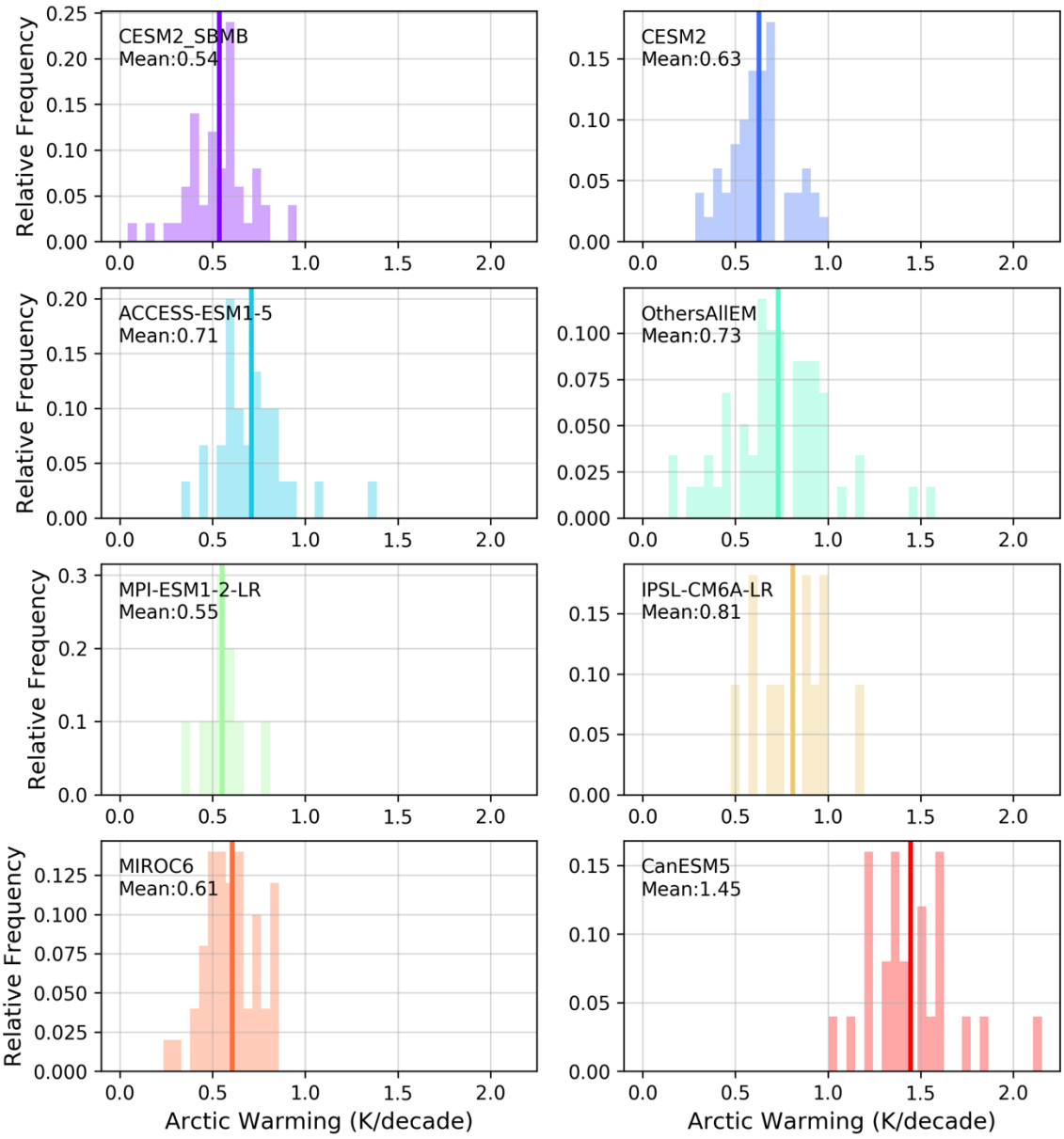

Fig. S4: Arctic warming from 1980-2022, simulated from each large ensemble shown individually (instead of overlaid on top of one another in Fig. 2A). The mean Arctic warming is shown as the vertical line.

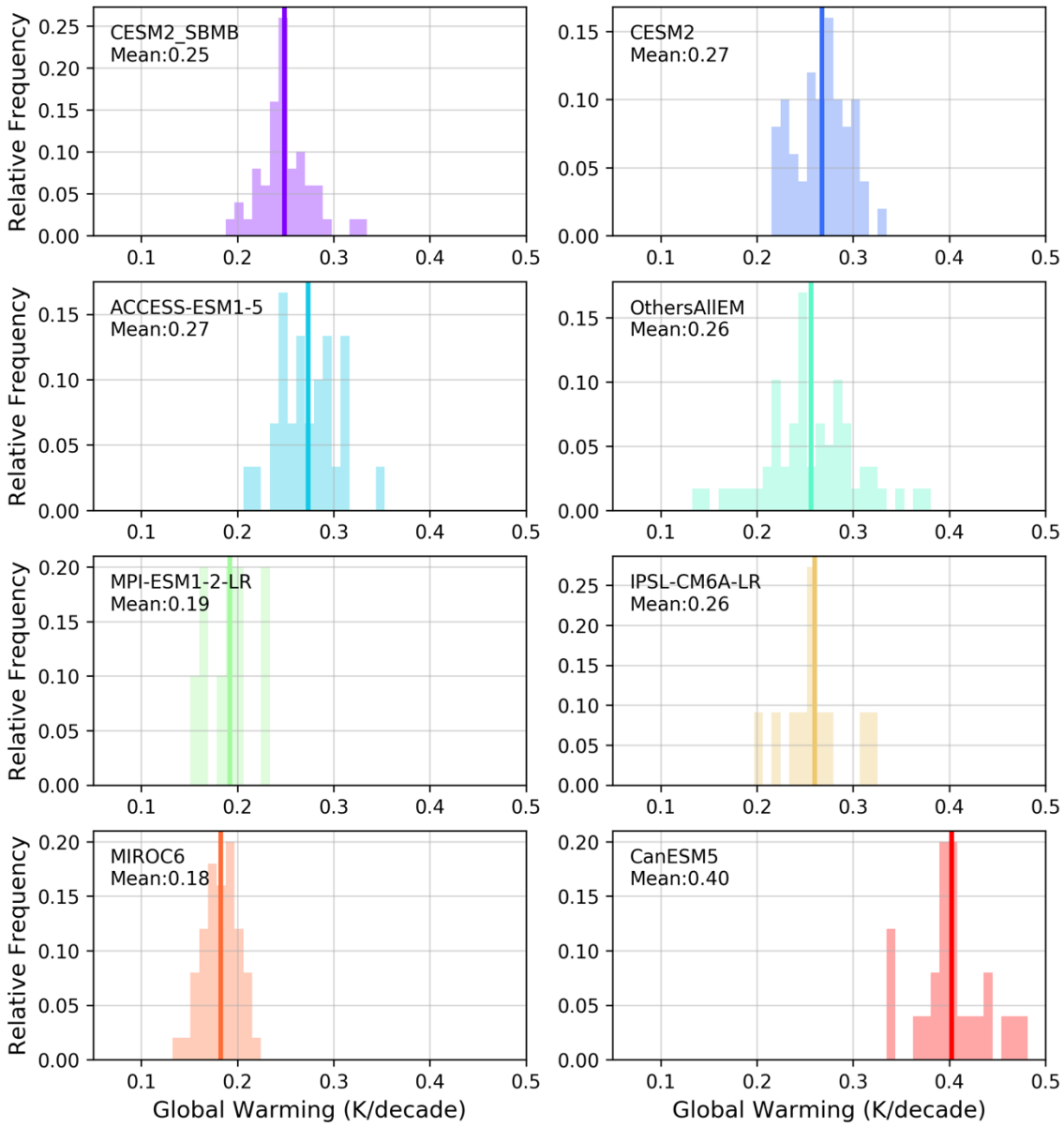

Fig. S5: Global warming from 1980-2022 simulated from each large ensemble shown individually (instead of overlaid on top of one another in Fig. 2B). The mean global warming is shown as the vertical line.

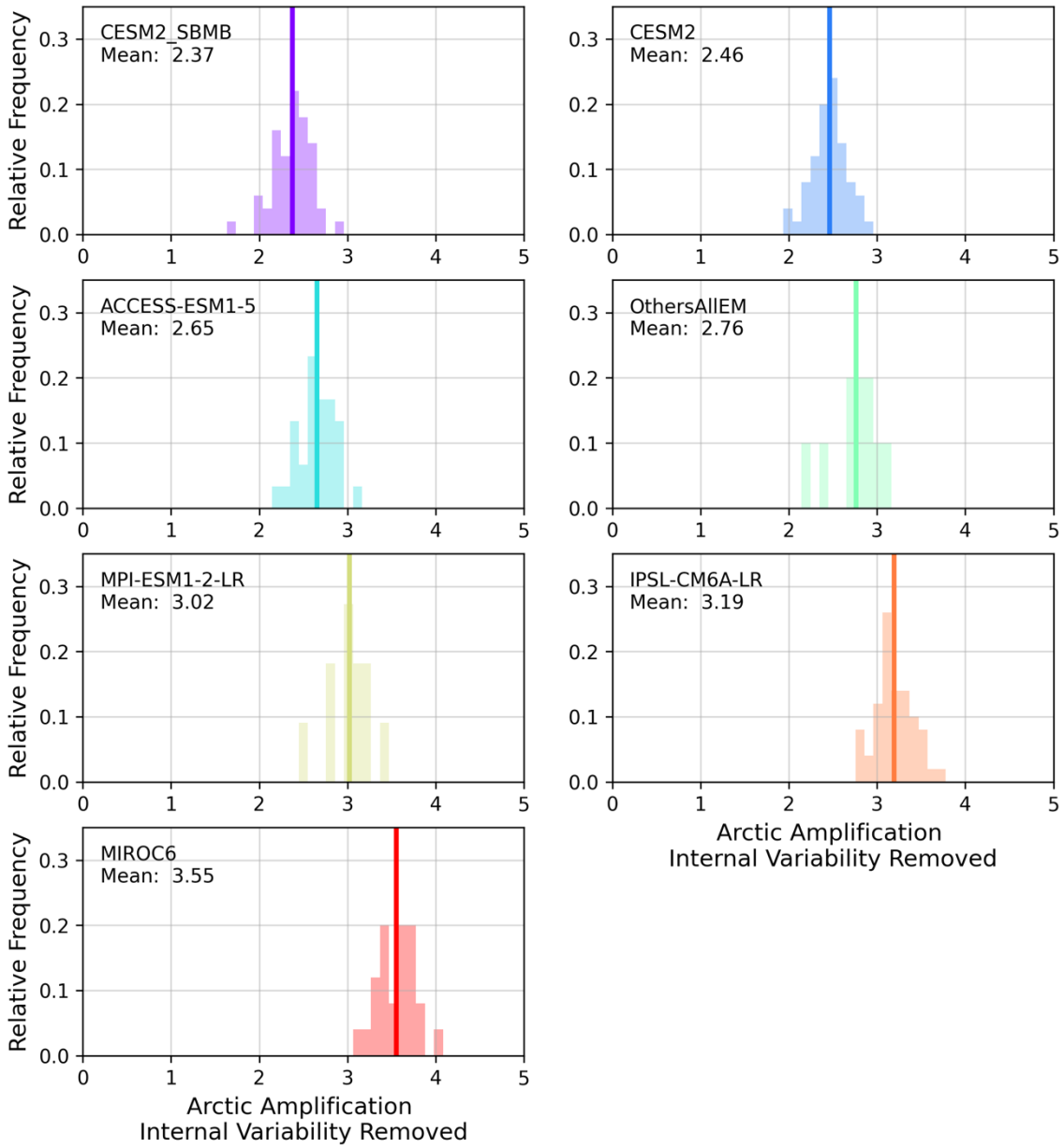

Fig. S6: Arctic Amplification simulated from each large ensemble after removing the internally generated trends predicted by the CNNs (instead of overlaid on top of one another in Fig. 4). The mean Arctic Amplification after removing internal variability is shown as the vertical line.

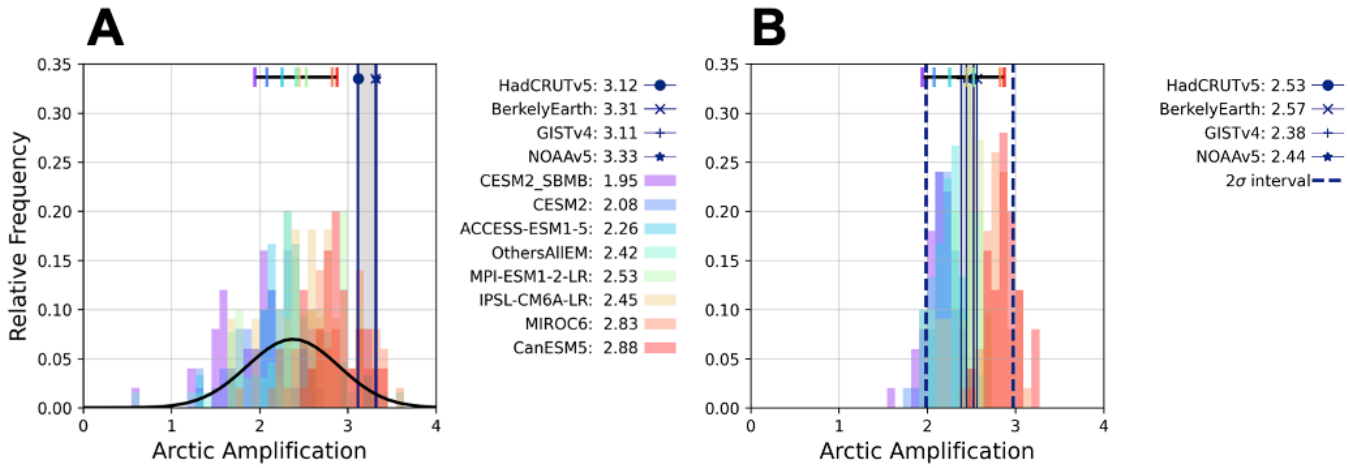

Fig. S7: (A) Comparison of observed and simulated AA from the four observational datasets and eight large ensembles using a definition of the Arctic from 60-90°. (B) Same as (A) but internal variability has been removed from observed Arctic and global temperature trends before calculating AA. Internal variability over 60-90°N is estimated using a newly trained CNN with the same architecture as for the 70-90° case. Internal variability for 60-90°N is estimated to be +0.082 K/Decade. The global internal variability is the same estimate used in the main text.

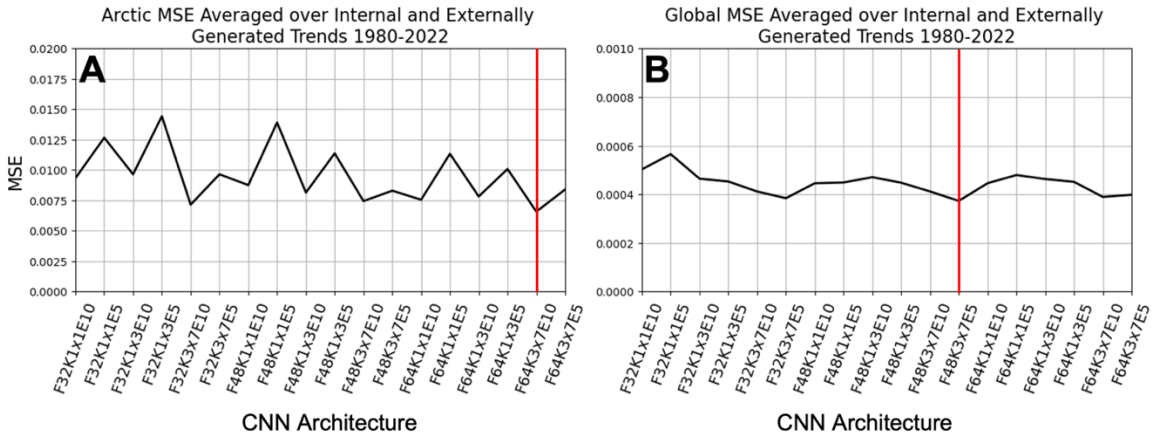

Fig. S8: Results from the CNN hyperparameter grid searches for the (A) Arctic and (B) global cases. MSE (y-axis) here is the average MSE over both the internally and externally generated predictions. Architecture indices (x-axis) specify the different filter sizes ('F'), kernel sizes ('K'), and the number of epochs used for training ('E'). Vertical red lines denote the architecture index that minimizes the average MSE. The optimal Arctic CNN architecture (architecture index 16 in A) uses 1 convolutional layer with 64 filters, a kernel size of 3x7 latitude-by-longitude pixels, trained with 10 epochs. The chosen optimal global CNN architecture (index 11) uses 1 convolutional layer, with 48 filters, a kernel size of 3x7, trained with 5 epochs.

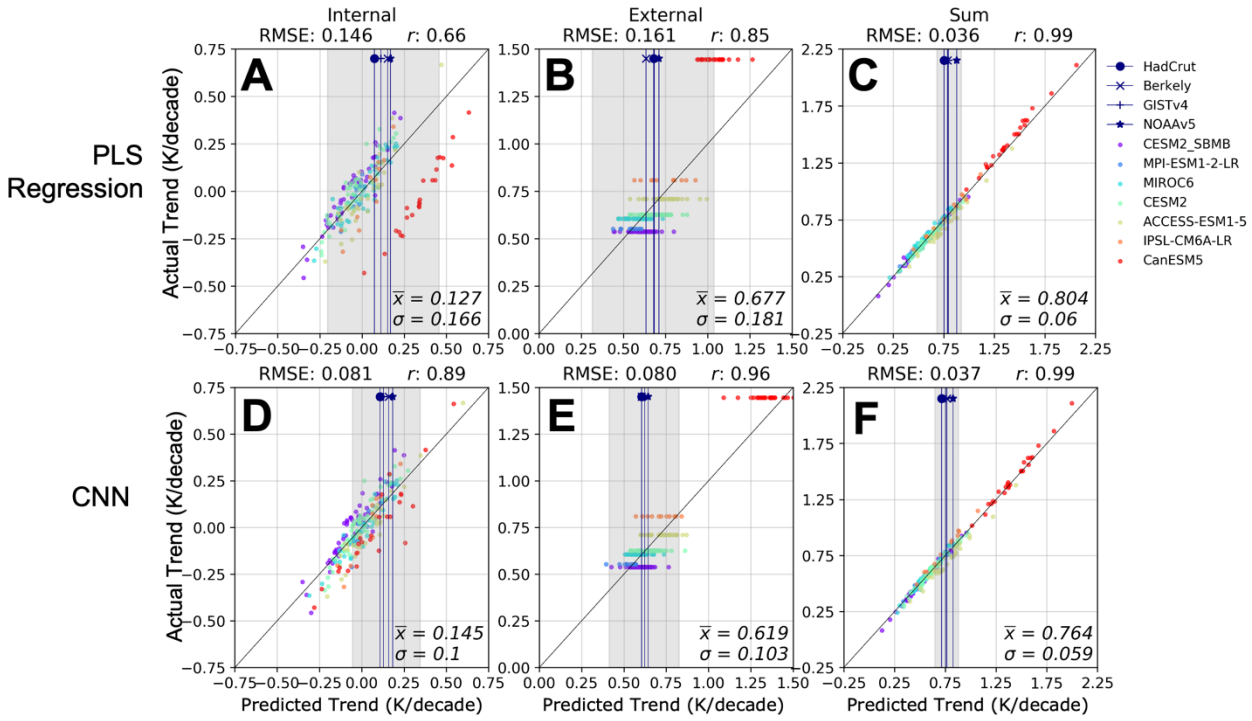

Fig. S9: Comparison of Arctic internal and external trend separation (A-C) using Partial Least Squares Regression (PLS-Regression) with SAT trend pattern poleward of 20°N and (D-F) using a Convolutional Neural Network (CNN) with both SAT and SLP trend patterns. PLS regression here used 6 components, which minimizes the MSE over the 1980-2022 period. The CNN architecture and results are the same as in Fig. 3 of the main text.

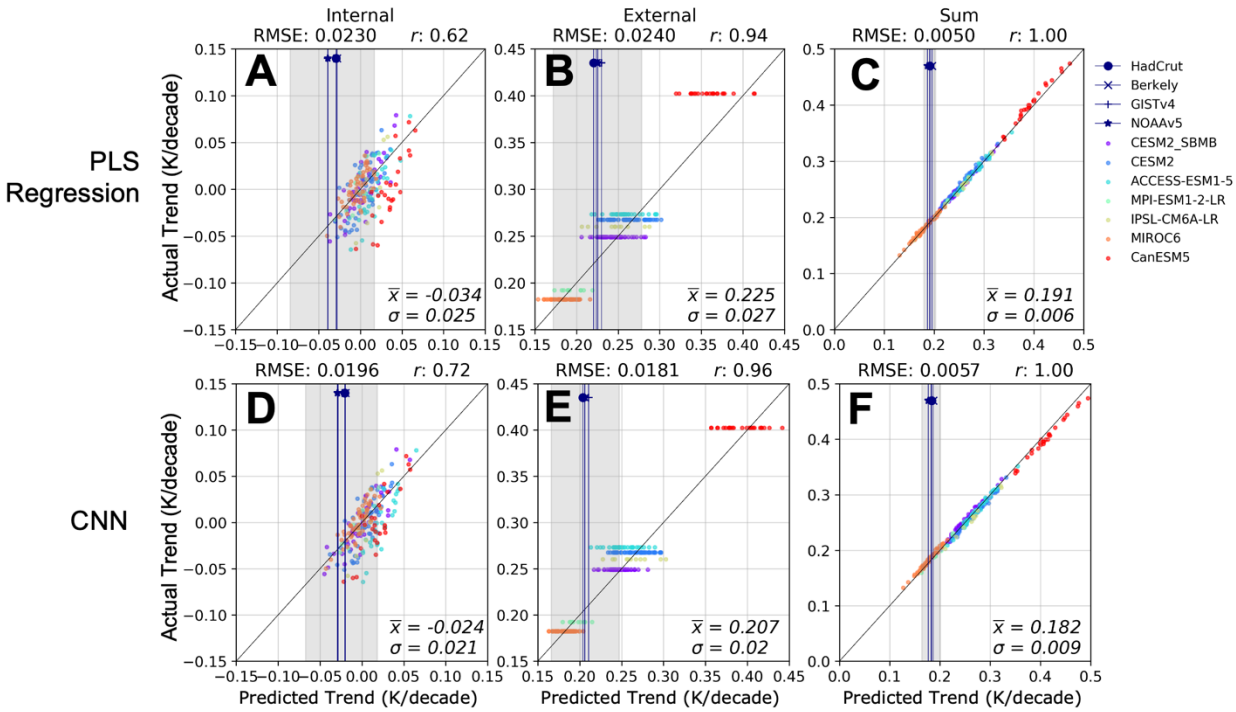

Fig. S10: The same as Fig. S9 but for global internal and external trend separation. PLS regression here used 3 components.

References

Chattopadhyay, A., Hassanzadeh, P., & Pasha, S. (2020). Predicting clustered weather patterns: A test case for applications of convolutional neural networks to spatio-temporal climate data. *Scientific Reports*, *10*(1), 1317. https://doi.org/10.1038/s41598-020-57897-9

Connolly, C., Barnes, E. A., Hassanzadeh, P., & Pritchard, M. (2023). Using Neural Networks to Learn the Jet Stream Forced Response from Natural Variability. *Artificial Intelligence for the Earth Systems*, *2*(2). https://doi.org/10.1175/AIES-D-22-0094.1

Gordon, E. M., & Barnes, E. A. (2022). Incorporating Uncertainty Into a Regression Neural Network Enables Identification of Decadal State-Dependent Predictability in CESM2. *Geophysical Research Letters*, *49*(15), e2022GL098635. https://doi.org/10.1029/2022GL098635

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*(7), e2021MS002502.

Yu, X., Millet, D. B., Henze, D. K., Turner, A. J., Delgado, A. L., Bloom, A. A., & Sheng, J. (2023). A high-resolution satellite-based map of global methane emissions reveals missing wetland, fossil fuel, and monsoon sources. *Atmospheric Chemistry and Physics*, *23*(5), 3325–3346. https://doi.org/10.5194/acp-23-3325-2023