Tipping-Points and Robustness of Sea Ice using Koopman Mode Decomposition

Alan B. Cao[1], Michael Planer[2], James Hogg[1], Maria Fonoberova[1], Lake Bookman[2], Tyler Macdonald[2], Jennifer Hemingway[2], Qinghua Ding[1,3], Igor Mezić[1,3]

[1]AIMdyn, Inc., Santa Barbara, CA 93101, USA.
[2]BAE Systems, Falls Church, VA 22042, USA.
[3]University of California, Santa Barbara, CA 93106, USA.

Abstract

We demonstrate the application of Koopman Operator Theory (KOT) to model Arctic sea ice concentrations on decadal timescales and to identify potential climate tipping-points. Koopman-based models are computationally inexpensive to train and evaluate compared to traditional climate models, enabling robustness analyses of long-term climate trends and sensitivity analyses of the trends to various assumptions and uncertainties. We identify a potential tipping-point in the Barents and Kara Sea through Koopman Mode Decomposition (KMD) and verify that the Koopman-based models are robust to the uncertainty in the data used to train the model.

Introduction

Sea ice forms as ocean water freezes. The formation and distribution of sea ice plays an important role in Earth's climate and thus large amounts of climate data related to sea ice have been collected since 1978. The decreasing extent of Arctic sea ice over the last several decades has been dramatic, with the annual September sea ice extent declining by about 40% during the past 40 years [1] and sea ice volume by about 70% over the same time period [2]. Figure 1 shows the decline in annual September (minimum) Arctic sea ice extents since 1979.

State-of-the-art global climate models (GCMs) such as the Community Earth Science Model (CESM) have enabled climate scientists to improve the understanding and prediction of Earth's climate. As these models are typically based on physical principals such as fluid motion and energy transfer, they require expert knowledge, are challenging to develop, and are computationally expensive to run. Moreover, climate models do not accurately model several relevant climate subsystems including surface albedo [3] and cloud fraction [4]. In the case of sea ice, the current best models are known to underestimate the sensitivity of sea ice loss due to warming [5].

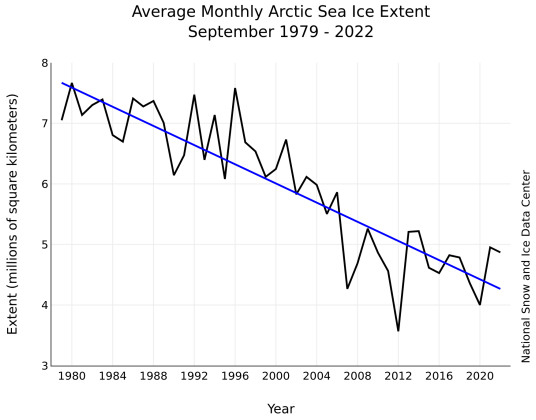

Figure 1: Average monthly September (minimum) Arctic sea ice extent from 1979-2022 obtained from the National Snow and Ice Data Center (NSIDC).

Recent advances in machine learning have shown potential for modeling high-dimensional climate dynamics directly from data [6]. While these black box methods relieve the need for expert knowledge, most are computationally expensive to train and reveal little about the underlying physical principals.

In this paper we demonstrate machine learning methods based on Koopman Operator Theory (KOT) [7] [8] [9] applied to climate data. Such methods have been previously applied to climate data and detected exponential decreases in sea ice concentration [10]. The Koopman-based climate models offer three main advantages over state-of-the-art GCMs: they (1) are completely data-driven and do not require prior climate knowledge, (2) are computationally inexpensive to train and evaluate, and (3) can capture underlying physics in the data. These benefits enable the Koopman-based models to run climate simulations under different forcing scenarios and to identify potential climate tipping-points. The Koopman-based models can also be used to improve existing GCMs by learning the underlying physics in the error between the GCM simulations and observational data.

Koopman Operator Theory

For some state space M we consider a nonlinear discrete-time *uncontrolled* dynamical system of the form:

$$\boldsymbol{x}_{k+1} = T(\boldsymbol{x}_k), \tag{1}$$

where $\boldsymbol{x}_k \in M$ is the state of the system at time k and $T : M \to M$ is the nonlinear state transition mapping between successive time steps. The Koopman operator U is an infinite-dimensional operator that describes the advancement of all scalar-valued observables $g : M \to \mathbb{C}$ forward in time through the dynamics:

$$Ug(\boldsymbol{x}_k) = g(T(\boldsymbol{x}_k)) = g(\boldsymbol{x}_{k+1}). \tag{2}$$

This enables embedding of nonlinear dynamics in high-dimensional function spaces. In practice, a set of observables $\boldsymbol{g}$ that span a subspace are chosen and an approximation to the action of the Koopman operator on that subspace is sought. Ideally, the subspace is chosen to be invariant under the action of the operator (this happens when the subspace is the span of the eigenfunctions). The observables can incorporate prior knowledge, resulting in an abstracted but understandable representation of the system. Figure 2 visually represents the action of the Koopman operator in the observable space. In contrast to many machine learning methodologies, the learned Koopman-based models can explain the coupling between different user-defined observables.

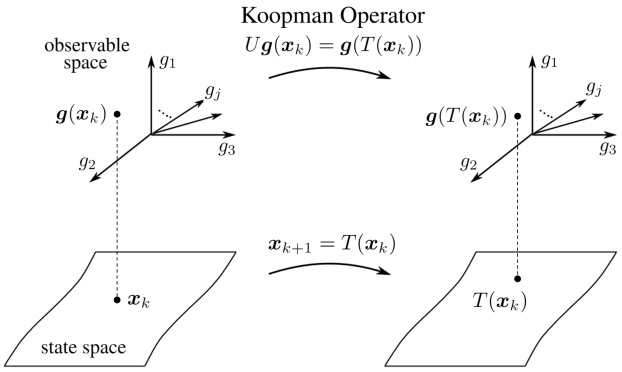

Figure 2: The nonlinear dynamics in state space can be represented by the Koopman operator in observable space, where the dynamics are linear but possibly infinite-dimensional.

Importantly, the Koopman operator is globally linear even if the underlying dynamical system is highly nonlinear. This allows the Koopman-based models to extrapolate well, even when trained on sparse data.

Koopman Mode Decomposition

The action of the Koopman operator on a set of observables $\boldsymbol{g}$ can be decomposed into decoupled dynamics in the form of Koopman eigenvalues, eigenfunctions, and modes:

$$U\boldsymbol{g}(\boldsymbol{x}_k) = \sum_{j=1}^{\infty} \boldsymbol{\nu}_j e^{(\lambda_j \Delta t)} \phi_j(\boldsymbol{x}_k). \tag{3}$$

The Koopman eigenvalues $e^{\lambda_j \Delta t}$ explain the evolution of the dynamics in time, where Δt denotes the time between consecutive time steps. The Koopman eigenfunctions ϕ_j explain the contribution of their corresponding modes and eigenvalues to the overall dynamics. The Koopman modes $\boldsymbol{\nu}_j$ are projections of fields of observables on the Koopman eigenfunctions, and indicate the spatial extent of the dynamics. Note that the Koopman eigenvalues and eigenfunctions are properties of the system, while the Koopman modes are associated with a specific observable and are only defined up to a constant.

While traditional statistical methods such as Empirical Orthogonal Function (EOF) analysis produce orthogonal components and are great for dimensional reduction of large datasets, Koopman Mode Decomposition (KMD) produces nonorthogonal Koopman modes while focusing on preserving the underlying dynamics. In linear systems with state observables, the Koopman modes are in fact eigenvectors of the system.

After training on N time steps, predictions of future behavior of the training data can be made using N dominant Koopman eigenvalues and modes:

$$\sum_{j=1}^{N} \boldsymbol{\nu}_j e^{(\lambda_j \Delta t)(n-1)} \tag{4}$$

For $n > N$, Equation (4) produces predictions for the nth time step based on the dynamics deduced from earlier observations.

Bifurcations and Tipping-Points

Bifurcations occur in dynamical systems when small smooth changes in parameters result in sudden qualitative changes in behavior. Several Koopman-based models can be constructed on different subintervals of the training dataset. Specifically, we train a series of models on a sliding window across the dataset to determine the occurrence of a bifurcation, or tipping-point [11]. We consider a supercritical pitchfork bifurcation:

$$\dot{y} = \mu y - y^3 \tag{5}$$

where y undergoes a bifurcation at $t = 30$, when μ abruptly changes from $+0.2$ to -0.2. Figure 3 shows the eigenvalues for each of the models as we slide the training interval over the dataset. After the bifurcation, the Koopman eigenvalues undergo a period of change before stabilizing to the new dynamics. Although the system converges from one stable fixed point to another stable fixed point, exponentially decaying and growing eigenvalues can appear during the transition.

Koopman Controlled Form

External forcing can be modelled as inputs to a nonlinear discrete-time *controlled* dynamical system of the form:

$$\boldsymbol{x}_{k+1} = S(\boldsymbol{x}_k, \boldsymbol{u}_k). \tag{6}$$

Considering the space of all input sequences P, $\boldsymbol{u}_k \in P$ is the input to the *controlled* system at time k and S :

$M \times P \rightarrow M$. The linearity of the Koopman operator in observable space can be extended to approximating the dynamics as a linear time-invariant (LTI) system:

$$g(\boldsymbol{x}_{k+1}) \approx Ug(\boldsymbol{x}_k) + B\boldsymbol{u}_k, \tag{7}$$

where the effect of the inputs on the dynamics is captured by the linear map B [12] [13]. Note that, in general, functions of $\boldsymbol{u}_k$ can be used in the linear representation, as well as products between functions of $\boldsymbol{u}_k$ and $\boldsymbol{x}_k$, capturing interactions between states and inputs in a more general framework. The Koopman Controlled Form (KCF) enables rapid simulations with different input sequences, allowing for robustness and sensitivity analysis. Furthermore, modeling the dynamics in this form immediately enables the application of linear control theory for determining inputs to drive the system to a desired state.

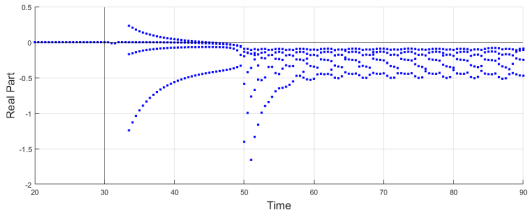

Figure 3: Koopman eigenvalues during a pitchfork bifurcation. A series of Koopman-based models are trained on a sliding window of 20 time steps using monomial observables. The vertical black line at $t = 30$ represents the occurrence of the bifurcation.

Climate Datasets

Due to the high-dimensionality of climate data, NASA's Earth science data archive is expected to hold more than 245 petabytes of data by 2025 [14]. Data-driven methods are well-suited for these large quantities of data.

Observational Datasets

Observational sea ice concentration data was obtained from the National Snow and Ice Data Center (NSIDC). These data include gridded daily and monthly averaged sea ice concentrations for both the north and south polar regions since 26 October 1978. The data shows the brightness temperature data derived from a few different sensors (microwave radiometers that sense emitted microwave radiation) that represent the sea ice concentration. The data are provided in the polar stereographic projection at a grid cell size of 25 x 25 km. In particular, we used the variable **seaice_conc** (NSIDC G02202_V4)[15].

Additional climate variables relevant to the sea ice dynamics that were used include:

- **T2M**: atmospheric surface temperature (ERA5) [16]
- **SST**: ocean surface temperature (ORAS5) [17]
- **Sea Ice Thickness**: sea ice thickness (PIOMAS v2.1)[18][19]

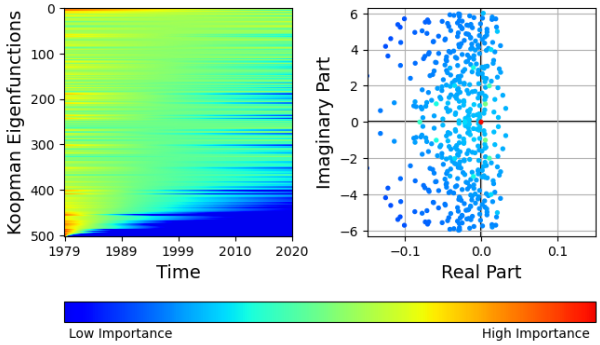

Figure 4: (Left) Each row represents a different Koopman eigenfunction evaluated on the NSIDC observational Arctic sea ice concentrations data from 1979 to 2020. The Koopman eigenfunctions are sorted with the greatest mean magnitude on top. The magnitudes of the Koopman eigenfunctions determine the contributions of their corresponding eigenvalues to the overall dynamics. The Koopman eigenvalues (right) are colored based on evaluating their corresponding eigenfunctions across the training data.

Climate Models Simulations

Although a second version of the Community Earth System Model (CESM) has been released, CESM2 has been found to overestimate the amount of clouds and rain in the Arctic when compared to CESM1, leading to higher mean temperatures that are not consistent with observational data [20]. In this work, we use CESM1 as the state-of-the-art GCM in the Arctic. As the CESM1 models are computationally expensive to run, we look at the CESM1 Large Ensemble Community Project (LENS)[21]. This project has produced a publicly available set of climate model simulations performed with the nominal 1-degree latitude/longitude version of the Community Earth System Model version 1 (CESM1). From the CESM1 LENS data, we used the following variables:

- **ICEFRAC**: sea ice concentration
- **TS**: atmospheric surface temperature
- **SST**: ocean surface temperature
- **HI**: sea ice thickness

Results

Koopman-based models were trained to learn the sea ice dynamics in the northern hemisphere. In addition to the NSIDC sea ice concentration data, the models were also trained on atmospheric temperature, ocean temperature, and sea ice thickness observational data. The Koopman-based climate models show potential for (1) identifing potential tipping-points, (2) predicting trends on decadal timescales, and (3) enabling robustness analysis of measurement uncertainty.

Koopman Mode Analysis

In contrast with many other machine learning methodologies, the Koopman operator framework produces a physically interpretable model. In particular, the method gives

insight to how the different observables are related to each other. By training our models on different combinations of climate variables, we gain insight into which variables are dependent on which. We have found sea ocean temperature to be the most correlated climate variable to sea ice concentration dynamics, while the atmospheric temperature and sea ice thickness dynamics were largely independent.

Applying KMD to the learned models reveals underlying low-dimensional dynamics in the form of Koopman modes and eigenvalues. The Koopman modes explain the spatial extent of the dynamics while their associated eigenvalues explain the temporal dynamics. Koopman-based models were trained on NSIDC observational sea ice concentration data. The learned Koopman eigenfunctions and eigenvalues are visualized in Figure 4

The *mean* mode is the mode for which the real and imaginary components of the associated eigenvalue are both equal or nearly equal to zero, indicating little to no time dependence. The *annual* mode is the mode whose associated eigenvalue has zero or nearly zero real component and whose imaginary component is very close to a frequency of 12 months. Figure 5 displays the mean and annual Koopman modes from an earlier period (1979-1983) compared to a more recent period (2016-2020). Comparison of the mean modes show that the mean sea ice extents have receded in later periods. The annual modes reveal similar results, with many of the Arctic marginal seas experiencing greater annual variation in the later periods. These modes suggest the existence of slow decaying behavior. The Koopman modes during these two time periods revalidate the results in [10].

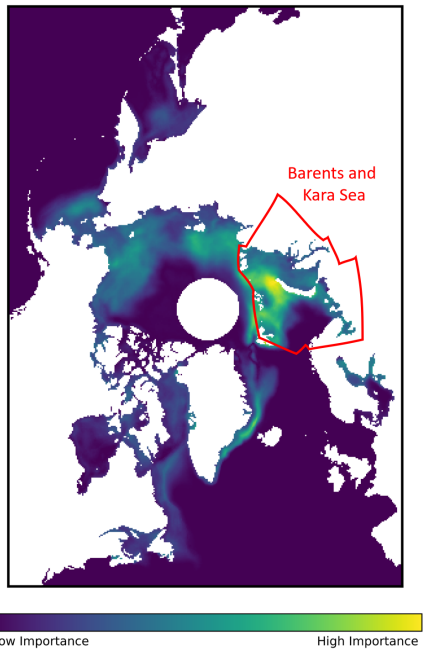

Figure 6: Most dominant Koopman mode of sea ice concentration over 1979-2020 in the Northern hemisphere whose associated eigenvalue has a nonzero real component, indicating exponential decay dynamics.

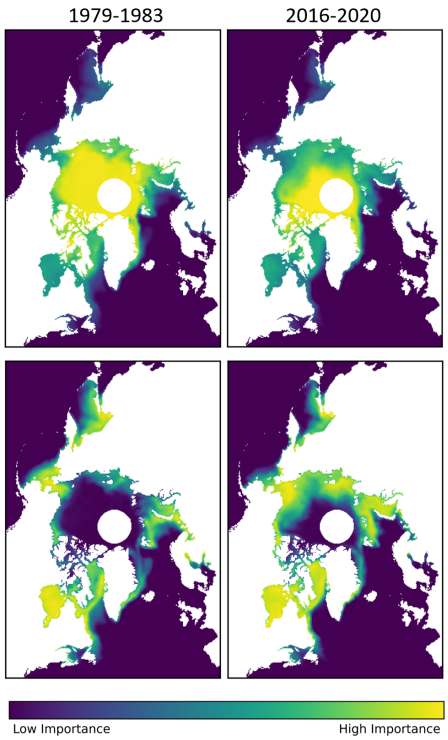

Figure 5: The *mean* (top) and *annual* (bottom) modes of sea ice concentration over different five-year periods (1979-1983 and 2016-2020) in the Northern hemisphere.

Potential Tipping-Points

Figure 6 shows one of the most dominant Koopman modes during 1979-2020 whose associated eigenvalue has a nonzero real component, representing exponential decay or growth dynamics. Because we see the sudden appearance of this mode exhibiting exponential decay on a relatively short time scale amidst previously highly stationary dynamics, we believe it indicates the existence of a climate tipping-point. This mode is mainly supported in the Barents and Kara Sea region, where an exponential decay with decay rate of 11.381 years is detected. This tipping-point has been confirmed by traditional climate research [22].

Koopman-Based Climate Simulations

The Koopman-based models learn the dynamics based on observations and can be used to produce future predictions. The global linearity of the Koopman operator allows the Koopman-based climate models to extrapolate well when faced with sparse data. To produce the predictions in Figure 7, a Koopman-based model was first trained on observational sea ice concentration, atmospheric temperature, ocean temperature, and sea ice thickness data over the period of 1979-2008. After training, the model is only given knowledge of forcing to build predictions of sea ice concentration. We compare our Arctic sea ice extent predictions (blue) to the mean of the CESM1 LENS predictions (orange) and the climatological mean (green). Each calendar

month in the climatological mean is the mean of the values of that calendar month over the entire training interval. On decadal timescales, the Koopman-based model seems to have learned the true dynamics, as can be seen by the qualitative similarity between the slowdown in sea ice loss in the NSIDC data and our predictions after 2008.

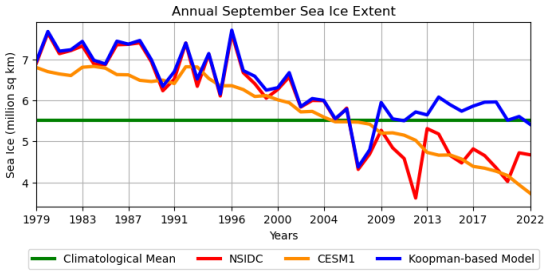

Figure 7: Annual September sea ice extent from 1979 to 2020. (Blue) Koopman-based model trained from 1979 to 2008. (Orange) Mean of CESM1 LENS predictions. (Green) Climatological mean of NSIDC observations.

Climate models such as CESM are driven by forcing terms such as greenhouse gases and other anthropogenic factors. The Koopman-based models in Koopman Controlled Form (KCF) have the potential to enable rapid climate simulations with different forcing scenarios. These experiments would enable understanding of how future anthropogenic activities might affect the climate.

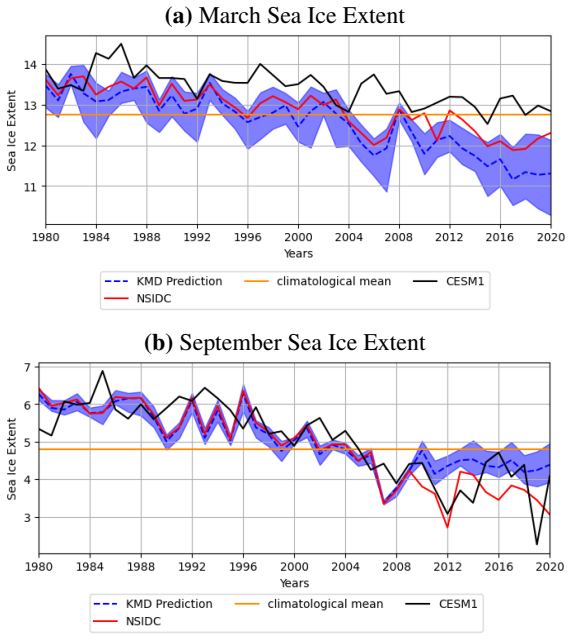

Figure 8: March (maximum) and September (minimum) sea ice extents from 1980 to 2020. (Blue) Koopman-based model trained from 01/1979 to 12/2009 with 2-σ uncertainty bands. (Black) CESM1 Large Ensemble member 002. (Orange) Climatological mean of NSIDC observations.

Robustness to Observational Uncertainty

Due to the chaotic nature of climate systems, predictions are typically generated in ensembles with small perturbations on the data or parameters. The Koopman-based models are well-suited for this ensemble approach due to their low computational cost. We generated 50 perturbed climate datasets by adding Gaussian noise to the observational measurements at the magnitude of measurement uncertainty. Koopman-based models were trained for each of these datasets, resulting in 50 different models. Figure 8 shows the average prediction of the ensemble of 50 models as well as the 2-σ envelope.

The ensemble of 50 models found that the flattening trend in the annual September sea ice extent was robust to statistical noise in the observational data. All 50 models also found exponential modes with 11 year decay rates in the Barents Sea region, confirming the robustness of the potential tipping-point.

Conclusions and Future Work

Methods based on Koopman Operator Theory show potential for modeling and prediction for the field of climate science. They are advantageous over state-of-the-art climate models and other emerging machine learning methods because they eliminate the need for expert knowledge and can learn directly from data while being computationally inexpensive to train and evaluate. The Koopman-based models in this paper were trained on observational data and successfully extract low-dimensional dynamics in the form of Koopman modes and eigenvalues. We also find the sudden appearance of an exponentially decaying Koopman eigenvalue, suggesting the occurrence of a climate tipping-point.

Predictions from the Koopman-based models show potential for rapid climate simulations under different forcing conditions, but direct comparison of our methods with other state-of-the-art methods is needed. Preliminary experiments training Koopman-based models on the error between observational data and climate simulations have shown potential for extracting dynamics that exist in the data but are not accounted for in the climate model, and warrants future investigation. Furthermore, work is needed to test different combinations of climate and forcing variables to understand their influence on the sea ice dynamics. By extending this work to other climate domains, we may gain further understanding of the coupling of different climate variables and the effect of anthropogenic forcing.

Acknowledgments

This material is based upon work supported by the Defense Advanced Research Projects Agency (DARPA) under Agreement No. HR00112290036 as part of the AI-Assisted Climate Tipping Point (ACTM) program. Alan B. Cao, James Hogg, Maria Fonoberova, and Igor Mezić were also supported by ONR project N00014-21-1-2384.

References

[1] NASA Global Climate Change, "Arctic sea ice minimum," 2021. [Online]. Available: https://climate.nasa.gov/vital-signs/arctic-sea-ice/

[2] Polar Science Center, "Piomas arctic sea ice volume reanalysis," 2021. [Online]. Available: http://psc.apl.uw.edu/research/projects/arctic-sea-ice-volume-anomaly/

[3] J. Karlsson and G. Svensson, "Consequences of poor representation of arctic sea-ice albedo and cloud-radiation interactions in the cmip5 model ensemble," *Geophys. Res. Lett.*, p. 4374– 4379, 2021. [Online]. Available: http://psc.apl.uw.edu/research/projects/arctic-sea-ice-volume-anomaly/

[4] P. C. Taylor, R. C. Boeke, Y. Li, and D. W. J. Thompson, "Arctic cloud annual cycle biases in climate models," *Atmospheric Chemistry and Physics*, vol. 19, no. 13, pp. 8759–8782, 2019. [Online]. Available: https://acp.copernicus.org/articles/19/8759/2019/

[5] J. Stroeve and D. Notz, "Changing state of arctic sea ice across all seasons," *Environmental Research Letters*, vol. 13, no. 10, p. 103001, 2018. [Online]. Available: iopscience.iop.org/article/10.1088/1748-9326/aade56/meta

[6] T. R. Andersson, J. S. Hosking, M. Pérez-Ortiz, B. Paige, A. Elliott, C. Russell, S. Law, D. C. Jones, J. Wilkinson, T. Phillips, J. Byrne, S. Tietsche, B. B. Sarojini, E. Blanchard-Wrigglesworth, Y. Aksenov, R. Downie, and E. Shuckburgh, "Seasonal arctic sea ice forecasting with probabilistic deep learning," *Nature Communications*, vol. 12, no. 1, aug 2021.

[7] I. Mezić, "Spectral properties of dynamical systems, model reduction and decompositions," *Nonlinear Dyn*, vol. 41, pp. 309–325, 2005.

[8] C. W. ROWLEY, I. MEZIĆ, S. BAGHERI, P. SCHLATTER, and D. S. HENNINGSON, "Spectral analysis of nonlinear flows," *Journal of Fluid Mechanics*, vol. 641, pp. 115–127, nov 2009.

[9] M. Budišić, R. Mohr, and I. Mezić, "Applied koopmanism," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 22, no. 4, p. 047510, dec 2012.

[10] J. Hogg, M. Fonoberova, and I. Mezić, "Exponentially decaying modes and long-term prediction of sea ice concentration using koopman mode decomposition," *Sci Rep*, vol. 10, p. 16313, 2020.

[11] T. M. Lenton, H. Held, E. Kriegler, J. W. Hall, W. Lucht, S. Rahmstorf, and H. J. Schellnhuber, "Tipping elements in the earth's climate system," *Proceedings of the National Academy of Sciences*, vol. 105, no. 6, pp. 1786–1793, feb 2008.

[12] M. Korda and I. Mezić, "Linear predictors for nonlinear dynamical systems: Koopman operator meets model predictive control," 2016.

[13] J. L. Proctor, S. L. Brunton, and J. N. Kutz, "Dynamic mode decomposition with control," 2014.

[14] NASA, "Nasa turns to the cloud for help with next-generation earth missions," 2021. [Online]. Available: https://www.jpl.nasa.gov/news/nasa-turns-to-the-cloud-for-help-with-next-generation-earth-missions?utm_source=iContact&utm_medium=email&utm_campaign=nasajpl&utm_content=earth20211013-1

[15] W. N. Meier, F. Fetterer, A. K. Windnagel, and J. S. Stewart., "Noaa/nsidc climate data record of passive microwave sea ice concentration, version 4," 2021. [Online]. Available: https://nsidc.org/data/G02202/versions/4

[16] Copernicus Climate Change Service, "Era5 monthly averaged data on single levels from 1979 to present," 2019. [Online]. Available: https://cds.climate.copernicus.eu/doi/10.24381/cds.f17050d7

[17] H. Zuo, M. A. Balmaseda, S. Tietsche, K. Mogensen, and M. Mayer, "The ecmwf operational ensemble reanalysis–analysis system for ocean and sea ice: a description of the system and assessment," *Ocean Science*, vol. 15, no. 3, pp. 779–808, 2019. [Online]. Available: https://os.copernicus.org/articles/15/779/2019/

[18] A. Schweiger, R. Lindsay, J. Zhang, M. Steele, H. Stern, and R. Kwok, "Uncertainty in modeled arctic sea ice volume," *Journal of Geophysical Research: Oceans*, vol. 116, no. C8, 2011. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011JC007084

[19] J. Zhang and D. Rothrock, "Modeling global sea ice with a thickness and enthalpy distribution model in generalized curvilinear coordinates," *Mon. Weather Rev.*, vol. 131, pp. 845–861, 2003.

[20] E. A. McIlhattan, J. E. Kay, and T. S. L'Ecuyer, "Arctic clouds and precipitation in the community earth system model version 2," *Journal of Geophysical Research: Atmospheres*, vol. 125, no. 22, nov 2020.

[21] J. E. Kay, C. Deser, A. Phillips, A. Mai, C. Hannay, G. Strand, J. M. Arblaster, S. C. Bates, G. Danabasoglu, J. Edwards, M. Holland, P. Kushner, J.-F. Lamarque, D. Lawrence, K. Lindsay, A. Middleton, E. Munoz, R. Neale, K. Oleson, L. Polvani, and M. Vertenstein, "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*, vol. 96, no. 8, pp. 1333–1349, 2015. [Online]. Available: https://journals.ametsoc.org/view/journals/bams/96/8/bams-d-13-00255.1.xml

[22] D. I. A. McKay, A. Staal, J. F. Abrams, R. Winkelmann, B. Sakschewski, S. Loriani, I. Fetzer, S. E. Cornell, J. Rockström, and T. M. Lenton, "Exceeding 1.5°c global warming could trigger multiple climate tipping points," *Science*, vol. 377, no. 6611, sep 2022.