<Supplementary Information>

Significant contribution of internal variability to recent Barents-Kara sea ice loss in winter

Peter Yu Feng Siew[1], Yutian Wu[1], Mingfang Ting[1], Cheng Zheng[1], Qinghua Ding[2], Richard Seager[1]

Corresponding author: Peter Yu Feng Siew
Email: pyfsiew@ldeo.columbia.edu

Affiliations:
[1] Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY, USA
[2] Department of Geography, and Earth Research Institute, University of California, Santa Barbara, Santa Barbara, CA, USA

Contents of this file:
Supplementary Figures S1 to S8
SI references

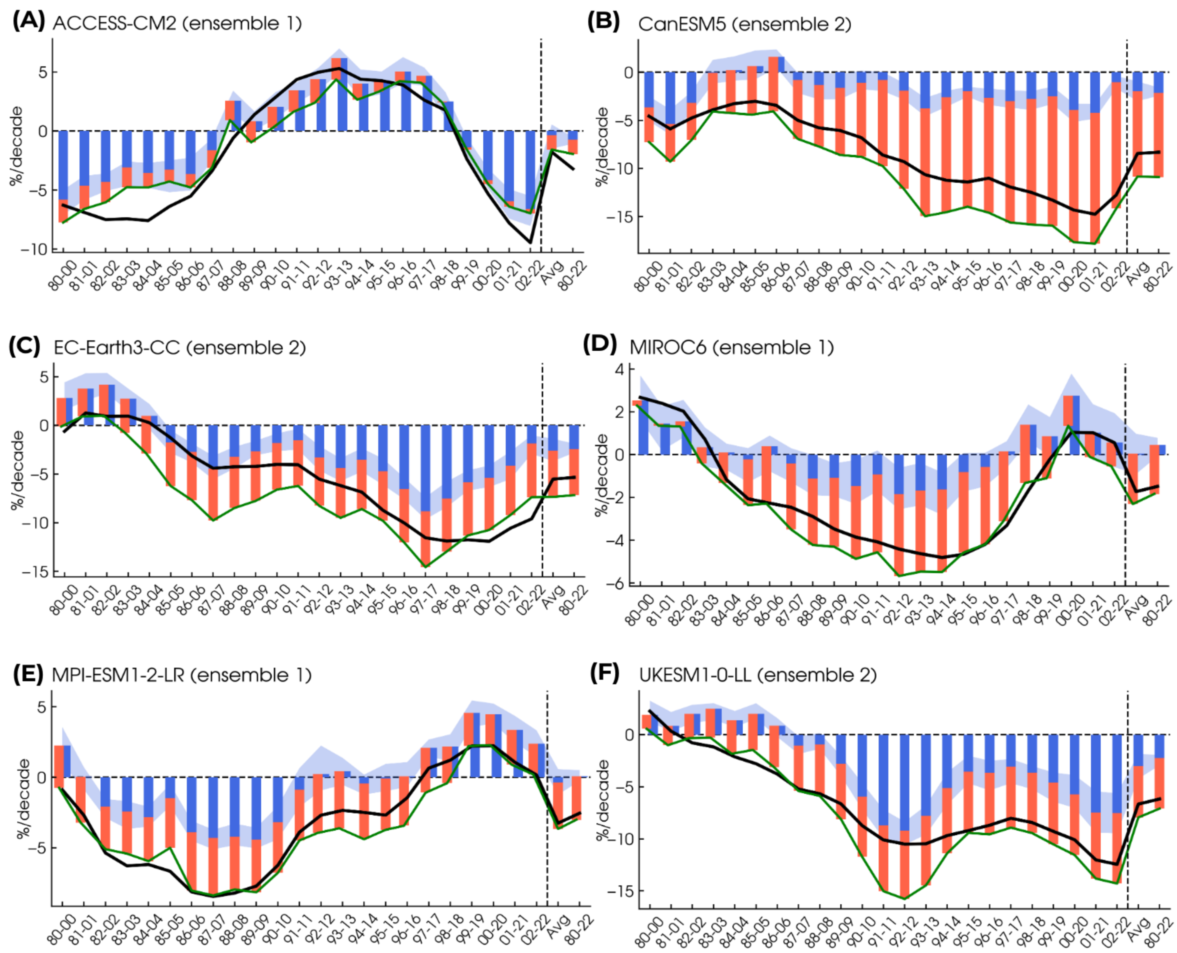

Figure S1: As in Figure 2, but estimating the anthropogenic (red bars) and circulation-driven (blue bars) Barents-Kara sea ice trends for either the first or second ensemble member of **(A)** ACCESS-CM2, **(B)** CanESM5, **(C)** EC-Earth3-CC, **(D)** MIROC6, **(E)** MPI-ESM1-2-LR and **(F)** UKESM1-0-LL forced simulation from CMIP6. The black line indicates linear trends of winter (DJF) Barents-Kara sea ice concentration (%/decade) from the simulation. The green line shows the sum of the estimated anthropogenic and circulation-driven components. The anthropogenic component is estimated by averaging 10 ensemble members from the testing model. The circulation-driven component is estimated by the machine-learning pattern recognition approach using all forced simulations (10 ensemble members each) but with the testing ensemble member removed from the training data.

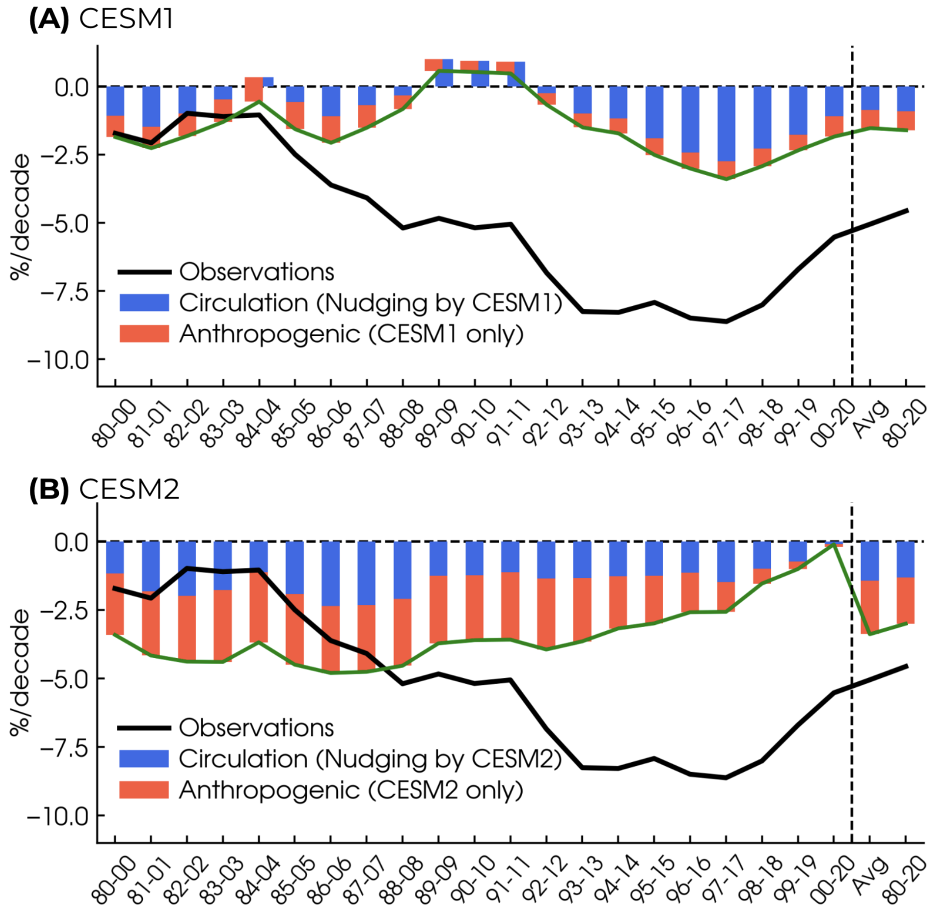

Figure S2: As in Figure 2, but estimating the anthropogenic (red bars) and circulation-driven (blue bars) Barents-Kara sea ice trends for observations using NCAR **(A)** CESM1 and **(B)** CESM2 models. The black line indicates linear trends of winter (DJF) Barents-Kara sea ice concentration (%/decade) from NSIDC satellite observations. The green line shows the sum of the estimated anthropogenic and circulation-driven components. The anthropogenic-driven component in (A) and (B) is estimated by averaging all available ensemble members from CESM1 (40 members) and CESM2 (100 members) forced simulations respectively. The circulation-driven component in (A) and (B) is estimated by dynamical nudging experiments performed by NCAR CESM1 (Ding et al. 2022) and CESM2 (Topál et al. 2023) respectively. The nudging experiments were only integrated up to 2020, so the last 20-year period (00-20) refers to DJF 2000/01 to 2019/2020.

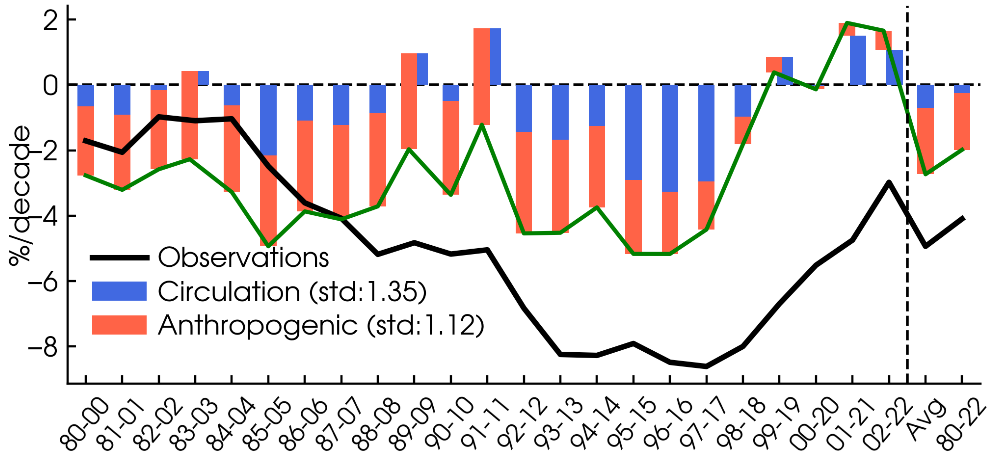

Figure S3: As in Figure 2, but estimating the anthropogenic (red bars) and circulation-driven (blue bars) Barents-Kara sea ice trends for observations using NCAR CESM2 model. The black line indicates linear trends of winter (DJF) Barents-Kara sea ice concentration (%/decade) from NSIDC satellite observations. The green line shows the sum of the estimated anthropogenic and circulation-driven components. The anthropogenic-driven component is estimated by averaging all available ensemble members (100 members) from CESM2 forced simulation. The circulation-driven component is estimated by the machine-learning pattern recognition approach using all available ensemble members (100 members) from CESM2 forced simulation.

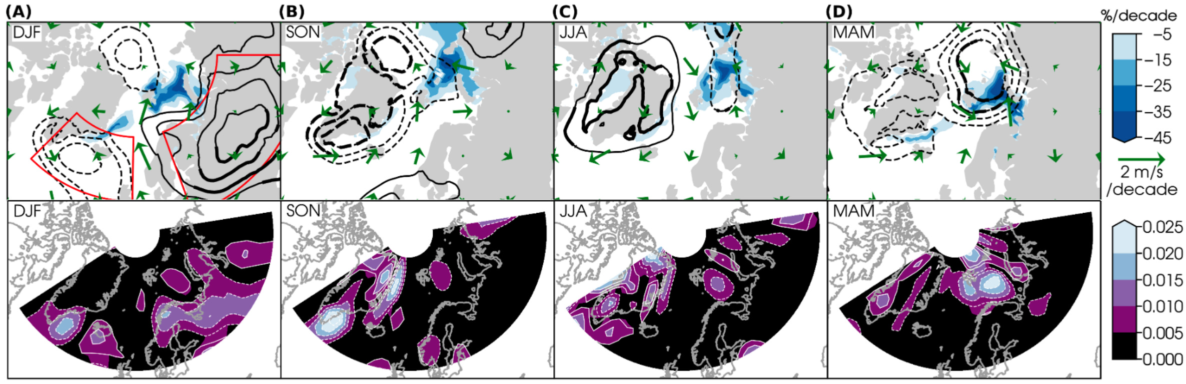

Figure S4: Upper row: As in Figure 3, but showing the linear trends of sea level pressure (black contours at 50 intervals > |100| Pa/decade; dashed contours indicate negative values; thick lines indicate significant values at 20% level using a two-tailed t-test) and 10-metre zonal and meridional winds (green vectors, m/s/decade; see scales at the right) from MERRA2 reanalysis for **(A)** DJF 1997/98-2016/17 and **(B)** SON 1997-2016 **(C)** JJA 1997-2016 and **(D)** MAM 1997-2016. Shading shows the corresponding linear trends of sea ice concentration (shading, %/decade) from NSIDC observations. Bottom row: "InputXGradient" attribution heatmaps (shading indicates absolute values; solid and dashed lines enclose positive and negative values respectively; unitless): regions with positive (negative) values indicate that the corresponding sea level pressure trends shown in the upper row contribute to an increase (a decrease) of the circulation-driven Barents-Kara sea ice trend for DJF 1997/98-2016/17 in the neural network estimation. Values at each grid point are smoothed by averaging the values with their neighboring grid points. The heatmaps are obtained by averaging across the trained models in all iterations under the "take-one-model-out" approach and random initializations. Another method "Integrated gradients" shows very similar results. "InputXGradient" and "Integrated gradients" calculations were performed by a Python package "Captum" (Kokhlikyan et al. 2020). Readers are referred to Mamalakis et al. 2022 for a review of neural network attribution methods in climate science.

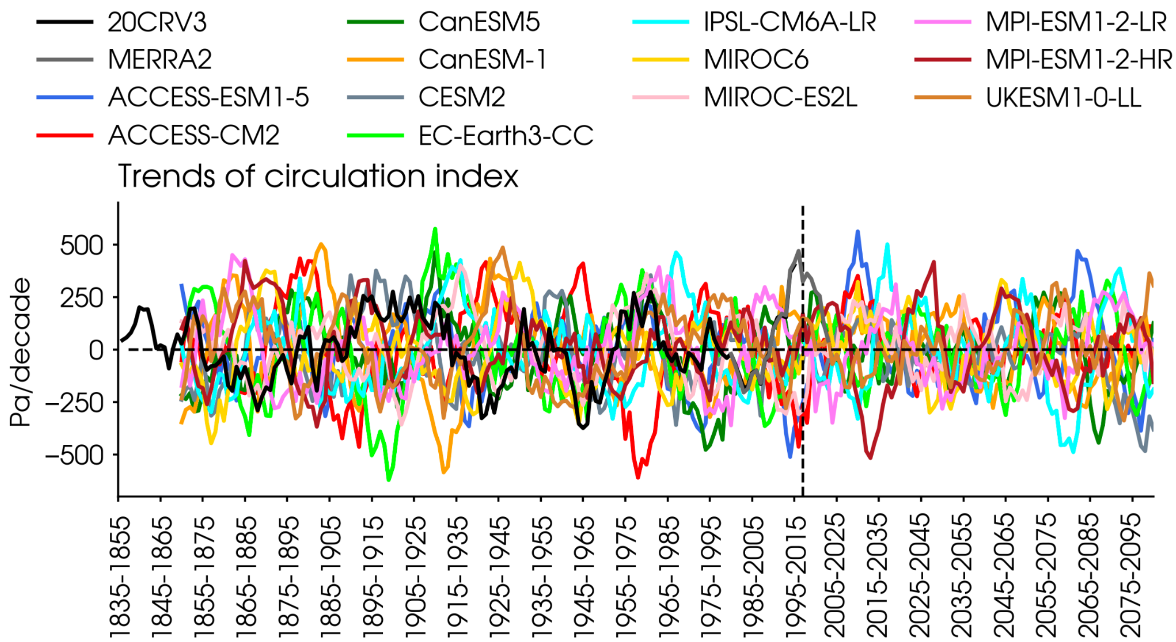

Figure S5: Linear trends of winter (DJF) circulation indices (Pa/decade) in all possible 20-year windows for 20th-century reanalysis V3 (black; from 1836-1856 to 1995-2015), MERRA2 (grey; from 1980-2000 to 2002-2022) and the first ensemble member of CMIP6 forced simulations (other colours; from 1850-1870 to 2080-2100). The vertical dashed line indicates the period of 1997-2017. The circulation index is defined as the difference in weighted area-averaged sea level pressures between the Urals and Icelandic regions shown in the red boxes in Figure 3A.

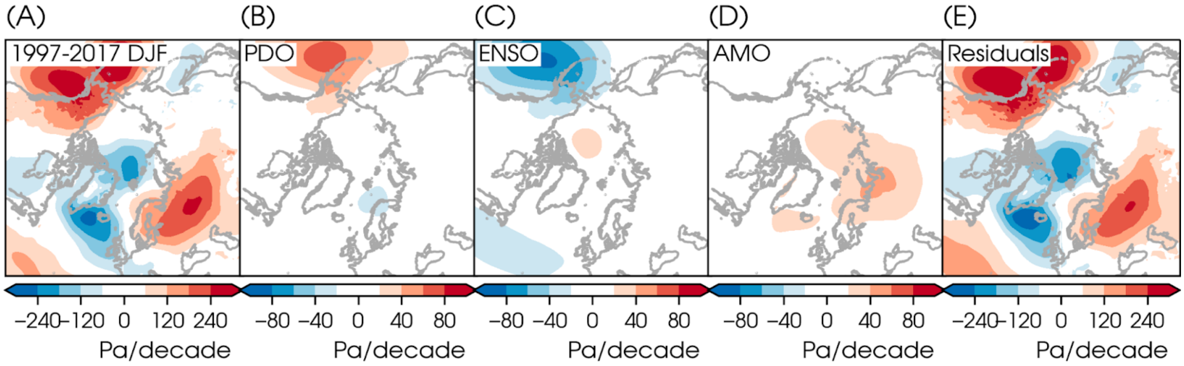

Figure S6: Linear trends of sea level pressure (shading, Pa/decade) over **(A)** DJF 1997/98-2016/17 and its decomposition into three dominant modes of sea surface temperature variability: **(B)** Pacific Decadal Oscillation (PDO), **(C)** El Niño-Southern Oscillation (ENSO), **(D)** Atlantic Multi-decadal Oscillation (AMO). **(E)** shows the unexplained residuals. The decomposition method follows Mckinnon and Deser 2018, where the monthly sea level pressure (20th-century reanalysis V3 from 1921-2015 and ERA5 from 2016 to 2018) in each grid point is linearly fitted to the three dominant modes of sea surface temperatures (represented by time series; derived from HadiSSTv1 from 1921-2018; Rayner 2003) by multiple linear regression by ordinary least squares method. Note that the colour bars in (B), (C) and (D) are different from (A) and (E).

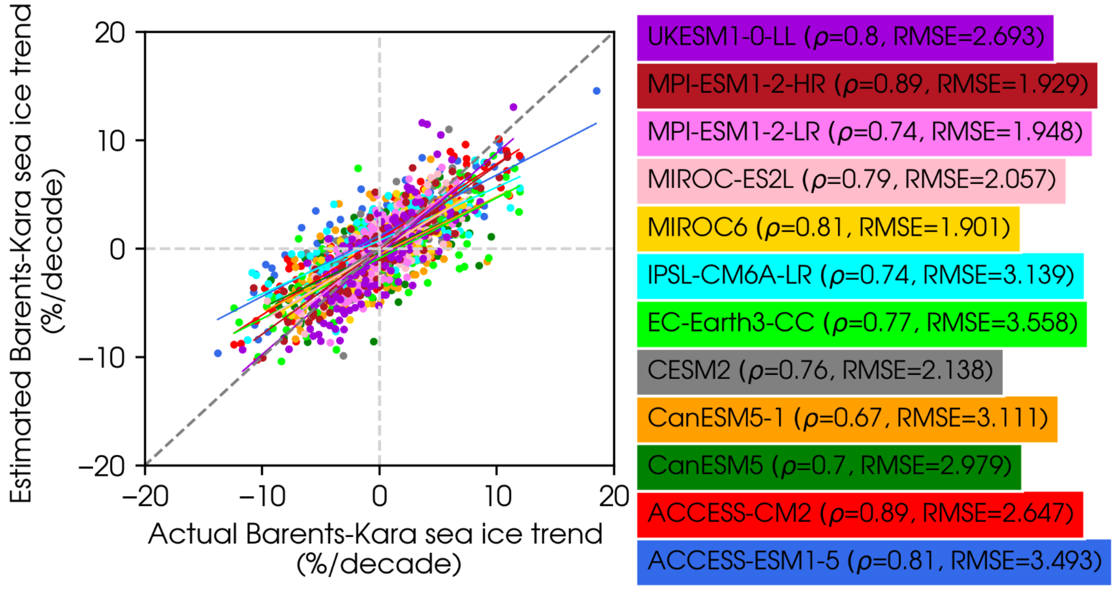

Figure S7: Relationships (dots) between the actual (X-axis) and estimated (Y-axis) Barents-Kara sea ice trends for each validating model (colours) under the “take-one-model-out” cross-validation approach. The actual trends are calculated directly from the model output with the ensemble mean removed. The estimated trends are obtained from the machine learning pattern recognition (or regression) approach. The straight, coloured lines represent the best-fit lines using standard least-squares regression. The coloured texts on the right correspond to the coloured dots and indicate the correlations (ρ) and root mean square error (RMSE) between the actual and estimated sea ice trends. The diagonal dashed line shows the one-to-one line.

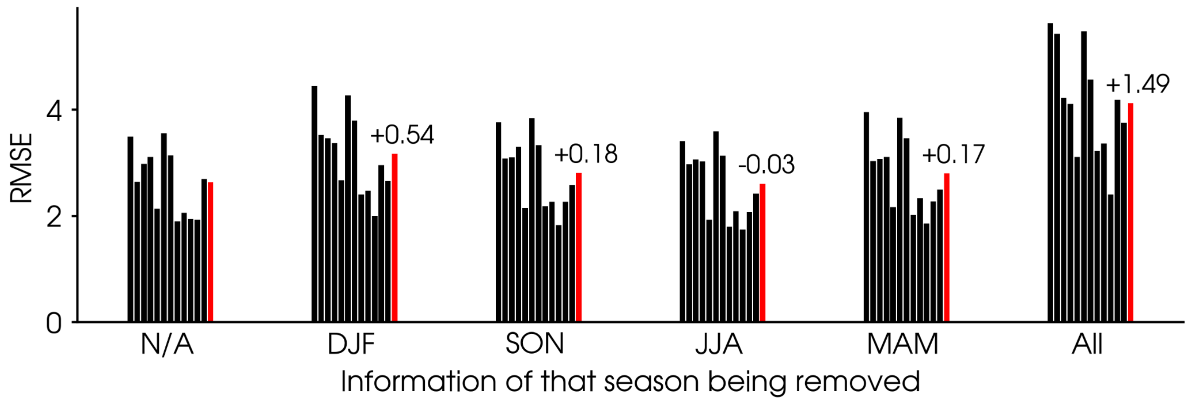

Figure S8: Root mean square errors (RMSEs) between actual and estimated Barents-Kara sea ice trends for each validating model under the "take-one-model-out" cross-validation approach. The first column shows the RMSEs from 12 individual models (black; from ACCESS-ESM1-5 to UKESM1-0-LL from left to right following the alphabetic orders shown in Figure S7) and their averages (red) without information on circulation trends being removed (hence the RMSES show the same values as in Figure S7). The second to last columns show the same RMSEs but with the information on circulation trends in DJF winter (2nd column), SON autumn (3rd column), JJA summer (4th column) and MAM spring (5th column) and all-season (last column) being removed when estimating the Barents-Kara sea ice trends for the validating model. The information on circulation trends in a season (or all seasons) is removed by randomly shuffling the information across the samples (10 ensemble and 13 20-year periods) in the validating models before the circulation trends are input into a trained network. The numbers above the red bars (2nd to last columns) indicate the increase of RMSEs compared to the mean-RMSE without any information being removed (red bar in the first column). A higher RMSE increase indicates that the sea ice trends depend more on the information on circulation trends from that season.

SI References:

Ding, Q., Schweiger, A. and Baxter, I., 2022. Nudging observed winds in the Arctic to quantify associated sea ice loss from 1979 to 2020. *Journal of Climate,* 35(20), pp.6797-6813.

Kokhlikyan, N., Miglani, V., Martin, M., Wang, E., Alsallakh, B., Reynolds, J., Melnikov, A., Kliushkina, N., Araya, C., Yan, S. and Reblitz-Richardson, O., 2020. Captum: A unified and generic model interpretability library for pytorch. *arXiv preprint* arXiv:2009.07896.

Mamalakis, A., Ebert-Uphoff, I. and Barnes, E.A., 2022. Neural network attribution methods for problems in geoscience: A novel synthetic benchmark dataset. *Environmental Data Science,* 1, p.e8.

McKinnon, K.A. and Deser, C., 2018. Internal variability and regional climate trends in an observational large ensemble. *Journal of Climate*, 31(17), pp.6783-6802.

Rayner, N.A.A., Parker, D.E., Horton, E.B., Folland, C.K., Alexander, L.V., Rowell, D.P., Kent, E.C. and Kaplan, A., 2003. Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research: Atmospheres,* 108(D14).

Topál, D. and Ding, Q., 2023. Atmospheric circulation-constrained model sensitivity recalibrates Arctic climate projections. *Nature Climate Change*, pp.1-9.