

# SANDIA REPORT

SAND2020-9932

Printed September 2020



Sandia  
National  
Laboratories

## Arctic Tipping Points Triggering Global Change LDRD Final Report

Kara Peterson, Amy Powell, Irina Tezaur, Erika Roesler, Jake Nichol, Matt Peterson, Warren Davis, John Jakeman, David Stracuzzi, Diana Bull

Prepared by  
Sandia National Laboratories  
Albuquerque, New Mexico 87185  
Livermore, California 94550

Issued by Sandia National Laboratories, operated for the United States Department of Energy by National Technology & Engineering Solutions of Sandia, LLC.

**NOTICE:** This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, make any warranty, express or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government, any agency thereof, or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof, or any of their contractors.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from

U.S. Department of Energy  
Office of Scientific and Technical Information  
P.O. Box 62  
Oak Ridge, TN 37831

Telephone: (865) 576-8401  
Facsimile: (865) 576-5728  
E-Mail: reports@osti.gov  
Online ordering: <http://www.osti.gov/scitech>

Available to the public from

U.S. Department of Commerce  
National Technical Information Service  
5301 Shawnee Road  
Alexandria, VA 22312

Telephone: (800) 553-6847  
Facsimile: (703) 605-6900  
E-Mail: orders@ntis.gov  
Online order: <https://classic.ntis.gov/help/order-methods>



## **ABSTRACT**

The Arctic is warming and feedbacks in the coupled Earth system may be driving the Arctic to tipping events that could have critical downstream impacts for the rest of the globe. In this project we have focused on analyzing sea ice variability and loss in the coupled Earth system. Summer sea ice loss is happening rapidly and although the loss may be smooth and reversible, it has significant consequences for other Arctic systems as well as geopolitical and economic implications. Accurate seasonal predictions of sea ice minimum extent and long-term estimates of timing for a seasonally ice-free Arctic depend on a better understanding of the factors influencing sea ice dynamics and variation in this strongly coupled system.

Under this project we have investigated the most influential factors in accurate predictions of September Arctic sea ice extent using machine learning models trained separately on observational data and on simulation data from five E3SM historical ensembles. Monthly averaged data from June, July, and August for a selection of ice, ocean, and atmosphere variables were used to train a random forest regression model. Gini importance measures were computed for each input feature with the testing data. We found that sea ice volume is most important earlier in the season (June) and sea ice extent became a more important predictor closer to September. Results from this study provide insight into how feature importance changes with forecast length and illustrates differences between observational data and simulated Earth system data.

We have additionally performed a global sensitivity analysis (GSA) using a fully coupled ultra-low resolution configuration E3SM. To our knowledge, this is the first global sensitivity analysis involving the fully-coupled E3SM Earth system model. We have found that parameter variations show significant impact on the Arctic climate state and atmospheric parameters related to cloud parameterizations are the most significant. We also find significant interactions between parameters from different components of E3SM. The results of this study provide invaluable insight into the relative importance of various parameters from the sea ice, atmosphere and ocean components of the E3SM (including cross-component parameter interactions) on various Arctic-focused quantities of interest (QOIs).

## **ACKNOWLEDGMENT**

This research was supported by Sandia Earth System Research Foundation Laboratory Directed Research and Development funds. Thanks to the following Sandia researchers who provided important contributions to the proposal and initial phases of research for this project: Eric Ackerman, Ray Bambha, Jennifer Frederick, Jasper Hardesty, Anastasia Ilgen, and Cosmin Safta and to Mike Parks for his project management.

## CONTENTS

1. Introduction .....	11
2. Observational and Reanalysis Data .....	13
3. E3SM Model Configuration .....	14
3.1. Sea Ice Variability in E3SM DECK Simulations .....	15
3.2. E3SM Ultra-Low Configuration Tuning .....	19
3.3. E3SM Medium-Low Configuration Tuning .....	23
4. Data Model and Feature Analysis .....	25
4.1. Overview of past work .....	25
4.2. Data and Methods .....	27
4.3. Feature Importance Results .....	29
5. Global Sensitivity Analysis .....	29
5.1. Overview of Past Work .....	30
5.2. Methodology .....	31
5.2.1. Design of Global Sensitivity Study .....	32
5.2.2. Variance-Based Global Sensitivity Analysis .....	33
5.2.3. GSA Workflow .....	36
5.3. Results from Ultra-Low (ne4) Resolution (ULR) .....	37
5.4. Results from Medium-Low (ne11) Resolution (MLR) .....	43
5.5. Comparison to Other Work .....	45
5.6. Significance .....	49
6. Enrichment Analysis for Simulation Error Status .....	49
7. Conclusion .....	51
References .....	79
Appendix A. Ultra-Low Sensitivity Runs .....	86
Appendix B. Instructions for Building and Submitting Ultra-Low Sensitivity Simulations on Skybridge .....	90
Appendix C. Outcomes and Impacts .....	92
C.1. Publications .....	92
C.2. Presentations .....	92
C.3. Posters .....	92

## LIST OF FIGURES

Figure 3-1.	Ultra-low resolution grid for atmosphere (a), medium-low resolution atmosphere grid (b), and ocean grid used in both ultra-low and medium-low configurations (c) . . . . .	15
Figure 3-2.	September sea ice extent (top) and volume (bottom) from E3SM pre-industrial control simulation in black. NSIDC satellite data in red (top) and PIOMAS reanalysis in red (bottom) from the observational record shown for comparison. . . . .	16
Figure 3-3.	September sea ice concentration, average over 500 years (left), average over 10 years with largest sea ice extent (middle), average over 10 years with smallest sea ice extent (right). . . . .	16
Figure 3-4.	Seasonal cycle for sea ice extent averaged over 500-year pre-industrial control simulation and 40-year NSIDC satellite record (a). Seasonal cycle for sea ice volume averaged over 500-year pre-industrial control simulation and 40-year PIOMAS reanalysis data from 1979-2018 (b). In both plots the solid lines represent the seasonal average and the shading represents the standard deviation. . . . .	17
Figure 3-5.	September sea ice extent (top) and volume (bottom) from E3SM historical ensembles, average in black and shading is standard deviation. E3SM 1% $CO_2$ simulation in blue. NSIDC satellite data in red (top) and PIOMAS reanalysis in red (bottom) from the observational record shown for comparison. . . . .	18
Figure 3-6.	Global surface temperature versus sea ice extent (left) and sea ice volume (right) trends. Trends from observational and reanalysis data from NSDIC and NCEP are shown blue, trends from the five historical E3SM ensembles are in red, and 35-year overlapping pseudo-ensembles from the pre-industrial control simulation are shown in black. . . . .	20
Figure 3-7.	Polar sea surface temperature versus sea ice extent (left) and sea ice volume (right) trends. Trends from observational and reanalysis data from NSDIC and NCEP are shown blue, trends from the five historical E3SM ensembles are in red, and 35-year overlapping pseudo-ensembles from the pre-industrial control simulation are shown in black. . . . .	20
Figure 3-8.	Yearly averaged global surface air temperature ( $^{\circ}C$ ) (a), yearly averaged net flux at top of atmosphere ( $W/m^2$ ) (b), September sea ice extent ( $10^6 km^2$ ) (c), and yearly averaged sea ice extent ( $10^6 km^2$ ) (d). Blue line from the 500-year E3SM pre-industrial control simulation and red line from 500-year baseline spin-up. For comparison NSIDC satellite observations are shown in black in (c) and (d) and NCEP atmospheric reanalysis data for surface temperature is shown in black in (a). . . . .	21
Figure 3-9.	Surface temperature ( $^{\circ}C$ ) (a), precipitation (b) top of atmosphere flux ( $W/m^2$ ) (c) for years 300-500 of the ultra-low resolution pre-industrial control spinup and for the 1 degree standard resolution pre-industrial control run. . . . .	22
Figure 3-10.	Zonal temperature ( $^{\circ}C$ ) (a), and zonal winds ( $m/s$ ) for years 300-500 of the ultra-low resolution pre-industrial control spinup and for the 1 degree standard resolution pre-industrial control run. . . . .	23

Figure 3-11.	Yearly averaged global surface air temperature ( $^{\circ}\text{C}$ ) (a), yearly averaged net flux at top of atmosphere ( $\text{W}/\text{m}^2$ ) (b), September sea ice extent ( $10^6\text{km}^2$ ) (c), and yearly averaged sea ice extent ( $10^6\text{km}^2$ ) (d). Blue line from the 500-year E3SM pre-industrial control simulation and red line from 180-year branch run using Golaz <i>et al.</i> tuning values. For comparison NSIDC satellite observations are shown in black in (c) and (d) and NCEP atmospheric reanalysis data for surface temperature is shown in black in (a) . . . . .	24
Figure 3-12.	Yearly averaged global surface air temperature ( $^{\circ}\text{C}$ ) (a), yearly averaged net flux at top of atmosphere ( $\text{W}/\text{m}^2$ ) (b), September sea ice extent ( $10^6\text{km}^2$ ) (c), and yearly averaged sea ice extent ( $10^6\text{km}^2$ ) (d). Blue line from the 500-year E3SM pre-industrial control simulation and red line from 180-year branch run using Golaz <i>et al.</i> tuning values. For comparison NSIDC satellite observations are shown in black in (c) and (d) and NCEP atmospheric reanalysis data for surface temperature is shown in black in (a). . . . .	26
Figure 4-1.	Comparison between E3SM simulations and NCEP observational data used in the data-driven model for (a) September sea ice extent, (b) June sea ice volume, (c) June Arctic surface air temperature, (d) June Arctic sea surface temperature, (e) June Arctic surface pressure, (f) June Arctic specific humidity, (g) June Arctic surface longwave flux, (h) June Arctic total cloud fraction. . . . .	53
Figure 4-2.	Box plot of June feature importance values for random 13 year predictions generated by 385 models. The $\overline{R^2}$ , anomaly correlation coefficient ( $\overline{ACC}$ ), and mean absolute error ( $\overline{MAE}$ ) is displayed towards the top. The blue line crossing each dataset is the importance of an included random variable in each feature set. . . . .	54
Figure 5-1.	“Spaghetti plots” of first six QOIs from Table 5-2 for $\text{ne4}$ runs that made it to year 100. The baseline run is distinguished from the others by the red markers. . . . .	55
Figure 5-2.	“Spaghetti plots” of last five QOIs from Table 5-2 for $\text{ne4}$ runs that made it to year 100. The baseline run is distinguished from the others by the red markers . . . . .	56
Figure 5-3.	Box plots of first six QOIs for perturbed $\text{ne4}$ runs of 100 years, averaged over the last 50 years. . . . .	57
Figure 5-4.	Box plots of last five QOIs for perturbed $\text{ne4}$ runs of 100 years, averaged over the last 50 years. . . . .	58
Figure 5-5.	“Spaghetti plots” of Arctic sea ice extent and Arctic sea ice volume for $\text{ne4}$ runs that made it to 100 years as a function of the month. The maximum/minimum Arctic sea ice extent occur in general April/October, respectively, and the maximum/minimum Arctic sea ice volume occur in general in May/October, respectively. The baseline run is distinguished from the others by the red markers. . . . .	59
Figure 5-6.	GSA results for $\text{ne4}$ resolution: main effects, total effects and Sobol indices for Sea Ice Extent (SIE) QOI calculated annually and by season . . . . .	60
Figure 5-7.	GSA results for $\text{ne4}$ resolution: main effects, total effects and Sobol indices for Sea Ice Volume (SIV) QOI calculated annually and by season . . . . .	61
Figure 5-8.	GSA results for $\text{ne4}$ resolution: main effects, total effects and Sobol indices for Sea Surface Temperature Averaged Over 60-90 $^{\circ}$ (SST) QOI calculated annually and by season . . . . .	62

Figure 5-9.	GSA results for ne4 resolution: main effects, total effects and Sobol indices for Surface Temperature Averaged Over 60-90° (TS) QOI calculated annually and by season . . . . .	63
Figure 5-10.	GSA results for ne4 resolution: main effects, total effects and Sobol indices for Specific Humidity Averaged Over 60-90° (QS) QOI calculated annually and by season . . . . .	64
Figure 5-11.	GSA results for ne4 resolution: main effects, total effects and Sobol indices for Low Cloud Coverage Averaged Over 60-90° (CLDLOW) QOI calculated annually and by season . . . . .	65
Figure 5-12.	GSA results for ne4 resolution: main effects, total effects and Sobol indices for Snow Precipitation Averaged Over 60-90° (PRECSL) QOI calculated annually and by season . . . . .	66
Figure 5-13.	GSA results for ne4 resolution: main effects, total effects and Sobol indices for Net Longwave Ice Ocean Surface Radiation Averaged Over 60-90° (FLNS) QOI calculated annually and by season . . . . .	67
Figure 5-14.	GSA results for ne4 resolution: main effects, total effects and Sobol indices for Sea Level Pressure Over Beaufort High (BH) QOI calculated annually and by season . . . . .	68
Figure 5-15.	GSA results for ne4 resolution: main effects, total effects and Sobol indices for Sea Level Pressure Over Aleutian Low (AL) QOI calculated annually and by season . . . . .	69
Figure 5-16.	GSA results for ne4 resolution: main effects, total effects and Sobol indices for Sea Level Pressure Over Siberian High (SH) QOI calculated annually and by season . . . . .	70
Figure 5-17.	“Spaghetti plots” of first six QOIs from Table 5-2 for ne4 and ne11 baseline runs. . . . .	71
Figure 5-18.	“Spaghetti plots” of last five QOIs from Table 5-2 for ne4 and ne11 baseline runs. . . . .	72
Figure 5-19.	“Spaghetti plots” of first six QOIs from Table 5-2 for ne11 runs that made it to year 75. The baseline run is distinguished from the others by the red markers. . . . .	73
Figure 5-20.	“Spaghetti plots” of last five QOIs from Table 5-2 for ne11 runs that made it to year 75. The baseline run is distinguished from the others by the red markers . . . . .	74
Figure 5-21.	“Spaghetti plots” of first six QOIs from Table 5-2 for ne4 and ne11 run #65, with baseline QOIs subtracted. . . . .	75
Figure 5-22.	“Spaghetti plots” of last five QOIs from Table 5-2 for ne4 and ne11 run #65, with baseline QOIs subtracted. . . . .	76
Figure 5-23.	Box plots of first six QOIs for perturbed ne11 runs of 75 years, averaged over the last 25 years. . . . .	77
Figure 5-24.	Box plots of last five QOIs for perturbed ne11 runs of 75 years, averaged over the last 25 years. . . . .	78
Figure A-1.	Parameter values and number of simulation years for ultra-low resolution E3SM sensitivity analysis, part 1. . . . .	86
Figure A-2.	Parameter values and number of simulation years for ultra-low resolution E3SM sensitivity analysis, part 2. . . . .	87

Figure A-3. Parameter values and number of simulation years for ultra-low resolution E3SM sensitivity analysis, part 3. ....	88
Figure A-4. Parameter values and number of simulation years for ultra-low resolution E3SM sensitivity analysis, part 4. ....	89

## LIST OF TABLES

Table 2-1. Sea ice, atmosphere, and ocean observational and reanalysis data fields, description and source. ....	14
Table 3-1. Sea ice extent and volume sensitivity computed as the linear least squares fit versus September sea ice yearly global average surface temperature. ....	19
Table 3-2. Sea ice volume sensitivity computed as the linear least squares fit of September sea ice yearly global average surface temperature and yearly polar sea surface temperature. ....	19
Table 3-3. Default atmospheric parameter values for ULR configuration and corresponding values from Golaz <i>et al.</i> [22]. ....	22
Table 3-4. Ultra-low resolution E3SM runs to investigate atmospheric parameter tunings. Branch runs were started from year 500 of the baseline simulation. ....	23
Table 5-1. Global sensitivity analysis parameters. ....	32
Table 5-2. Global sensitivity analysis quantities of interest (QOIs). ....	33
Table 5-3. Table of correlation coefficients between the eleven QOIs considered (Table 5-2), averaged annually over the last 50 years, for ne4 runs that made it to year 100. Large positive correlation coefficients ( $\geq 0.85$ ) are colored blue. Large negative correlation coefficients ( $\leq -0.85$ ) are colored yellow. ....	39
Table 5-4. Table of correlation coefficients between the eleven QOIs considered (Table 5-2), averaged during the winter season (January–March) over the last 50 years, for ne4 runs that made it to year 100. Large positive correlation coefficients ( $\geq 0.85$ ) are colored blue. Large negative correlation coefficients ( $\leq -0.85$ ) are colored yellow. ....	40
Table 5-5. Table of correlation coefficients between the eleven QOIs considered (Table 5-2), averaged during the spring season (April–June) over the last 50 years, for ne4 runs that made it to year 100. Large positive correlation coefficients ( $\geq 0.85$ ) are colored blue. Large negative correlation coefficients ( $\leq -0.85$ ) are colored yellow. ....	40
Table 5-6. Table of correlation coefficients between the eleven QOIs considered (Table 5-2), averaged during the summer season (July–September) over the last 50 years, for ne4 runs that made it to year 100. Large positive correlation coefficients ( $\geq 0.85$ ) are colored blue. Large negative correlation coefficients ( $\leq -0.85$ ) are colored yellow. ....	41
Table 5-7. Table of correlation coefficients between the eleven QOIs considered (Table 5-2), averaged during the autumn season (October–December) over the last 50 years, for ne4 runs that made it to year 100. Large positive correlation coefficients ( $\geq 0.85$ ) are colored blue. Large negative correlation coefficients ( $\leq -0.85$ ) are colored yellow. ....	41

Table 5-8.	Table of correlation coefficients between the eleven QOIs considered (Table 5-2), averaged annually over the last 25 years, for ne11 runs that made it to year 75. Large positive correlation coefficients ( $\geq 0.85$ ) are colored blue. Large negative correlation coefficients ( $\leq -0.85$ ) are colored yellow. . . . .	45
Table 5-9.	Table of correlation coefficients between the eleven QOIs considered (Table 5-2), averaged during the winter season (January–March) over the last 25 years, for ne11 runs that made it to year 75. Large positive correlation coefficients ( $\geq 0.85$ ) are colored blue. Large negative correlation coefficients ( $\leq -0.85$ ) are colored yellow. . . . .	46
Table 5-10.	Table of correlation coefficients between the eleven QOIs considered (Table 5-2), averaged during the spring season (April–June) over the last 25 years, for ne11 runs that made it to year 75. Large positive correlation coefficients ( $\geq 0.85$ ) are colored blue. Large negative correlation coefficients ( $\leq -0.85$ ) are colored yellow. . . . .	46
Table 5-11.	Table of correlation coefficients between the eleven QOIs considered (Table 5-2), averaged during the summer season (July–September) over the last 25 years, for ne11 runs that made it to year 75. Large positive correlation coefficients ( $\geq 0.85$ ) are colored blue. Large negative correlation coefficients ( $\leq -0.85$ ) are colored yellow. . . . .	47
Table 5-12.	Table of correlation coefficients between the eleven QOIs considered (Table 5-2), averaged during the autumn season (October–December) over the last 25 years, for ne11 runs that made it to year 75. Large positive correlation coefficients ( $\geq 0.85$ ) are colored blue. Large negative correlation coefficients ( $\leq -0.85$ ) are colored yellow. . . . .	47
Table 6-1.	Parameter-bin error pairing for pairings with corrected p-values less than 0.05. .	51

## 1. INTRODUCTION

The Arctic is warming at more than twice the rate of the rest of the globe [1, 2]. Arctic physical and biological systems are strongly coupled and feedbacks may be driving the Arctic to tipping events that could have critical downstream impacts for the rest of the world [43, 44, 7]. Permafrost thaw is resulting in infrastructure damage and coastal erosion and will eventually lead to significant greenhouse gas release [55, 65]. Melt of the Greenland ice sheet will lead to global sea level rise with risks to coastal infrastructure [25]. Sea ice loss will lead to increased maritime activity with the potential to lead to geopolitical conflict as more nations vie for access to the region [69]. In addition, there is evidence that loss of sea ice and Arctic warming lead to changes in mid-latitude weather and precipitation [12, 13, 15]. Both sea ice and land ice melt may also lead to ocean current disruptions due to freshwater flux into the North Atlantic [68]. As changes progress in each of these physical systems, interactions between them as well as with the atmosphere and ocean can lead to feedbacks that accelerate the warming.

Recently, considerable research efforts have been focused on predicting critical transitions in geophysical and other complex systems [41, 7, 16, 44, 64, 45, 42, 63, 43]. Tipping events often arise due to self-amplification processes within the system dynamics, thus prioritizing accurate capture of these feedbacks. Lenton and others [43, 30]. have found that triggers of the tipping events may be the magnitude of the forcing, the rate of the forcing, or system noise; hence, data analytic techniques should be employed to track these triggers. A large committed change due to a tipping event can occur rapidly. Alternatively, a rate limitation in the system dynamics may cause the regime change to occur very slowly [41]. From a socioeconomic perspective both the timing of the point of commitment as well as the range and magnitude of the state-altering effect are important.

In the final two years of this project we have focused on analyzing sea ice in the coupled Earth system. Arctic sea ice extent has been declining over the past four decades with the most significant loss in summer months [52]. Contributing to sea ice loss is the well-known ice-albedo feedback, which is the process where highly reflective sea ice is lost reducing the surface albedo and increasing solar radiation absorption in the darker ocean water [24]. This positive feedback is counteracted by a negative feedback mechanism whereby thinner sea ice grows more quickly in response to thermodynamic forcing from the ocean and atmosphere. The interactions of these feedbacks have led researchers to consider whether sea ice loss is a tipping point in the global climate [43]. In simplified low-order single column models of Arctic sea ice, parameter regimes have been found where a bifurcation occurs as the Arctic transitions to a seasonally ice-free state and where a more robust bifurcation point occurs with the loss of winter sea ice [3, 29]. However, when researchers have looked at more detailed models of sea ice, including the types of models used in coupled Earth system models, the bifurcations associated with a transition to a seasonal ice-free state are no longer observed. In particular, Wagner and Eisenman found that instabilities in low-order models are due to lack of meridional heat transport and the seasonal cycle of solar forcing, explaining why general circulation models display a smooth loss of Arctic sea ice that is reversible in contrast to the simplified models [78].

Although the transition to a seasonally ice-free state in the Arctic may be smooth and reversible, observational data have shown that the loss of Arctic sea ice is happening rapidly and that this loss will have a significant influence on Arctic coastal systems and global climate. Recent research

has suggested that Arctic sea ice loss contributes to mid-latitude weather changes, including increases in winter storms and drought [12, 13, 15]. Sea ice loss can also impact Atlantic Meridional Overturning Circulation [68] and has significant geopolitical and economic implications due to increasing shipping activity [69].

Accurate seasonal predictions of sea ice minimum extent and long-term estimates of timing for a seasonally ice-free Arctic depend on a better understanding of the factors influencing sea ice dynamics and variation in this strongly coupled system. This includes loss due to greenhouse gas forcing and the superimposed internal variability of the complex Earth system. Studies have found a linear relationship between sea ice loss and global average surface temperature in both observational data and simulation data with most predictions indicating that the Arctic will be seasonally ice free by mid-century [79, 71, 52]. The correlation is generally explained by a common dependency of temperature and sea ice concentration on green house gas concentration. However, there is evidence that internal variability may strongly influence the timing for an ice-free Arctic. Screen and Deser show that changes in the Interdecadal Pacific Oscillation (IPO) can shift this timing by seven years [67]. Ding *et al.* conclude through a fingerprint analysis with the Large Ensemble Community Project (LENS) data set that internal variability contributes between 40-50 % of the recent decline in Arctic sea ice [17]. Additionally, there is no consensus in the literature on the dominant factors impacting internal variability and sea ice loss and this continues to be an active area of research. Potential drivers include atmospheric temperature fluctuations [53], decadal sea surface temperature changes driving convection and teleconnections [47], and ocean mixed layer processes [54].

Accurate long-term predictions from coupled Earth systems models are required to estimate Arctic sea ice changes and find the dominant drivers in that loss. Previous generations of climate models have on average underestimated sea ice loss [72] and it has been shown that among CMIP5 models that matched the sea ice trends, the atmospheric warming has been overpredicted [62]. Given of the complexity of the coupled Earth system and the feedbacks in the warming Arctic, there are opportunities to investigate these interactions in more detail with advanced approaches that include investigating dominant drivers of Arctic change using machine learning techniques and evaluating how the parametric uncertainty in coupled models impacts projections.

Under this project we have investigated the main drivers of sea ice loss using both observational data and simulation data from the Energy Exascale Earth System Model (E3SM). In Section 2 we describe the observational data and reanalysis products used in our analyses. In Section 3, we provide a description of E3SM and an initial analysis of sea ice behavior in the standard resolution simulation data for CMIP6. We also introduce in this section an ultra-low and medium-low resolution of the E3SM, both of which are used for our global sensitivity analysis, and provide results from spin-up and tuning runs using these models. In Section 4 we introduce a data-driven model for September sea ice extent and provide results of a feature analysis to investigate the most important factors driving sea ice extent in both observations and E3SM. In Section 5 we discuss a global sensitivity analysis for the fully coupled E3SM using an ensemble set of successful perturbed simulations at two low-resolution configurations. This analysis allows us to identify which parameters from the fully-coupled E3SM have the greatest influence on several Arctic-focussed quantities of interest. In Section 6, we attempt to extract some additional information from the aforementioned low-resolution simulations by performing an enrichment analysis aimed at identi-

fyng if any of the model parameters in our experiments are correlated with whether the simulation completed successfully. Finally, we provide some overall conclusions in Section 7. Outcomes of this project including presentation, posters and publications are listed in Appendix C.

## 2. OBSERVATIONAL AND REANALYSIS DATA

Observational data and reanalysis data products were used in our analysis as comparison for simulation data from the Energy Exascale Earth System Model (E3SM) (Sections 3.1 and 3.2) as well as for training the machine learning model for sea ice extent (Section 4). For Arctic sea ice coverage, monthly sea ice extent values from the National Snow and Ice Data Center (NSIDC) were used [19]. The monthly extent is derived from gridded daily passive microwave satellite observations of sea ice concentration at 25 km resolution [48, 56]. The concentration data set provides the fraction of ice coverage in each cell and extent is computed as the sum of grid cell areas with at least 15% sea ice coverage. The data set runs from 1979 through the present and was downloaded March, 2019.

For sea ice volume we utilize data from the Pan-Arctic Ice Ocean Modeling and Assimilation System (PIOMAS) [66]. PIOMAS is a coupled sea ice and ocean numerical model that assimilates daily sea ice concentration data from NSIDC as well as sea surface temperatures from National Center for Environmental Protection (NCEP). The model computes an estimate of sea ice thickness, which is used to derive an average value for monthly total Arctic sea ice volume. The data are available from 1979 to the present and we used data through 2018 for our analysis.

For atmospheric data we use the NCEP-DOE Reanalysis II [37]. We accessed data for 4-times daily values on a T62 Gaussian grid for 2 meter air temperature, surface pressure, surface downward longwave flux, 10 meter meridional wind, and 10 meter zonal wind. We additionally downloaded monthly average values of 2 meter specific humidity, surface precipitation rate, and cloud fraction also on a T62 Gaussian grid. In all cases we used data from years 1979-2018. For our analysis we typically used either a global area average of the field obtained by computing the area sum over the entire grid or an Arctic area average obtained by computed the area sum over all cells with latitude greater than 60. The National Oceanic and Atmospheric Administration (NOAA) Extended Reconstructed Sea Surface Temperature (ERSSTv4) dataset, which uses historical ship and buoy sea surface temperature (SST) measurements to reconstruct a monthly SST field on a 2 degree global latitude longitude grid is used for sea SST data [31]. As with the atmospheric reanalysis data, we compute either a global area average or an Arctic area average from the gridded product over the years 1979-2018 for our analysis.

A summary of the observational and data fields we used in this project including a short name, description and data source are provided in Table2-1.

Name	Description	Source
SIE	sea ice extent	<a href="ftp://sidads.colorado.edu/DATASETS/NOAA/G02135/">ftp://sidads.colorado.edu/DATASETS/NOAA/G02135/</a> [48]
SIV	sea ice volume	<a href="http://apl.uw.edu/research/projects/arctic-sea-ice-volume-anomaly/">apl.uw.edu/research/projects/arctic-sea-ice-volume-anomaly/</a> [66]
TS	surface air temperature	<a href="http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis2.html">www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis2.html</a> [37]
FLWS	downward longwave flux at surface	<a href="http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis2.html">www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis2.html</a> [37]
PS	surface air pressure	<a href="http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis2.html">www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis2.html</a> [37]
UWIND	Eastward wind	<a href="http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis2.html">www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis2.html</a> [37]
VWIND	Northward wind	<a href="http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis2.html">www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis2.html</a> [37]
CLT	total cloud cover percentage	<a href="http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis2.html">www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis2.html</a> [37]
SSH	near-surface specific humidity	<a href="http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis2.html">www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis2.html</a> [37]
SST	sea surface temperature	<a href="ftp://ftp.ncdc.noaa.gov/pub/data/cmb/ersst/v4">ftp://ftp.ncdc.noaa.gov/pub/data/cmb/ersst/v4</a> [31]

**Table 2-1 Sea ice, atmosphere, and ocean observational and reanalysis data fields, description and source.**

### 3. E3SM MODEL CONFIGURATION

Version 1 of the Energy Exascale Earth System Model (E3SM)<sup>1</sup> was used to investigate changes in Arctic sea ice in response to internal variability related to ocean and atmosphere modes as well as in response to perturbations in the model parameters. E3SM consists of component models for atmosphere, ocean, ice, land, and river transport. The E3SM Atmosphere Model (EAM) has a spectral element dynamical core discretized on a cubed sphere grid using 72 vertical levels. The standard resolution E3SM configuration uses a 1 degree grid for both EAM and the E3SM Land Model (ELM), which corresponds to approximately 110 km at the equator. The ocean and sea ice models are based on the Model Prediction Across Scales (MPAS) framework. MPAS-ocean uses a finite volume discretization on an unstructured Voronoi grid, which is shared with MPAS-Sea Ice. At the standard resolution the ocean and sea ice grid has a resolution varying between 60 km at midlatitudes and 30 km at the poles. The Model for Scale Adaptive River Transport (MOSART) has a resolution of 50 km.

A detailed description of E3SM version 1 and simulation skill under the Diagnostic, Evaluation and Characterization of Klima (DECK) simulations for the Sixth Coupled Model Intercomparison Project (CMIP6) are provided in Golaz *et al.* [22]. The set of standard resolution CMIP6 simulation data, produced by the E3SM development team, is available for downloading<sup>2</sup> and was used in our analysis. Details of CMIP6 are given in [18]. In particular, we have used the following data sets

- 500-year pre-industrial control (“piControl”) simulation,
- 165-year historical simulation with five ensemble members,
- 150-year 1% CO<sub>2</sub> increase simulation.

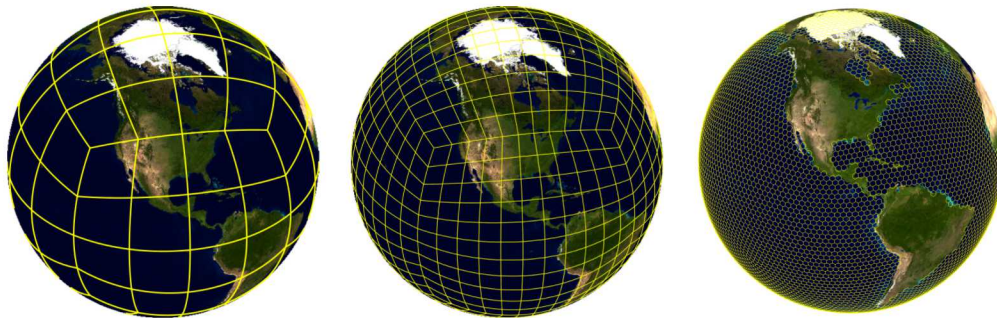
A summary of Arctic sea ice variability and trends from the E3SM v1 DECK data sets is provided in Section 3.1. The five historical ensembles are used as training data in the machine learning feature analysis described in Section 3.2.

In addition to the data from the E3SM v1 released DECK simulations, we have generated our own data sets using an ultra-low resolution (ULR) and medium-low resolution (MLR) configuration of E3SM v1. We chose an ULR configuration that would provide a computationally tractable way

<sup>1</sup>[www.e3sm.org](http://www.e3sm.org)

<sup>2</sup><https://e3sm.org/data/get-e3sm-data/released-e3sm-data/v1-1-deg-data-cmip6/>

to generate large numbers of ensemble runs to explore the parameter space in the coupled model. This ULR configuration has a grid resolution of approximately 7.5 degrees (ne4) for EAM and ELM and 240 km or approximately 2.2 degrees for MPAS-ocean and MPAS-Sea Ice. To develop robust predictive understanding of sea ice stability, we additionally performed computations with a medium-low resolution (MLR) configuration that included a higher resolution atmosphere model (ne11) with resolution approximately 2.7 degrees. The ocean model grid remained the in both the ULR and MLR configurations. Importantly, key dynamics, such as baroclinic instability, are resolved in this configuration. Baroclinic instability dynamics give rise to large-scale and mesoscale motion in the atmosphere and ocean. Plots of the grids in the low resolution configurations are provided in Figure (3-1).



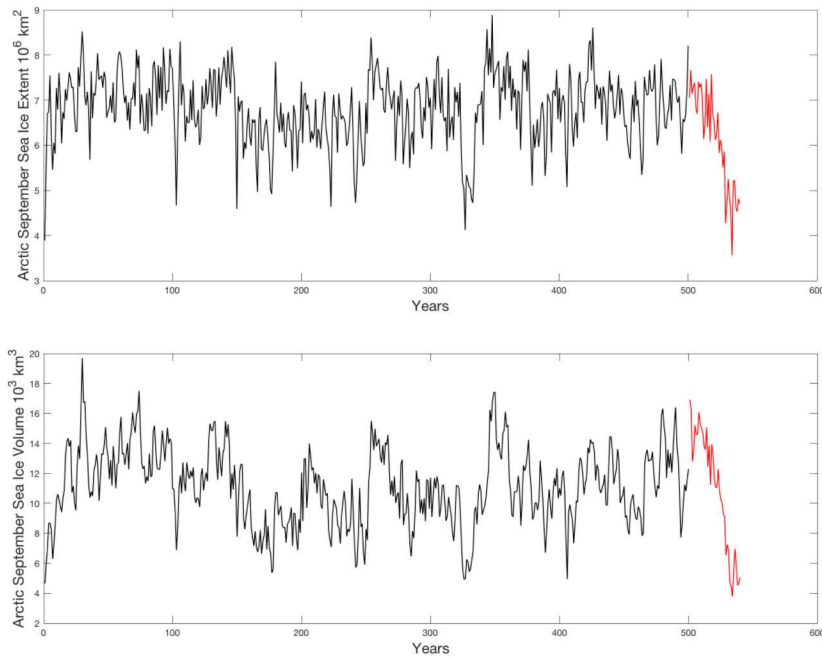
**Figure 3-1** Ultra-low resolution grid for atmosphere (a), medium-low resolution atmosphere grid (b), and ocean grid used in both ultra-low and medium-low configurations (c)

E3SM v1 at standard resolution has been scientifically validated with tuning parameters chosen to optimize top-of-atmosphere radiative equilibrium, cloud and aerosol optical depths, sea and land ice extent, and circulations of the atmosphere and ocean. Therefore, this configuration is used as a reference for our E3SM ULR and MLR simulations for comparing quantities of interest.

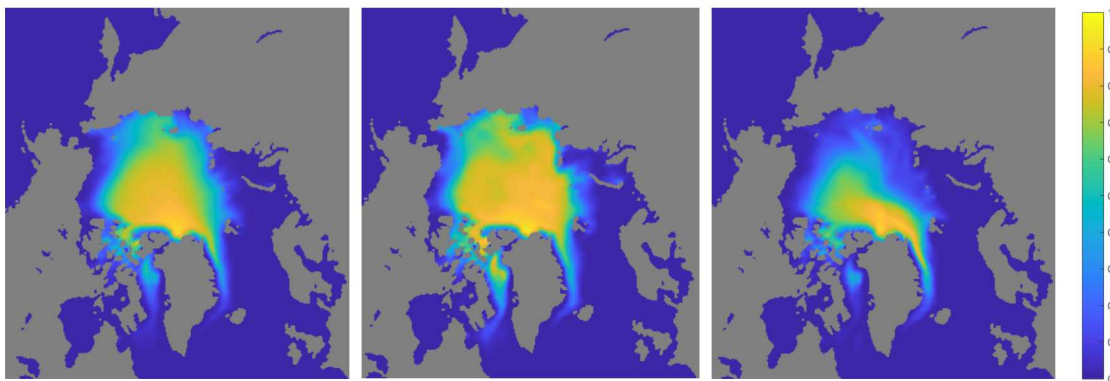
### **3.1. Sea Ice Variability in E3SM DECK Simulations**

Here we provide a preliminary evaluation of September sea ice extent and volume trends in a subset of the DECK simulations including the pre-industrial control simulation, the five historical ensembles, and the 1% CO<sub>2</sub> increase simulation. The 500-year pre-industrial control simulation provides a view of the range of variability of sea ice extent and the historical and 1% CO<sub>2</sub> simulations provide a view of the sensitivity of Arctic sea ice cover to greenhouse gas forcing and Arctic warming. We focus primarily on the September sea ice extent and volume as changes in these quantities are most important for predictions of seasonally ice-free conditions in the Arctic. In the 500-year pre-industrial control simulation the September sea ice extent shows considerable yearly and decadal variability (Figure 3-2). NSIDC observational data from the historical record is shown in red and shows that the E3SM September sea ice extent is reasonable in magnitude compared to the historical observations. The range of the variations in September sea ice extent can be seen more directly in plots of September sea ice concentration averaged over all years, the ten smallest

extent years, and the ten largest sea ice extent years (Figure 3-3). This represents variability in E3SM without anthropogenic greenhouse gas forcing.



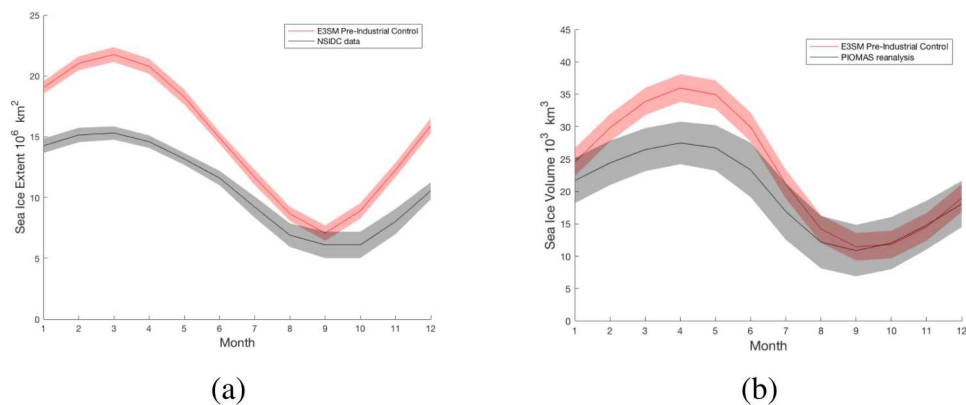
**Figure 3-2** September sea ice extent (top) and volume (bottom) from E3SM pre-industrial control simulation in black. NSIDC satellite data in red (top) and PIOMAS reanalysis in red (bottom) from the observational record shown for comparison.



**Figure 3-3** September sea ice concentration, average over 500 years (left), average over 10 years with largest sea ice extent (middle), average over 10 years with smallest sea ice extent (right).

Although E3SM closely matches the observational Arctic September sea ice extent, it exhibits a bias in winter sea ice extent and volume overestimating the values compared against NSIDC data and the PIOMAS reanalysis product. This can be seen by averaging the seasonal cycle over simulation and observation years as shown in Figure 3-4. Figure 3-4a shows the seasonal cycle for

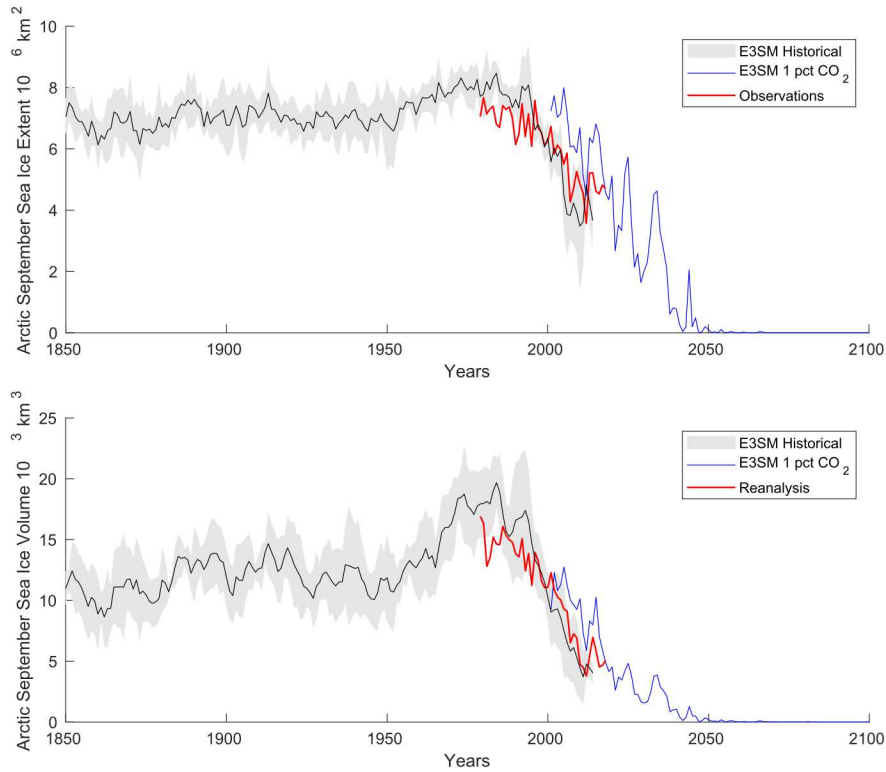
sea ice extent where the solid red line represents the 500-year average of the E3SM piControl sea ice extent and the shading represents the standard deviation. Similarly the solid black line represents the 40-year average of the NSDIC sea ice extent and the gray shading the standard deviation. E3SM has a pronounced overestimation of the sea ice extent in all months except September. The seasonal cycle for sea ice volume is shown in Figure 3-4b for a 500-year average of the E3SM piControl Arctic sea ice volume and a 40-year average of the PIOMAS reanalysis data. The E3SM seasonal cycle for volume also overestimates the volume during winter and late spring, but matches quite well in summer and fall. The variability in sea ice extent and volume as measured by standard deviation of the seasonal cycle is smaller in general for the piControl simulation. This is not surprising as the observational record does show more variability due to forced changes from greenhouse gases.



**Figure 3-4 Seasonal cycle for sea ice extent averaged over 500-year pre-industrial control simulation and 40-year NSIDC satellite record (a). Seasonal cycle for sea ice volume averaged over 500-year pre-industrial control simulation and 40-year PIOMAS reanalysis data from 1979-2018 (b). In both plots the solid lines represent the seasonal average and the shading represents the standard deviation.**

Data released from the E3SM project for CMIP6 includes five ensemble members initialized from different years of the pre-industrial control simulation and run with historical forcing from 1850 through 2014 as well as a 1%  $CO_2$  forced simulation of 150 years. September sea ice extent and volume for these simulation are shown in Figure 3-5. The plots show that the historical ensembles slightly overestimate the sea ice extent and volume loss decline of the last 30 years. The trends are investigated in more detail below. In the 1%  $CO_2$  simulation, September sea ice extent and volume go to near zero within 50 years of the start of the simulation.

Following Winton [79], we calculate the sea ice sensitivity with respect to global surface temperature (TS). Observational and reanalysis data from NSIDC, PIOMAS, and NCEP for September sea ice extent (SSIE), volume (SSIV), and TS as described in Section 2 are used in the comparison with simulations. TS is calculated from gridded observational and simulation data as an area-averaged sum of surface temperature over the entire globe. Using a linear least squares fit we compute slope and  $R^2$  of SSIE and SSIV versus TS. The slopes are computed using the final 35 years of the historical simulation to provide a direct comparison with the observational data. In the case of the 1%  $CO_2$  simulation, the first 40 years are used to compute the slope because after that point the September sea ice extent reduces to nearly zero. As shown in Table 3-1 all E3SM historical



**Figure 3-5 September sea ice extent (top) and volume (bottom) from E3SM historical ensembles, average in black and shading is standard deviation. E3SM 1%  $CO_2$  simulation in blue. NSIDC satellite data in red (top) and PIOMAS reanalysis in red (bottom) from the observational record shown for comparison.**

simulations as well as the E3SM 1%  $CO_2$  simulation show larger slopes than the observational data. This is in contrast to previous models in CMIP5, which generally show smaller sensitivity to global surface temperature [72]. The  $R^2$  values for the observations, historical simulations, and 1%  $CO_2$  simulation indicate a good fit supporting previous analyses that found strong linear correlations between sea ice extent and global surface temperature. In contrast, the pre-industrial control simulation has  $R^2$  values of 1.35 and 1.24 for SSIE and SSIV, respectively, much lower than the other simulations. This suggests that the linear relationship between TS, SST, and SSIE, SSIV are weaker in this simulation, perhaps because other sources of variability dominate in the pre-industrial control forcing case.

We also investigate sea ice sensitivity with respect to Arctic sea surface temperature (SST). Arctic sea surface temperature is computed as an area sum over the region with latitude greater than 60 degrees. Observational and reanalysis data from NSIDC, PIOMAS, and ERSSTv4 for September sea ice extent (SSIE), volume (SSIV), and SST as described in 2 are used in the comparison. As before we use a linear least squares fit to compute slope and  $R^2$  of SSIE versus SST and SSIV versus SST using the final 35 years of the historical simulation and the first 40 years of the 1%  $CO_2$  simulation. As shown in Table 3-2 all E3SM historical simulations as well as the E3SM 1%  $CO_2$  simulation show larger slopes than the observational data similar to what was seen in the SSIE sensitivity results (Table 3-1). The  $R^2$  values for the observations and historical simulations indicate a good fit and in many cases are larger than the corresponding  $R^2$  values for sensitivity

Data Source	SSIE vs Global TS		SSIV vs Global TS	
	Slope ( $10^6 km^2/K$ )	$R^2$	Slope ( $10^3 km^3/K$ )	$R^2$
Observations	-3.54	0.644	-13.7	0.712
E3SM Hist 1	-5.65	0.810	-18.3	0.788
E3SM Hist 2	-3.77	0.656	-10.6	0.449
E3SM Hist 3	-4.94	0.728	-17.1	0.695
E3SM Hist 4	-4.02	0.705	-17.2	0.766
E3SM Hist 5	-4.73	0.805	-12.4	0.690
E3SM piControl	-2.65	0.135	-8.42	0.124
E3SM 1%CO <sub>2</sub>	-4.39	0.758	-6.92	0.791

**Table 3-1 Sea ice extent and volume sensitivity computed as the linear least squares fit versus September sea ice yearly global average surface temperature.**

with respect to TS. As before the pre-industrial control simulation does not have as good a fit and interestingly neither does the 1% CO<sub>2</sub> simulation. correlations between sea ice extent and global surface temperature.

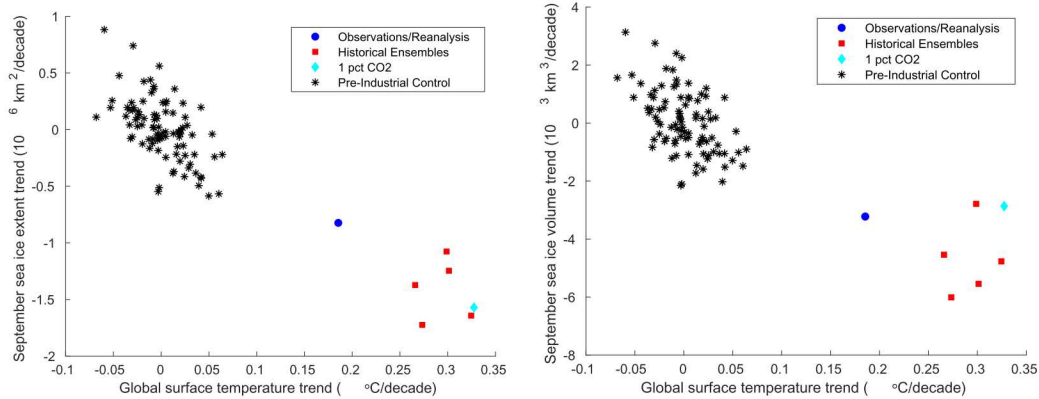
Data Source	SSIE vs Polar SST		SSIV vs Polar SST	
	Slope ( $10^6 km^2/K$ )	$R^2$	Slope ( $10^3 km^3/K$ )	$R^2$
Observations	-3.60	0.790	-13.9	0.879
E3SM Hist 1	-4.86	0.798	-15.7	0.755
E3SM Hist 2	-3.97	0.877	-11.4	0.792
E3SM Hist 3	-4.64	0.739	-14.2	0.672
E3SM Hist 4	-4.46	0.828	-17.3	0.824
E3SM Hist 5	-4.75	0.852	-11.8	0.724
E3SM piControl	-3.00	0.275	-7.24	0.142
E3SM 1%CO <sub>2</sub>	-4.47	0.571	-8.50	0.550

**Table 3-2 Sea ice volume sensitivity computed as the linear least squares fit of September sea ice yearly global average surface temperature and yearly polar sea surface temperature.**

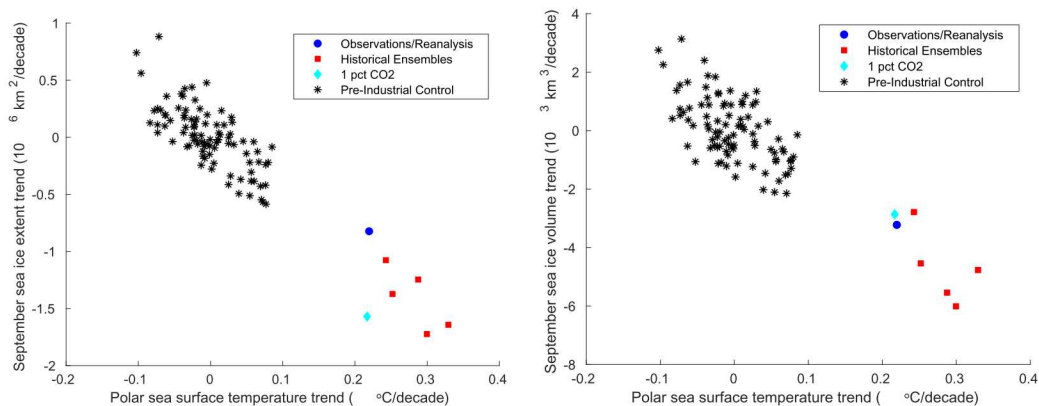
We can additionally compare trends between the E3SM simulations and observational and reanalysis data. In Figure 3-6 we plot SSIE versus TS trends and SSIV versus TS trends for observational and simulation data. In this figure we plot trends for overlapping 35-year pseudo-ensembles from the pre-industrial control simulation starting from every five years in the 500-year time series, shown as black stars. As would be expected for a run in equilibrium, these points center around zero trends in surface temperature and September sea ice extent. It is interesting, however, that the SSIE versus TS trends from observations (dark blue) and historical (red) and 1% CO<sub>2</sub> (light blue) simulations do not overlap with the trends from the piControl run, which are significantly smaller. These plots also demonstrate that the historical simulations overpredict the recent observational trends in sea ice loss. Similar results can be seen in the SSIV versus TS trend plot as well as the SSIE and SSIV versus SST plots (Figure 3-7).

### 3.2. E3SM Ultra-Low Configuration Tuning

For ULR simulations, we performed a spin-up (i.e., running the model until equilibrium state is achieved) using the pre-industrial control forcing. Our initial spin-up was a simulation of 500 years



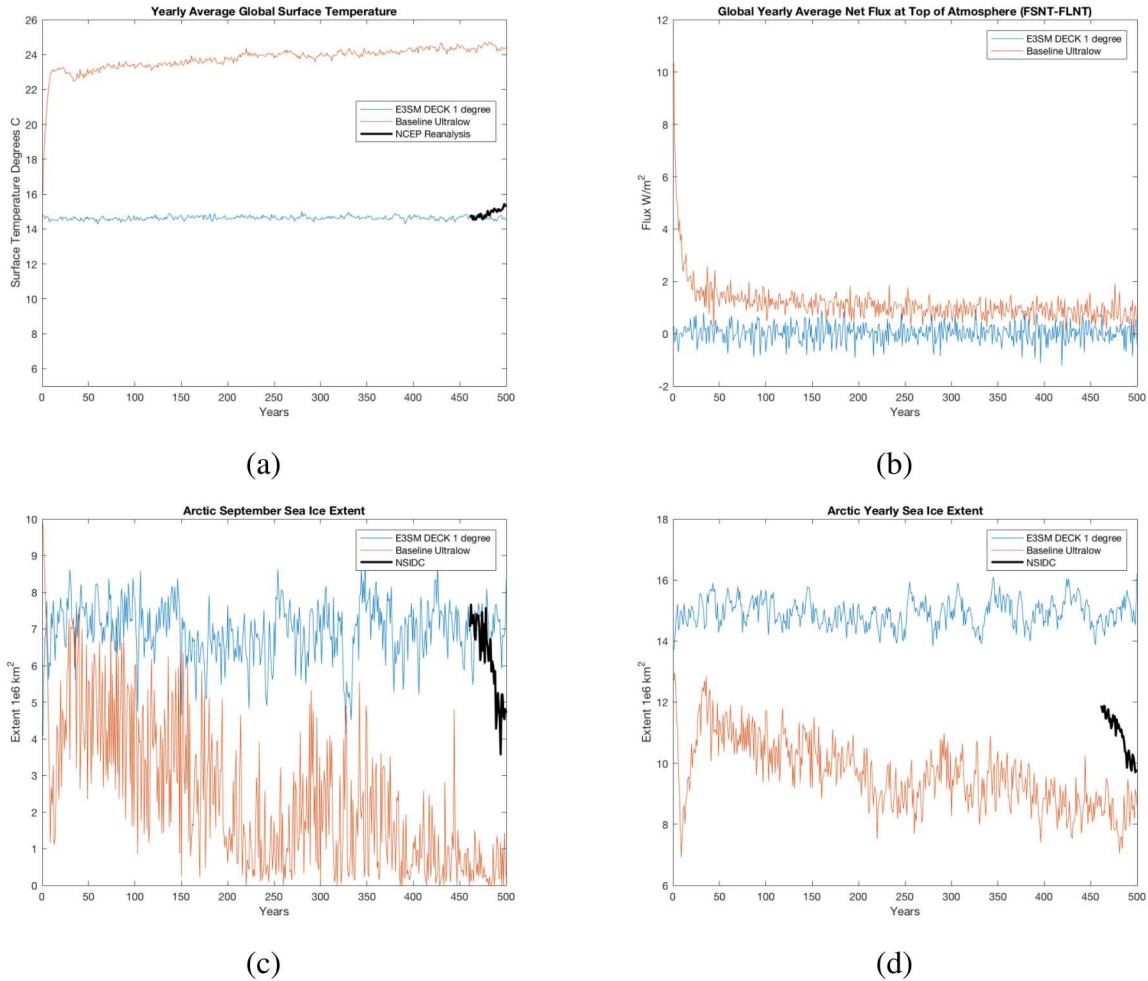
**Figure 3-6 Global surface temperature versus sea ice extent (left) and sea ice volume (right) trends. Trends from observational and reanalysis data from NSDIC and NCEP are shown blue, trends from the five historical E3SM ensembles are in red, and 35-year overlapping pseudo-ensembles from the pre-industrial control simulation are shown in black.**



**Figure 3-7 Polar sea surface temperature versus sea ice extent (left) and sea ice volume (right) trends. Trends from observational and reanalysis data from NSDIC and NCEP are shown blue, trends from the five historical E3SM ensembles are in red, and 35-year overlapping pseudo-ensembles from the pre-industrial control simulation are shown in black.**

with default parameter values. This simulation exhibited a warm bias, with surface temperature elevated, compared with observations and declining sea ice over the 500-year period. In Figure 3-8 time series plots of values from the baseline run show the warm bias is air surface temperature. Corresponding to the warm bias we see a positive net top of atmosphere flux and declining sea ice extent. It is also noteworthy that sea ice year-to-year variability is quite large in this case. It has been found in other models that variability increases as the Arctic warms and sea ice is lost [49].

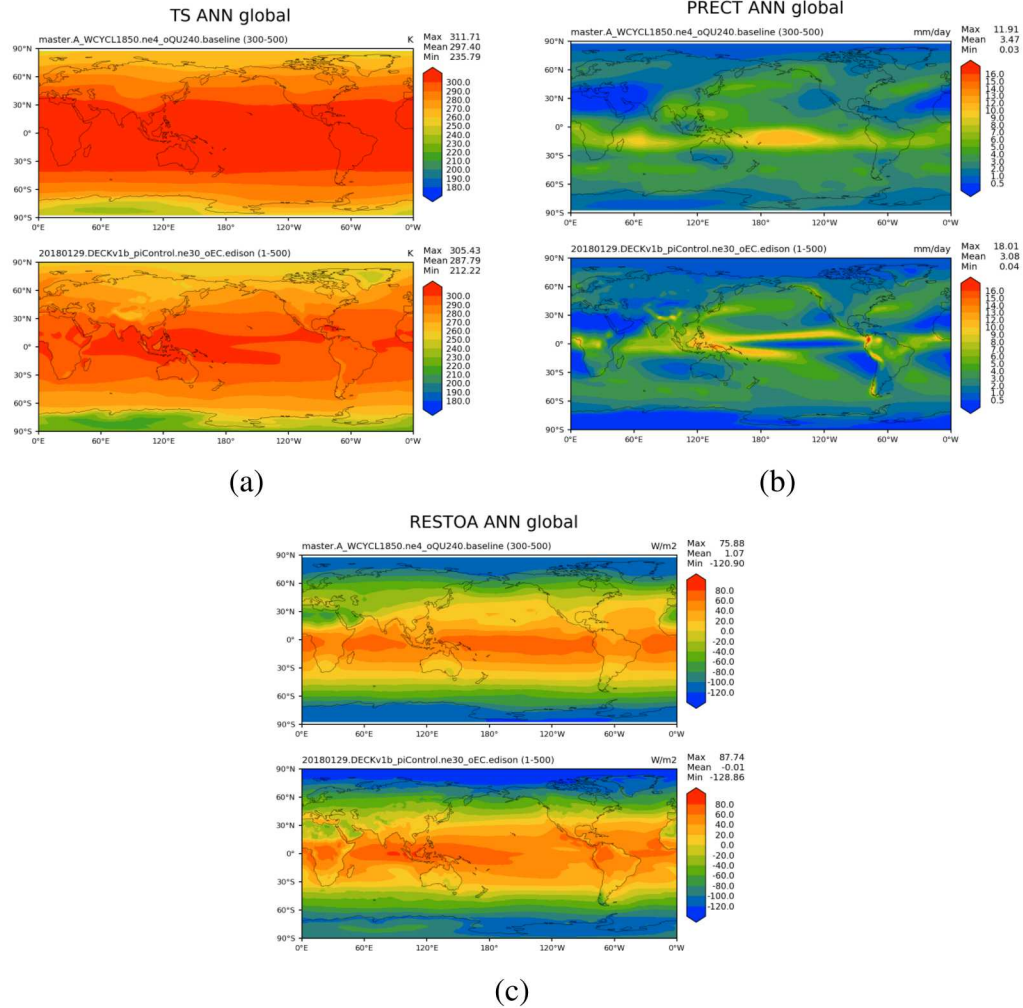
Annual climatologies have been computed for the baseline ULR simulation from the final 200 years in the spin-up to investigate spatial variation of selected fields. Plots of the annual average global surface temperature, precipitation, and top of atmosphere flux and shown in Figure 3-9 where the final 200 years of the ULR baseline spin-up is compared to the average of the 500 year pre-industrial control simulation from the 1 degree E3SM CMIP6 simulations. In Figure 3-



**Figure 3-8** Yearly averaged global surface air temperature ( $^{\circ}C$ ) (a), yearly averaged net flux at top of atmosphere ( $W/m^2$ ) (b), September sea ice extent ( $10^6 km^2$ ) (c), and yearly averaged sea ice extent ( $10^6 km^2$ ) (d). Blue line from the 500-year E3SM pre-industrial control simulation and red line from 500-year baseline spin-up. For comparison NSIDC satellite observations are shown in black in (c) and (d) and NCEP atmospheric reanalysis data for surface temperature is shown in black in (a).

10 zonal mean values of surface temperature and zonal wind are shown. These plots show that the ULR simulation is not able to capture some of the smaller features that a higher-resolution simulation can. However, the ULR simulation does capture the large-scale features of the flow providing support that the ULR simulation can provide useful information for higher-resolution modeling.

Tuning the coupled model, we want near-zero long-term average net top-of-atmosphere (TOA) energy flux, constant global average mean surface air temperature, and a reasonable absolute global mean surface air temperature. However, this is not achieved with the initial baseline computation of the ULR model. In an attempt to make the ULR model more Earth-like with respect to key QOI, we tuned the atmosphere component, investigating several changes to atmosphere parameters, as well as modified parameters to match the final tuning from the Golaz *et al.* paper [22]. Parameter



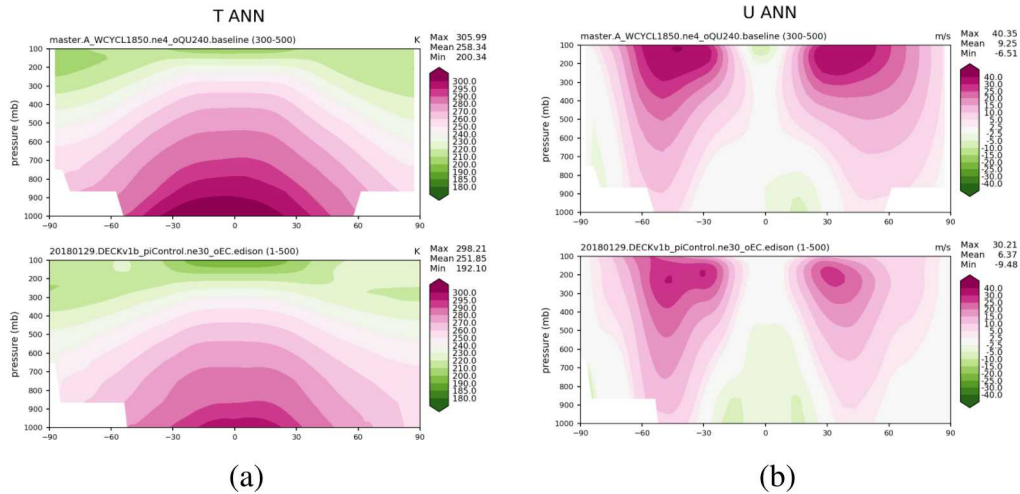
**Figure 3-9** Surface temperature ( $^{\circ}\text{C}$ ) (a), precipitation (b) top of atmosphere flux ( $\text{W}/\text{m}^2$ ) (c) for years 300-500 of the ultra-low resolution pre-industrial control spinup and for the 1 degree standard resolution pre-industrial control run.

values are given in Table 3-3.

Parameter	Baseline Value	Golaz <i>et al.</i> Value
zmconv_ke	$1.5 \times 10^{-6}$	$5.0 \times 10^{-6}$
so4_sz_thresh_icenuc	$7.53 \times 10^{-8}$	$5.0 \times 10^{-8}$
clubb_c14	1.3	1.06
cldfrc_dp1	0.045	0.045

**Table 3-3** Default atmospheric parameter values for ULR configuration and corresponding values from Golaz *et al.*[22].

In addition to the changes to Golaz *et al.* values shown in Table 3-3. We ran a series of perturbed simulations to see if we could obtain a spin-up with more realistic climate as shown in Table 3-4. We found that the Cloud Layers Unified By Binormals (CLUBB) parameter, cldfrc\_dp1, related to cloud parameterizations in EAM, has a strong impact on surface temperature as well as sea ice



**Figure 3-10 Zonal temperature ( $^{\circ}\text{C}$ ) (a), and zonal winds ( $m/s$ ) for years 300-500 of the ultra-low resolution pre-industrial control spinup and for the 1 degree standard resolution pre-industrial control run.**

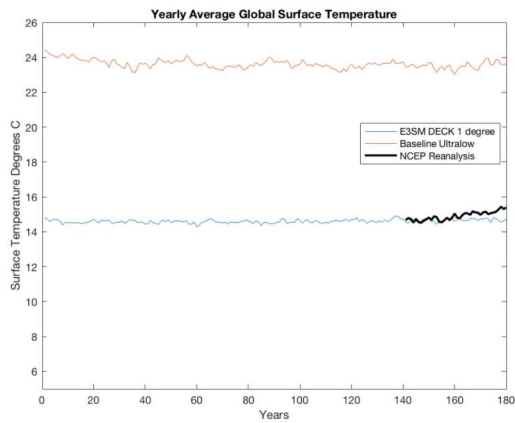
Short Case Name	Years	Description
baseline	1–500	Baseline simulation using default parameters
tunecldfrc_dp1p1	1–148	Baseline with cldfrc_dp1 = 0.1
tunecldfrc_dp1p05	1–201	Baseline with cldfrc_dp1 = 0.05
tune_agu_cldfrc_dp1_0.05	1–200	Golaz defaults with cldfrc_dp1 = 0.05
tune_agu_cldfrc_dp1_0.075	1–200	Golaz defaults with cldfrc_dp1 = 0.075
branch_tune_clubb_c141p06	501-750	Branch from baseline, with clubb_c14 = 1.06
branch_tune_Golaz	501-683	Branch from baseline, Golaz defaults
branch_tune_dp1p125	501-593	Branch from baseline, Golaz defaults with cldfrc_dp1 = 0.125
branch_tune_dp1p15	501-576	Branch from baseline, Golaz defaults with cldfrc_dp1 = 0.15
branch_tune_dp1p175	501-614	Branch from baseline, Golaz defaults with cldfrc_dp1 = 0.175
branch_tune_dp1p2	501-531	Branch from baseline, Golaz defaults with cldfrc_dp1 = 0.2
branch_tune_dp1p3	501-504	Branch from baseline, Golaz defaults with cldfrc_dp1 = 0.3
branch_tune_dp1p4	501-517	Branch from baseline, Golaz defaults with cldfrc_dp1 = 0.4

**Table 3-4 Ultra-low resolution E3SM runs to investigate atmospheric parameter tunings. Branch runs were started from year 500 of the baseline simulation.**

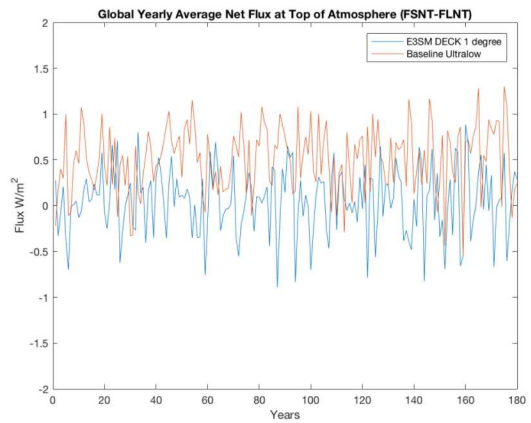
concentration and extent. However, despite numerous attempts, we were unable to reproduce the 1-degree diagnostic QOI behaviors and validated E3SM configuration with different tunings of key atmospheric parameters. In order to move forward with the sensitivity analysis, we chose to use a branch of the baseline ULR simulation with the Golaz *et al.* parameter values as the initialization. This set of parameters did result in a more realistic climate with the net top of atmosphere flux was closer to zero. Time series plots of surface temperature, top of atmosphere flux, and sea ice extent for this simulation compared with the standard 1-degree E3SM results are shown in Figure 3-11.

### 3.3. E3SM Medium-Low Configuration Tuning

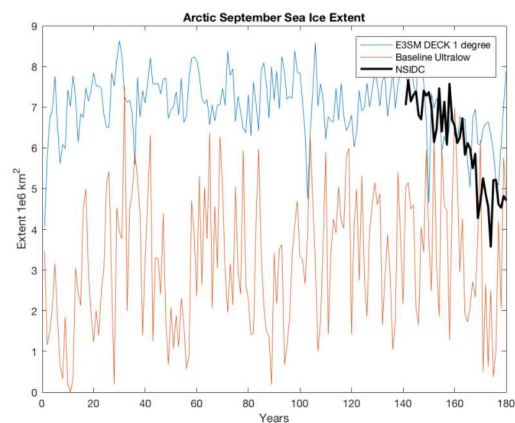
To enable multi-fidelity studies for rigorous uncertainty estimation, we began developing tunings at a “medium-low” (MLR) resolution –  $2.7(^{\circ})$  latitude, corresponding to the ne11 EAM grid. To



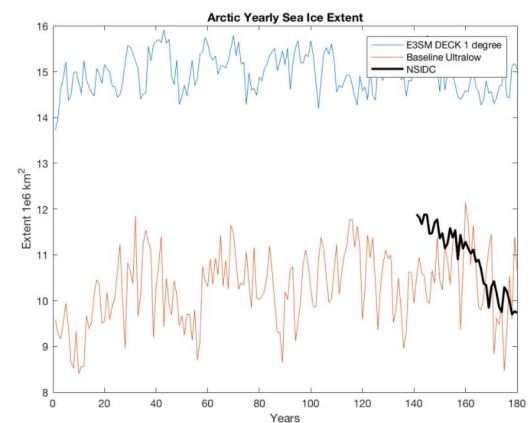
(a)



(b)



(c)



(d)

**Figure 3-11** Yearly averaged global surface air temperature ( $^{\circ}C$ ) (a), yearly averaged net flux at top of atmosphere ( $W/m^2$ ) (b), September sea ice extent ( $10^6 km^2$ ) (c), and yearly averaged sea ice extent ( $10^6 km^2$ ) (d). Blue line from the 500-year E3SM pre-industrial control simulation and red line from 180-year branch run using Golaz *et al.* tuning values. For comparison NSIDC satellite observations are shown in black in (c) and (d) and NCEP atmospheric reanalysis data for surface temperature is shown in black in (a).

establish the baseline necessary for parameter sensitivity studies described in Section 5, we tested two configurations – no tuning, and tuning as in Golaz *et al.* Our baseline ended up being the Golaz *et al.* tuning; 242 years were simulated with this tuning, and parameter sensitivity studies branched from this MLR at simulated year 200. This run had reached equilibrium by approximately simulation year 80. The pre-industrial control component set was also used for the MLR spin up. Parameter values from Golaz *et al.* from Table 3-3 were used in this case as well.

The Golaz *et al.* tuning for MLR runs resulted in a climate with a cold bias (with respect to the surface temperature QOI), in contrast to the warm bias seen in the ULR baseline configuration. Because simulations at MLR are approximately five times more computationally expensive than those for ULR, project time constraints precluded further investigation of additional tunings. We therefore proceeded with this MLR run as a baseline, and branched from simulation year 200 for the perturbed sensitivity runs at this resolution. Time series plots of surface temperature, top of atmosphere flux, and sea ice extent for this simulation compared with the standard 1-degree E3SM results are shown in Figure 3-12.

## 4. DATA MODEL AND FEATURE ANALYSIS

In this analysis we develop machine learning models of September sea ice extent trained on both E3SM historical simulation data and observations. The objective of this work is to identify the key model input features that impact model predictability.

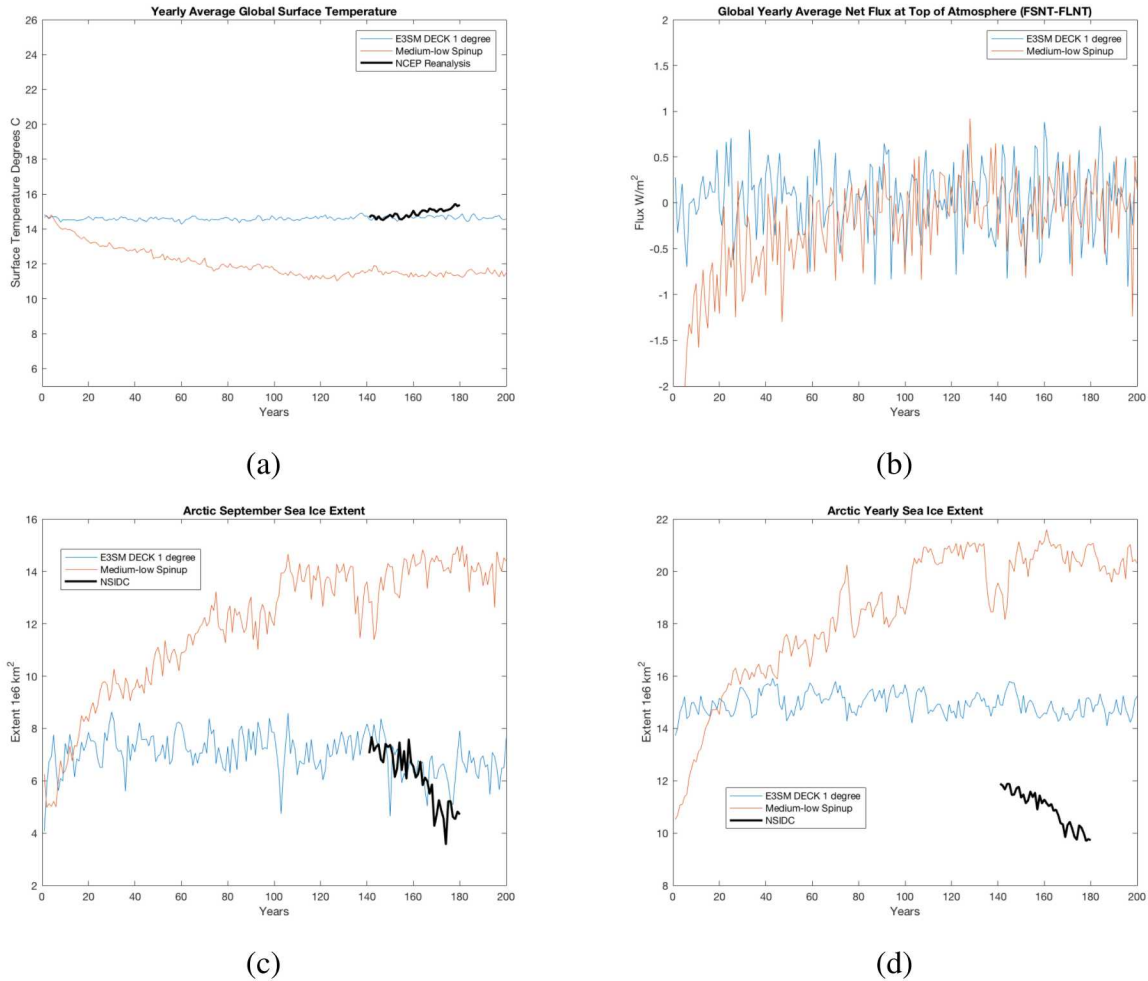
### 4.1. Overview of past work

Seasonal predictions of Arctic sea ice concentration and extent are important for many stakeholders. Accurate seasonal predictions of these quantities depend on better understanding of the factors influencing sea ice variability including interactions and feedbacks between the ocean, atmosphere, and sea ice as well as warming due to greenhouse gas forcing. Since 2008, the Sea Ice Outlook<sup>3</sup> funded by the Sea Ice Prediction Network project<sup>4</sup> has worked to gather predictions of Arctic September sea ice extent and concentration to provide a forum for demonstrating different modeling methods and evaluating the accuracy and skill of these predictions, which come from physics-based models, statistical approaches, as well as machine learning models. The Sea Ice Outlook effort has demonstrated some of the difficulties in generating skillful predictions of seasonal sea ice extent, leading to a number of studies investigating the limitations of sea ice predictability. In their review Guemas *et al.* identify mechanisms influencing predictability that include persistence, or the short term memory of anomalies, advection of sea ice quantities due to ocean currents and atmospheric winds, and reemergence, where seasonal or subseasonal anomalies in quantities like sea ice thickness and extent impact the predictability in the following season [26]. In their investigation of sea ice reemergence, Bushuk and Giannakis computed the covariability of Arctic sea ice extent, sea surface temperature, sea level pressure, and sea ice thickness using nonlinear Laplacian spectral analysis finding a set of coupled modes of variability that influence the memory of Arctic

---

<sup>3</sup><https://www.arcus.org/sipn/sea-ice-outlook>

<sup>4</sup><https://www.arcus.org/sipn/overview>



**Figure 3-12** Yearly averaged global surface air temperature ( $^{\circ}\text{C}$ ) (a), yearly averaged net flux at top of atmosphere ( $\text{W}/\text{m}^2$ ) (b), September sea ice extent ( $10^6\text{km}^2$ ) (c), and yearly averaged sea ice extent ( $10^6\text{km}^2$ ) (d). Blue line from the 500-year E3SM pre-industrial control simulation and red line from 180-year branch run using Golaz *et al.* tuning values. For comparison NSIDC satellite observations are shown in black in (c) and (d) and NCEP atmospheric reanalysis data for surface temperature is shown in black in (a).

sea ice [11]. More recently, Ordonez *et al.* have performed a lag-correlation analysis using the CESM Large Ensemble for predicability of pan-Arctic and pan-Antarctic sea ice area. The authors find in the case of September Arctic sea ice area, August sea ice area is the best predictor of September area and sea ice volume or thickness-related variables are the best predictors when estimating from prior months [54].

Research into modes of variability and the subset of atmosphere, ocean, and ice fields that strongly influence predictability are key to more skillful seasonal forecasts. Coupled physics models include many of the complex feedbacks in the Arctic system, but identifying the critical quantities leading to predictability can be difficult due to the many coupled interactions in the model. Machine learning models provide an opportunity to investigate a smaller subset of important features to determine

their role in generating skillful predictions. Ionita *et al.* used multiple linear regression to successfully predict September sea ice extent based on sea level pressure, air temperature, precipitable water content, surface zonal wind, surface meridional wind, ocean heat content, and sea surface temperature [33]. J. Kim *et al.* developed a deep neural network (DNN) model of Arctic sea ice concentration using regional climate model data from the Coordinated Regional Downscaling Experiment (CORDEX). The inputs to their model included near-surface temperature, near-surface specific humidity, surface downwelling longwave radiation, and surface upwelling longwave radiation [38]. Y. J. Kim *et al.* compared persistence, random forests, and convolutional neural networks in a one-month sea ice concentration prediction model and found that sea ice concentration from the previous month was an important factor in the model [40].

## 4.2. Data and Methods

In this project we investigate the most influential features in predicting September average sea ice extent in the Arctic using machine learning models (MLMs) trained separately on observational data and on simulation data. The intent is to determine which factors are dominant and how they differ between five E3SM historical ensemble members and observations. Monthly averaged data from June for sea ice extent (SIE), sea ice volume (SIV), surface air temperature (TS), surface pressure (PS), cloud fraction (CLT), specific humidity (HSS), surface downward long wave flux (FLWS), sea surface temperature (SST), and meridional and zonal winds (UWIND, VWIND) were used for the training. Observational and reanalysis data including NSIDC monthly sea ice extent, NCEP Reanalysis II atmospheric fields, and ERSST v4 sea surface temperatures were used for the analysis and are described in more detail in Section 2.

Five historical ensembles from E3SM at one-degree resolution generated for CMIP6 were used as training data for comparison with observations. The five ensemble members were initialized from different years of a 500-year pre-industrial control run of E3SM and run with greenhouse gas forcing from 1850 to 2014. For this analysis we used a consistent set of available years 1979-2014 for training and testing in both the historical simulations and the observations. Plots of the June data for training and the September sea ice extent for both the observational data and historical ensembles are shown in Figure 4-1. In most cases the observational and reanalysis data sets are close in magnitude to the simulation data. However, for Arctic cloud fraction (CLT) the NCEP reanalysis is significantly lower than the E3SM values. This is a known bias with the NCEP reanalysis data and future work could investigate feature analysis for alternative reanalysis data sets.

There are a number of possible data-driven models to choose from for the analysis. Neural networks have many advantages, but are not well-suited to small data sets like ours. Initial investigation of multiple linear regression methods found that some of the nonlinear relationships between features limited the ability to fit the data well. Support vector machines were also investigated, however, they do not inherently provide a means to measure feature importance. There are some model-agnostic feature analysis methods such as permutation and drop-column importance, which operate after model training by iteratively manipulating the dataset one feature at a time. Because of that, they are highly susceptible to the effects of multicollinearity and our feature set has strong correlations due to the coupled interactions of the Earth system.

For this analysis, we used random forest regression models because they are a relatively simple, intuitive model that can learn nonlinear relationships between features and provide an inherent importance measure, Gini importance [10]. Random forest regression is an ensemble learning technique, which is similar to a combination of bootstrap aggregation (bagging) and decision tree regression. Bagging is a method to combine the knowledge of many naive estimators, or trees provided a subset of the full feature set. The result is the average of many noisy, but unbiased, estimators, reducing overall variance. Random forests improve the bagging method by adding a process to reduce the correlation of trees. Instead of each tree being given a subset of features, each tree node is given a random set of features to consider when it is splitting. The number of random features each node considers, and when to split are tuned hyperparameters. The final forest’s estimate is the average prediction from the random trees. For  $N$  trees,  $T_1, \dots, T_N$ , random forest regression prediction is computed as follows:

$$RF(N) = \frac{1}{N} \sum_{n=1}^N T_n(x)$$

given the training sample,  $x$ .

One of the benefits of random forests is that they inherently provide an interpretation of feature importance via Gini importance. Because Gini importance is generated by the nonlinear process of decision tree learning, the measure is not heavily influenced by multicollinearity. Given that all of our features come from the same complex system, it would be difficult to eliminate features by simple correlation measures. Most of the time Gini importance is presented normalized to compare relative importance. We preserve the importance values as their raw values to compare across datasets. In our case, the importance is computed as the total reduction in mean absolute error (MAE) caused by each feature.

Ensemble size is an important hyperparameter to tune because the number of trees in forest directly impacts the possible feature sets the forest can explore and it has been shown that too many trees can reduce a random forest’s performance while also sacrificing run-time. Our forests comprised 250 decision trees and the trees use mean squared error as a splitting criterion. In analyzing the fit for the testing and training of the RFS models we used the following three error measures. Given true values,  $y$ , with a mean value of,  $\bar{y}$ , and forecasted values,  $\hat{y}$ , we define

$$R^2 = 1 - \frac{\sum(y - \hat{y})^2}{\sum(\bar{y} - \hat{y})^2} \quad ACC = \frac{\sum(\hat{y} - \bar{y})(y - \bar{y})^2}{M\sigma_{\hat{y}}\sigma_y} \quad MAE = \frac{\sum(y - \hat{y})}{n} \quad (1)$$

where  $\sigma_y$  is the standard deviation of the true values and  $\sigma_{\hat{y}}$  is the standard deviation of the forecasted values.

For this analysis, the data-driven modeling and feature analysis was done with Python scripts utilizing the package `scikit-learn`<sup>5</sup>.

---

<sup>5</sup><https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html>

### 4.3. Feature Importance Results

We found that random forests performed best when predicting 13 random years and training on the remaining 23, as opposed to splitting the data differently. The “best” performance was determined by the highest maximum and minimum test  $\overline{R^2}$  scores among datasets for the June input data. This train-test-split resulted in maximum and minimum  $\overline{R^2}$  scores of 0.88 and 0.77, respectively.  $\overline{R^2}$  denotes the average  $R^2$  of the 385 models. We replicated our analysis for each month between June and August, predicting September SIE. Each succeeding month generated less error and the importance of each feature changed. Figure 4-2 shows June’s feature importance values as box plots and the average error of the 385 models at the top. The boxes show the data’s quartiles and the whiskers indicate the rest of the distribution. The outliers marked are determined using a function of the inter-quartile range. The average train and test error values indicate that the models generally learn the data well. The dotted line shows the mean feature importance of a random variable included in each model’s feature set. The random variable indicates a lower bound importance; any feature with an importance near this line has effectively no importance. We found that adding a random variable does decrease individual model performance but the effect is minimized when taking the mean over all models.

We find that SIV, TS, SSH, SIE, FLWS, and SST are important in each dataset. The datasets, except for simulation 3, share the same list of unimportant features as well; these are CLT, PS, uwind, and vwind. The combined error metrics and general consistency of results between each dataset suggests that our models have learned the data well and the feature analysis can identify key patterns. It is meaningful that the same six features are considered important across datasets and input-months. Although the observed and simulated data share overall patterns, there is a clear difference between them. In June, simulated data relies too heavily on almost all of its features and an extreme reliance on sea ice volume. Towards the end of the summer in August, simulated data develops some more similarities with the observed dataset. Sea ice extent in August should be the most important feature in forecasting sea ice extent only a month later. All datasets correctly rely more on August’s mean sea ice extent and the importance value distributions overlap much better. Still, simulated datasets put noticeably more importance on sea ice volume than the observed data. The remaining features’ importance values are diminished in all datasets and their distributions overlap more than they did in June, but the observed dataset still puts the least importance on FLWS, SSH, TS, and SST.

Additional feature importance results and further discussion are included in [51].

## 5. GLOBAL SENSITIVITY ANALYSIS

As mentioned earlier, one of the primary contributions of this LDRD project is the performance of a global sensitivity analysis (GSA) using several resolutions of the E3SM. This GSA study can be viewed as the first step towards characterizing the parametric uncertainty within this model. To the authors’ knowledge, we perform herein the first global sensitivity analysis involving the fully-coupled E3SM Earth system model, and the first sensitivity study involving version 1 of the fully-coupled E3SM. A true GSA study is possible in part by our consideration of coarser grids

than typically used for production runs using the fully-coupled E3SM, which affords us far more simulations than typically possible using this model. The results of this study provide invaluable insight into the relative importance of various parameters from the sea ice, atmosphere and ocean components of the E3SM (including cross-component parameter interactions) on various Arctic-focused quantities of interest (QOIs). Additionally, our GSA results can inform targeted studies at higher resolutions, as well as help guide model spin-ups at various resolutions.

## 5.1. Overview of Past Work

Recent years have seen a number of studies addressed at understanding the sensitivity of various climate models to relevant parameters. The vast majority of this work has focused on individual components of a global ESM, e.g., the ocean, sea ice and atmosphere components. Numerous studies have investigated the sensitivity to parameters in ocean models, most of them examining subgrid mixing parameterizations, wind drag, model domain and grid resolution, numerical formulations and topography [5, 9, 27, 28, 32, 46, 6, 61]. Several studies have examined the sensitivity to model parameters in stand-alone configurations of sea-ice models, including [39, 57, 75, 77]. In the most recent of these works [77], Urrego-Blanco *et al.* conducted a comprehensive sensitivity analysis of sea ice thickness and area to 39 sea ice model parameters using Sobol sequences together with a fast emulator for the Los Alamos sea ice model, CICE. Similar sensitivity studies have been done also for stand-alone atmosphere models. In [81], Zhao *et al.* evaluated the sensitivity of radiative fluxes at the top of the atmosphere to various cloud microphysics and aerosol parameters. Covey *et al.* [14] used the elementary effects method applied to an ensemble of model simulations to perform a sensitivity analysis with respect to 27 atmospheric parameters. In [58], Qian *et al.* conducted a sensitivity study using the one-at-a-time method and short (three day) simulations using a  $1^\circ$  resolution of the E3SM atmosphere model, EAM, studying the sensitivities of the model with respect to 18 parameters from various parts of the atmospheric dycore, including parameterizations of deep convection, shallow convection and cloud macro/microphysics. Another recent work that focused on the EAM is [60], where attention was focused on how various EAM version 1 parameters should be tuned based on the grid resolution.

While the aforementioned studies provide much insight into individual ESM components, without considering a fully coupled ESM, it is impossible to identify the interaction among various climate components. Hence, studies focusing on a single climate component have the danger of significantly overlooking relevant climate feedbacks. Performing sensitivity studies on fully-coupled climate models is far more challenging than considering an individual climate component for several reasons. First, until recently, fully-coupled ESMs were not readily available to the general public. An additional challenge stems from the fact that running a fully-coupled ESM is far more computationally expensive than running a single climate component. Since sensitivity studies typically require many samples or model realizations, sensitivity analyses using these models are in general intractable without the use of efficient surrogates, especially at “production” grid resolutions (e.g., the  $1^\circ$  or ne30 resolution). The authors are aware of only one reference focusing on a sensitivity study involving several climate components using a fully coupled ESM, namely [76]. In [76], Urrego-Blanco *et al.* use the  $1^\circ$  resolution of the E3SM v0-HiLAT (EHV0) fully coupled climate system to identify emerging relationships between sea ice area, net surface longwave

radiation and atmospheric circulation over the Beaufort gyre. The authors consider five model parameters, two from the atmosphere model (CAM5), two from the sea ice model (CICE5) and two from the ocean model (POP), and initialize their model using preindustrial forcing. By employing an elementary effects or Morris-One-At-a-Time (MOAT) method [50] for their sensitivity analysis (an approach that perturbs one input parameter at a time, rather than all parameters together), the authors are able to keep the number of ensemble members (or E3SM simulations) required down to 24.

It is worthwhile to note that there are other works utilizing global climate models for sensitivity analyses targeting a single climate component. For instance, the reference [59] focuses on performing a sensitivity study of the sea ice simulation within the global coupled climate model HadGEM3, which contains atmosphere, ice and ocean components. Here, both the Arctic and Antarctic are considered. In a similar vein, Uotila *et al.* [75] explore the sensitivity of the global sea-ice distribution of the Australian Climate Ocean Model (AusCOM) to a range of sea-ice physics-related parameters within a global ocean-ice model comprised of AusCOM coupled with the Los Alamos CICE model.

Our present work is primarily motivated by the recent study in [76], but differs in several important ways:

- We consider version 1 of the E3SM, also referred to as E3SM v1 described in Section 3. To the best of our knowledge, this is the first global sensitivity analysis using the fully-coupled E3SM, and the first sensitivity study using E3SM v1.
- We employ much lower spatial resolution grids than those considered in [76], which we refer to as the “Ultra-Low Resolution” (ULR) and “Medium-Low Resolution” (MLR). These grids correspond to a  $7.5^\circ$  (ne4) and  $2.7^\circ$  (ne11) grid resolution in the atmosphere, respectively, and were described in Sections 3.2 and 3.3, respectively. By using lower resolution grids, we are able to afford far more simulations, allowing us to consider more parameters and employ more sophisticated sensitivity analysis approaches than the MOAT method.
- We study the effect of ten parameters, spanning three E3SM components, the sea ice model (MPAS-Sea Ice), the atmosphere model (CAM) and the ocean model (MPAS-Ocean). These parameters are given and discussed in Section 5.2.1.
- We perform a variance-based global sensitivity analysis which uses polynomial surrogates constructed by `PyApprox` [34]. The algorithm is described in more detail in Section 5.2.2.

In the following subsections, we first describe the design of our global sensitivity study and overview how variance-based GSA works. We then present some results obtained from this study using both the ULR and MLR grids considered.

## 5.2. Methodology

Here, we briefly describe the design and mechanics of our global sensitivity analysis study. The goal of this analysis is to quantify the impacts of fluctuations in various parameters in a mathematical model of a physical system (in this case the E3SM) on various model outputs, or QOIs. In

this way, sensitivity analysis can be used to apportion the changes in outputs of a system to different sources of uncertainty in its inputs, and can thus be viewed as a first step towards parametric uncertainty quantification (UQ).

### 5.2.1. Design of Global Sensitivity Study

Before beginning a global sensitivity study, it is necessary to perform a model spin-up at the selected grid resolution, that is, to run the model until an equilibrium state is achieved. This step was described earlier in Sections 3.2 and 3.3.

The first step in designing a sensitivity study involves selecting the parameters to be perturbed, and a set of relevant QOIs, on which the parameters are expected to have an effect. Here we consider ten parameters spanning three E3SM v1 components: the sea ice component (MPAS-Sea Ice), the atmosphere component (CAM), and the ocean component (MPAS-Ocean). A description of the parameters along with the ranges of their values is shown in Table 5-1. The parameters were chosen based on their significance in previous sensitivity studies involving both individual component as well as fully-coupled climate simulations, most notably [77, 76, 61, 6, 58, 60]. Of the ten parameters, three are from the sea ice component, three are from the ocean component and four are from the atmosphere model.

In our GSA, each of the ten parameters was randomly selected from a uniform distribution between the “Min” and “Max” values given in Table 5-1 using the DAKOTA library [4]. Much like the parameters themselves, the selection of the parameter ranges was guided by past analyses [77, 76, 61, 6, 58, 60]. It is worthwhile to note that the three MPAS-Sea Ice parameters selected in our GSA were hard-coded to their default values in the `master` branch of the E3SM<sup>6</sup>. In order to enable the straight-forward specification of these parameters in the relevant input file, a fork of the E3SM was created<sup>7</sup> and used in the present study. Instructions for cloning this fork as well as building the code and submitting a perturbed run on Skybridge are provided in Appendix B.

Component	Parameter	Min	Max	Description
MPAS-Sea Ice	ksno	$2.0 \times 10^{-1}$	$6.0 \times 10^{-1}$	Snow conductivity
	lambda_pond	$1.15 \times 10^{-8}$	$1.15 \times 10^{-4}$	Drainage timescale of ponds
	dragio	$2.0 \times 10^{-4}$	$1.6 \times 10^{-1}$	Ocean-ice drag
CAM	clubb_c1	1.0	5.0	Constant assoc. w/ dissipation of variance $w^2$ (Clubb)
	clubb_c8	2.0	8.0	Const assoc. with Newtonian damping of $w^3$ (Clubb)
	gamma_coeff	$1.0 \times 10^{-1}$	$5.0 \times 10^{-1}$	Const width of PDF in $w$ coord (Clubb)
MPAS-Ocean	cldrf_dp1	$2.0 \times 10^{-2}$	$1.0 \times 10^{-1}$	Deep convection cloud fraction parameter (Clubb)
	standardgm_tracer_kappa	$6.0 \times 10^2$	$1.8 \times 10^3$	Bolus coefficient of GM parameterization of eddy transport
	cvmix_kpp_criticalbulkrichardsonnumber	$2.0 \times 10^{-1}$	1.0	Bulk Richardson number used in KPP vertical mixing scheme
	salinity_restoring_constant_piston_velocity	$7 \times 10^{-7}$	$3 \times 10^{-6}$	Rate at which salinity is restored to a month climatology

**Table 5-1 Global sensitivity analysis parameters.**

In the present study, we compute sensitivity metrics for a set of eleven QOIs, summarized in Table 5-2. Similar QOIs were considered in past works [60] and [76]. Much like the parameters, our QOIs span the three climate components targeted by the study, namely the sea ice, ocean and atmosphere components; see Table 5-2. Also note that the QOIs selected are focused around the

<sup>6</sup>Available at: <https://github.com/E3SM-Project/E3SM>.

<sup>7</sup>Available at: [https://github.com/karapeterson/E3SM\(add\\_namelist\\_params\\_branch\)](https://github.com/karapeterson/E3SM(add_namelist_params_branch)).

Arctic, the geographical area targeted in this study. Following the approach in [77, 76], we look at the QOIs in Table 5-2 annually as well as seasonally.

The sensitivity of the chosen QOIs depend on the following three variables:

- $n_{\text{GSA-yrs}}$  the number of years each perturbed run will be required to complete.
- $n_{\text{GSA-spinup}}$ : the number of years each perturbed run is allowed to spin-up after it has started.
- $n_{\text{GSA-avg}}$ : the year up to which the QOIs will be computed.

Each perturbed run was given a spin-up period, to equilibrate the simulation, that is, to get past the inevitable transient period that occurs when the run commences. Hence, the QOIs are calculated using data from year  $n_{\text{GSA-spinup}} + 1$  to year  $n_{\text{GSA-avg}}$ , where, as expected,  $n_{\text{GSA-spinup}} < n_{\text{GSA-avg}} \leq n_{\text{GSA-yrs}}$ . The values of the variables  $n_{\text{GSA-yrs}}$ ,  $n_{\text{GSA-spinup}}$  and  $n_{\text{GSA-avg}}$  used in the present study are given in Sections 5.3 and 5.4.

QOI	Units	Description	Component
SIE	$km^2$	Total Arctic sea ice extent	sea ice
SIV	$km^3$	Total Arctic sea ice volume	sea ice
SST	$^{\circ}C$	Sea surface temperature averaged over 60-90° N	ocean
TS	$^{\circ}C$	Surface air temperature averaged over 60-90° N	atmosphere
QS	$kg/kg$	Surface air specific humidity averaged over 60-90° N	atmosphere
FLNS	$W/m^2$	Net longwave flux at surface over 60-90° N	atmosphere
CLDLOW	—	Low cloud coverage below 700 hPa averaged over 60-90° N	atmosphere
PRECSL	$m/s$	Large-scale snow precipitation averaged over 60-90° N	atmosphere
BH	$hPa$	Mean sea level pressure over the Beaufort Sea	atmosphere
AL	$hPa$	Mean sea level pressure over the Aleutian Low	atmosphere
SH	$hP$	Mean sea level pressure over the Siberian High	atmosphere

**Table 5-2 Global sensitivity analysis quantities of interest (QOIs).**

A final ingredient an E3SM sensitivity study is the selection of the grid resolution and environmental forcing. For the present study, we use the Ultra-Low and Medium-Low Resolution grids, discussed results. We also used a preindustrial forcing, sometimes referred to as piControl (Section 3). Considering alternate forcing, such as one of the forcings in [22] would be an interesting and useful follow-on exercise to the present study.

### 5.2.2. Variance-Based Global Sensitivity Analysis

Variance-based global sensitivity analysis (VGSA) is to measure the effect that combinations of parameters have on the variance of model output. In this section we describe how to use polynomial chaos expansions to compute Sobol indices as a measure of parameter sensitivities.

#### *Sobol Indices*

Sobol sensitivity indices can be used to quantify the relative importance of parameter combinations on a given QOI. With this goal let  $f$  denote a model output QOI that depends on some model

parameters  $z$ . Any function  $f$  with finite variance parameterized by a set of independent variables  $z$  with probability distribution  $\rho(z) = \prod_{j=1}^d \rho(z_j)$  and support  $\Gamma = \otimes_{j=1}^d \Gamma_j$  can be decomposed into a finite sum, referred to as the ANOVA decomposition,

$$f(z) = \hat{f}_0 + \sum_{i=1}^d \hat{f}_i(z_i) + \sum_{i,j=1}^d \hat{f}_{i,j}(z_i, z_j) + \cdots + \hat{f}_{1,\dots,d}(z_1, \dots, z_d) \quad (2)$$

or more compactly

$$f(z) = \sum_{\mathbf{u} \subseteq \mathcal{D}} \hat{f}_{\mathbf{u}}(z_{\mathbf{u}}) \quad (3)$$

where  $\hat{f}_{\mathbf{u}}$  quantifies the dependence of the function  $f$  on the variable dimensions  $i \in \mathbf{u}$  and  $\mathbf{u} = (u_1, \dots, u_s) \subseteq \mathcal{D} = \{1, \dots, d\}$ .

The functions  $\hat{f}_{\mathbf{u}}$  can be obtained by integration, specifically

$$\hat{f}_{\mathbf{u}}(z_{\mathbf{u}}) = \int_{\Gamma_{\mathcal{D} \setminus \mathbf{u}}} f(z) d\rho_{\mathcal{D} \setminus \mathbf{u}}(z) - \sum_{\mathbf{v} \subseteq \mathbf{u}} \hat{f}_{\mathbf{v}}(z_{\mathbf{v}}), \quad (4)$$

where  $d\rho_{\mathcal{D} \setminus \mathbf{u}}(z) = \prod_{j \notin \mathbf{u}} d\rho_j(z)$  and  $\Gamma_{\mathcal{D} \setminus \mathbf{u}} = \otimes_{j \notin \mathbf{u}} \Gamma_j$ .

The first-order terms  $\hat{f}_{\mathbf{u}}(z_i)$ ,  $\|\mathbf{u}\|_0 = 1$  represent the effect of a single variable acting independently of all others. Similarly, the second-order terms  $\|\mathbf{u}\|_0 = 2$  represent the contributions of two variables acting together, and so on.

The terms of the ANOVA expansion are orthogonal, i.e. the weighted  $L^2$  inner product  $(\hat{f}_{\mathbf{u}}, \hat{f}_{\mathbf{v}})_{L^2_{\rho}} = 0$ , for  $\mathbf{u} \neq \mathbf{v}$ . This orthogonality facilitates the following decomposition of the variance of the function  $f$ .

$$\mathbb{V}[f] = \sum_{\mathbf{u} \subseteq \mathcal{D}} \mathbb{V}[\hat{f}_{\mathbf{u}}], \quad \mathbb{V}[\hat{f}_{\mathbf{u}}] = \int_{\Gamma_{\mathbf{u}}} f_{\mathbf{u}}^2 d\rho_{\mathbf{u}} \quad (5)$$

where  $d\rho_{\mathbf{u}}(z) = \prod_{j \in \mathbf{u}} d\rho_j(z)$ .

The quantities  $\mathbb{V}[\hat{f}_{\mathbf{u}}] / \mathbb{V}[f]$  are referred to as Sobol indices [70] and are frequently used to estimate the sensitivity of  $f$  to single, or combinations of input parameters. Note that this is a *global* sensitivity, reflecting a variance attribution over the range of the input parameters, as opposed to the local sensitivity reflected by a derivative. Two popular measures of sensitivity are the main effect and total effect indices given respectively by

$$S_i = \frac{\mathbb{V}[\hat{f}_{\mathbf{e}_i}]}{\mathbb{V}[f]}, \quad S_i^T = \frac{\sum_{\mathbf{u} \in \mathcal{J}} \mathbb{V}[\hat{f}_{\mathbf{u}}]}{\mathbb{V}[f]} \quad (6)$$

where  $\mathbf{e}_i$  is the unit vector, with only one non-zero entry located at the  $i^{\text{th}}$  element, and  $\mathcal{J} = \{\mathbf{u} : i \in \mathbf{u}\}$ .

Sobol indices (5) can be computed using a number of different ways. Herein, we employ the software library `PyApprox` [34], described in more detail below, which utilizes polynomial chaos expansions as in [73].

### *Polynomial Chaos Emulator*

In this study we use polynomial chaos expansions to compute the Sobol indices of each 11 QOI considered. Polynomial chaos expansions (PCE) [21, 36, 80] represent the model output  $f(z)$  as an expansion of orthonormal polynomials

$$f(z) \approx f_N(z) = \sum_{\lambda \in \Lambda} \gamma_\lambda \phi_\lambda(z), \quad |\Lambda| = N, \quad (7)$$

The basis functions  $\phi_\lambda$  are typically constructed to be orthonormal with respect to the density  $\rho$ . In the following we set  $\phi_\lambda$  to be the tensor product of univariate Legendre polynomials which are orthogonal to the uniformly distributed input variables of our model, i.e.

$$\phi_\lambda(z) = \prod_{i=1}^d \lambda_i \phi_i(z_i)$$

Here  $\lambda = (\lambda_1, \dots, \lambda_d)$  are multivariate indices denoting the degree of the univariate polynomials in each variable dimension. The orthogonality of the PCE basis can be exploited to compute Sobol indices analytically from the expansion via

$$\mathbb{V}[f_{N,\mathbf{u}}] = \sum_{\lambda \in \Lambda_{\mathbf{u}}} \gamma_\lambda^2 \quad \Lambda_{\mathbf{u}} = \{\lambda \mid \lambda_i > 0, i \in \mathbf{u}, \lambda_j = 0, j \notin \mathbf{u}\} \quad (8)$$

Given a set of samples of the random variables  $\mathcal{Z} = \{z^{(m)}\}_{m=1}^M$  and a set of corresponding model outputs  $y = [y^{(1)}, \dots, y^{(M)}]^T$  we compute the coefficients of the PCE using regression with regularization that promotes sparsity. Specifically we begin by setting using a quadratic PCE and solve the LASSO problem

$$\min_{\gamma} \frac{1}{2N} \|y - \Phi\gamma\|_2^2 + \alpha \|\gamma\|_1$$

where  $\Phi$  is a vandermonde-type matrix with entries  $\Phi_{mn} = \phi_n(z^{(m)})$  where there is a one-to-one correspondence between the integers  $n$  and the multi-indices  $\lambda$ . We use 10 fold cross validation to determine the best regularization parameter  $\alpha$ . Once the coefficients of the quadratic expansion have been computed we use the magnitude of these coefficients in an adaptive algorithm [35] which refines the index set  $\Lambda$  to determine the best basis that represents the data.

### *The `PyApprox` Toolkit*

The aforementioned PCE based Sobol indices were computed using a recently-released open-source<sup>8</sup> python library known as `PyApprox` [34]. `PyApprox` provides flexible and efficient tools

<sup>8</sup>Available at: <https://github.com/sandia-labs/pyapprox>.

for high-dimensional approximation and uncertainty quantification. In the present work, we utilize `PyApprox` as a post-processing tool. Following the completion of a given set of perturbed runs using a selected grid resolution of the E3SM (ULR or MLR), we post-process the model output to extract the QOIs in Table 5-2. We feed these data into `PyApprox`, which uses them to perform surrogate-based global sensitivity analysis, as described in Section 5.2.1 and returns the relevant Sobol indices (5). Our entire GSA workflow, including the `PyApprox` step, is summarized below, in Section 5.2.3.

### 5.2.3. GSA Workflow

We end this subsection by summarizing our GSA workflow, including the relevant software that were used.

- *Step 0:* Spin-up the relevant E3SM resolution so that it is in equilibrium with present-day data, following the procedure discussed in Section 3.2.
- *Step 1:* Select a relevant set of parameters  $\{z_i\}$  and appropriate parameter ranges or probability distributions for the parameters (see Table 5-1).
- *Step 2:* Determine a maximum number of runs<sup>9</sup> you can afford to perform using the E3SM, call it  $N$ . Employ DAKOTA to generate  $N$  random samples of the parameters  $\{z_i\}$  from the selected parameter ranges or probability distributions.
- *Step 3:* Create namelist files for the E3SM v1 runs corresponding to each of the  $N$  randomly selected parameter sets. For our study, the relevant namelist files are `user_nl_cam`, `user_nl_mpas`, `user_nl_mpscice`.
- *Step 4:* Set off  $N$  runs of the E3SM, one per each set of perturbed parameter values, branching off the spun-up model run for the selected grid resolution. One should ensure the runs are long-enough to avoid the initial transient period that is inevitable in branch runs such as these (i.e., the value of  $n_{\text{GSA-spinup}}$  introduced in Section 5.2.1 is sufficiently large).
- *Step 5:* Post-process the perturbed runs to extract from them the relevant QOIs (see Table 5-2).
- *Step 6:* Provide the QOI and parameter data to `PyApprox` and use it to perform a surrogate-based GSA to obtain the relevant Sobol indices.

In the present study, *Steps 1–5* were largely automated through the creation of shell scripts that execute the relevant commands comprising these steps. These scripts are stored in an internal repository<sup>10</sup>. All the runs for the ULR and MLR grids were performed on the Skybridge high-capacity cluster located at Sandia National Laboratories, which contains 1848 nodes, each having 16 2.6 GHz Intel Sandy Bridge processors.

---

<sup>9</sup>Note that in our current workflow, one can always create additional samples if one wishes to perform additional runs after having generated the original set of parameter samples, as the samples are independent.

<sup>10</sup>Available at: <https://gitlab.sandia.gov/ArcticTippingPts/ArcticTippingPts>.

### 5.3. Results from Ultra-Low ( $ne4$ ) Resolution (ULR)

In this section we summarize the results of our sensitivity analysis of the ULR configuration of the E3SM v1, which corresponds to an  $ne4$  or  $7.5^\circ$  grid resolution for the atmosphere. For this study, a total of  $N = 212$  parameter values randomly selected from uniform distributions given by the “Min” and “Max” values found in Table 5-1. We set off 212 perturbed runs of E3SM v1, one for each set of parameter values. The aim was to run each case for  $n_{GSA-yr} = 100$  years with a pre-industrial control forcing. We selected  $n_{GSA-spinup} = 50$  years and  $n_{GSA-avg} = 100$  years, that is, in calculating the QOIs (Table 5-2), we performed averaging over years 51-100, so as to give each run a spin-up period of 50 years. In addition to over-writing the default values of the parameters in Table 5-1, two atmospheric parameters were modified in all the perturbed runs: `clubb_c14` and `so4_sz_thresh_icenuc` (see Section 3.2). These parameters were given the values from [22], found in the third column of Table 3-3, so as to be consistent with the values used in the initial model spin-up, discussed in Section 3.2. The values of all 212 perturbed sets of parameters are given in Appendix A. Also included in this Appendix are parameters for the so-called “baseline” run, in which the parameters took on the same values as used in the spin-up run. For this run, we over-write the default values of the atmospheric parameters `zmconv_ke`, `so4_sz_thresh_icenuc`, and `clubb_c14` to the values given in Table 3-3. Additionally, we over-write the default value of the atmospheric parameter `cldfrc_dp1 = 0.045`. All of our  $ne4$  simulations were run on 96 processors (6 nodes) of the Skybridge high-capacity cluster described earlier in Section 5.2.3.

Of the  $N = 212$  perturbed runs, a total of 104 runs made it to  $n_{GSA-yr} = 100$  years. The baseline run also made it to  $n_{GSA-yr} = 100$  years, totalling 105 successful runs. The majority of the runs that did not complete terminated with the following error: `ERROR: negative layer thickness. timestep or remap time too large`, which is indicative of a CFL violation encountered during the run. Attempts to adjust the time-step to circumvent this error were not made for the  $ne4$  resolution runs.

Figures 5-1 and 5-2 shows “spaghetti plots” all eleven QOIs considered (Table 5-2) for each of the runs that made it to 100 years. The QOIs are averaged over each year and plotted as a function of the year since the start of each perturbed run. The baseline run is distinguished from the others by the red markers. The following observations from Figures 5-1 and 5-2 are noteworthy:

- The relationships between the QOIs are as expected, e.g., runs giving rise to a large sea ice volume also give rise to a smaller surface temperature.
- Most of the perturbed runs appear to have reached equilibrium by year 20.
- All eleven QOIs are effectively close to equilibrium for the baseline run, shown using the red markers.
- Significant changes to the QOIs are seen in the perturbed runs, including several cases in which a total loss of Arctic sea ice is observed (`SIE`, `SIV`  $\rightarrow 0$  with time).
- The “spaghetti plots” showing the sea level pressures over the Beaufort High, the Aleutian Low and the Siberian High (Figure 5-2(c)–(e)) appear to all resemble white noise.

We now look at some statistics for the perturbed runs that made it to year 100. Figures 5-3 and 5-4 show box plots for each of the eleven QOIs considered, calculated by season. Here, the seasons are defined as follows: “Winter” is comprised of the months of January to March, “Spring” is comprised of the months April to June, “Summer” is comprised of the months July to September, and “Autumn” is comprised of the months October to December. The red central mark indicates the median of the data, whereas the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers<sup>11</sup> are plotted using the ‘+’ symbol. The following observations from Figures 5-3 and 5-4 are noteworthy:

- The number of outliers depends on the season and also the QOI. The sea ice volume QOI has the most outliers spanning all four seasons. In general, there are more outliers in the summer and autumn seasons, except for the sea level pressure over the Aleutian Low QOI. Here, the season with the most outliers is winter.
- The maximum and minimum sea ice extent is observed in the “Spring” and “Autumn” seasons, respectively. This result may seem surprising, as observational data (see Chapter 4) has shown that the maximum and minimum sea ice extent in general occur in March and September, respectively, which would correspond to the “Winter” and “Autumn” seasons based on our definition. A closer inspection reveals that, for the majority of the ULR<sub>ne4</sub> runs, including the baseline run for this resolution, the maximum and minimum sea ice extent occurs in April and October, respectively (Figure 5-5(a)). Similarly, the maximum and minimum sea ice volume occurs in May and October, respectively (Figure 5-5(b)).

Next, we look at correlations between the various QOIs (Table 5-2) based on our perturbed<sub>ne4</sub> dataset. A similar study was done in [76] by Urrego-Blanco *et al.* to identify so-called “emerging relationships” in the ensemble runs. Tables 5-3 – 5-7 report correlation coefficients between each pair of QOIs considered. The correlation coefficient is a statistical measure of the strength of the relationship between the relative movements of two variables. The values of this coefficient range between  $-1.0$  and  $1.0$ . A correlation of  $-1.0$  is indicative of a perfect negative correlation (i.e., as one variable goes up, the second variable goes down), whereas a correlation of  $1.0$  indicates a perfect positive correlation (i.e., as one variable goes up, so does the second variable).

We first examine Table 5-3, which gives the correlation coefficients between our eleven QOIs averaged annually over the last 50 years, for runs that made it to year 100. The reader can observe that there are strong positive correlations between the following sets of QOIs: (SIE, SIV), (SIE, CLDLOW), (SIE, PRECSL), (SST, TS), (SST, QS), (SST, FLNS), (TS, QS), (QS, FLNS) and (CLDLOW, PRECSL). In contrast, there are strong negative correlations between: (SIE, SST), (SIE, TS), (SIE, QS), (SIE, FLNS), (SIV, TS), (SST, CLDLOW), (SST, PRECSL), (TS, CLD), (TS, PRECSL), (QS, PRECSL), (CLD, FLNS) and (PRECSL, FLNS). In general, these results are as expected: Arctic sea ice extent goes up when snow precipitation increases, and goes down as sea surface temperature and surface temperature go up, low cloud coverage has a strong positive correlation with snow precipitation; for the most part, QOIs that are strongly correlated with each other have similar strong correlations to other QOIs. It is interesting to remark that Arctic sea ice

---

<sup>11</sup>Outliers are defined as values that are more than 1.5 times the interquartile range away from the top or bottom of the box in a given box plot.

extent has a strong positive correlation with low cloud coverage, which indicates that clouds have an insulating effect on Arctic sea ice. Additionally, we observe a negative relationship between FLNS and SIE across all four seasons, indicating that, in general, as longwave radiation at the ocean-ice surface increases, the SIE decreases. Also interesting is the fact that the sea level pressure QOIs (BH, AL and SH) are not highly correlated with any of the other QOIs or each other.

Tables 5-4 – 5-7 give the correlation coefficients between our eleven QOIs, but now averaged seasonally, rather than annually. The reader can observe by examining these tables that the correlations differ by season. The various QOIs are most correlated in the spring season (Table 5-5). One can observe there are non-trivial correlations between BH and the other QOIs in the winter, spring and autumn seasons, and SH and the other QOIs in the spring season. One can note that the correlations between the sea level pressure QOIs and the other QOIs changes drastically based on the season. For example, there is a strong positive relationship between BH and SIE in the winter and autumn (freezing) seasons, a strong negative relationship between these QOIs in the spring season, and almost no relationship between these two QOIs in the summer season. Additionally, Table 5-7 reveals that the SIV QOI is not strongly correlated with any of the other QOIs in the autumn season, presumably because this is the season with minimal sea ice volume (see Figures 5-3 and 5-5. For a brief discussion about how these results compare to the results of Urrego-Blanco *et al.* in [76], the reader is referred to Section 5.5.

**Table 5-3 Table of correlation coefficients between the eleven QOIs considered (Table 5-2), averaged annually over the last 50 years, for ne4 runs that made it to year 100. Large positive correlation coefficients ( $\geq 0.85$ ) are colored blue. Large negative correlation coefficients ( $\leq -0.85$ ) are colored yellow.**

	SIE	SIV	SST	TS	QS	CLDLow	PRECSL	FLNS	BH	AL	SH
SIE	1.0	0.86	-0.94	-0.99	-0.94	0.91	0.96	-0.90	0.61	0.42	-0.57
SIV		1.0	-0.69	-0.85	-0.73	0.61	0.72	-0.60	0.41	0.43	-0.36
SST			1.0	0.96	1.0	-0.99	-0.99	0.96	-0.55	-0.28	0.74
TS				1.0	0.98	-0.93	-0.98	0.90	-0.55	-0.36	0.65
QS					1.0	-0.97	-0.99	0.94	-0.51	-0.27	0.75
CLDLow						1.0	0.98	-0.98	0.56	0.31	-0.69
PRECSL							1.0	-0.96	0.59	0.35	-0.68
FLNS								1.0	-0.61	-0.35	0.62
BH									1.0	0.68	-0.02
AL										1.0	0.39
SH											1.0

Finally, we present and discuss the results of the GSA study using the method described in Section 5.2.2. For the  $M = 105$  runs that made it to year 100 (including the baseline run), we calculated the eleven QOIs in Table 5-2, averaging annually as well as seasonally over the last 50 years of the runs. These data were fed into PyApprox, where they were used to build a polynomial surrogate, which was employed in the calculation of the Sobol indices. The main results are summarized in Figures 5-6–5-16 below. For each figure and subfigure, three plots are reported: a pie chart showing the “main effects”, a bar plot showing the “total effects” and a pie chart showing the individual Sobol indices corresponding to each of the ten parameters considered (Table 5-1). Also reported in each

**Table 5-4 Table of correlation coefficients between the eleven QOIs considered (Table 5-2), averaged during the winter season (January–March) over the last 50 years, for  $n_{e4}$  runs that made it to year 100. Large positive correlation coefficients ( $\geq 0.85$ ) are colored blue. Large negative correlation coefficients ( $\leq -0.85$ ) are colored yellow.**

	SIE	SIV	SST	TS	QS	CLDLow	PRECSL	FLNS	BH	AL	SH
SIE	1.0	0.81	-0.90	-0.98	-0.96	0.51	0.86	-0.26	0.96	0.47	0.23
SIV		1.0	-0.60	-0.89	-0.74	-0.02	0.48	0.17	0.79	0.59	0.48
SST			1.0	0.88	0.98	-0.72	-0.98	0.45	-0.87	-0.30	0.05
TS				1.0	0.96	-0.40	-0.81	0.15	-0.94	-0.50	-0.27
QS					1.0	-0.60	-0.94	0.33	-0.92	-0.39	-0.08
CLDLow						1.0	0.81	-0.78	0.46	-0.002	-0.29
PRECSL							1.0	-0.56	0.83	0.25	-0.11
FLNS								1.0	-0.26	0.12	0.31
BH									1.0	0.48	0.19
AL										1.0	0.82
SH											1.0

**Table 5-5 Table of correlation coefficients between the eleven QOIs considered (Table 5-2), averaged during the spring season (April–June) over the last 50 years, for  $n_{e4}$  runs that made it to year 100. Large positive correlation coefficients ( $\geq 0.85$ ) are colored blue. Large negative correlation coefficients ( $\leq -0.85$ ) are colored yellow.**

	SIE	SIV	SST	TS	QS	CLDLow	PRECSL	FLNS	BH	AL	SH
SIE	1.0	0.84	-0.96	-0.98	-0.97	0.96	0.99	-0.93	-0.92	-0.47	-0.87
SIV		1.0	-0.74	-0.89	-0.83	0.72	0.86	-0.64	-0.87	-0.34	-0.74
SST			1.0	0.96	0.99	-0.99	-0.96	0.97	0.91	0.57	0.92
TS				1.0	0.99	-0.95	-0.99	0.90	0.96	0.51	0.91
QS					1.0	-0.98	-0.99	0.93	0.95	0.56	0.93
CLDLow						1.0	0.96	-0.98	-0.90	-0.52	-0.88
PRECSL							1.0	-0.92	-0.96	-0.47	-0.88
FLNS								1.0	0.84	0.48	0.82
BH									1.0	0.44	0.86
AL										1.0	0.79
SH											1.0

subfigure of each figure is an  $R^2$  value, which is a measure of how well the constructed surrogate is able to represent the given data. A perfect fit corresponds to an  $R^2 = 1$ . Note the values reported, in all Figures, for main effect, and Sobol indices are multiplied by 100 so that the values listed are percentages of the total variance.

The main effect values measure the effect of individual parameters acting alone and can sum to at most 1. As the sum approaches 1, the contribution of all parameter combinations involving two or more variables decreases. A value of 1 indicates that the function is purely additive and there is no interaction between any parameters. For a number of QOI it appears that the strength of the parameter interactions between parameters is seasonally dependent. For example, compare the main effect values in the winter and summer months in Figure 5-6.

**Table 5-6 Table of correlation coefficients between the eleven QOIs considered (Table 5-2), averaged during the summer season (July–September) over the last 50 years, for ne4 runs that made it to year 100. Large positive correlation coefficients ( $\geq 0.85$ ) are colored blue. Large negative correlation coefficients ( $\leq -0.85$ ) are colored yellow.**

	SIE	SIV	SST	TS	QS	CLDLOW	PRECSL	FLNS	BH	AL	SH
SIE	1.0	0.90	-0.88	-0.90	-0.85	0.87	0.98	-0.88	0.07	0.41	-0.72
SIV		1.0	-0.68	-0.71	-0.65	0.66	0.94	-0.66	-0.16	0.24	-0.68
SST			1.0	1.0	0.99	-1.0	-0.81	0.97	-0.21	-0.37	0.74
TS				1.0	0.99	-0.99	-0.84	0.96	-0.17	-0.35	0.76
QS					1.0	-0.99	-0.78	0.94	-0.18	-0.30	0.76
CLDLOW						1.0	0.80	-0.98	0.23	0.40	-0.72
PRECSL							1.0	-0.80	-0.03	0.34	-0.73
FLNS								1.0	-0.26	-0.50	0.66
BH									1.0	0.77	0.23
AL										1.0	0.17
SH											1.0

**Table 5-7 Table of correlation coefficients between the eleven QOIs considered (Table 5-2), averaged during the autumn season (October–December) over the last 50 years, for ne4 runs that made it to year 100. Large positive correlation coefficients ( $\geq 0.85$ ) are colored blue. Large negative correlation coefficients ( $\leq -0.85$ ) are colored yellow.**

	SIE	SIV	SST	TS	QS	CLDLOW	PRECSL	FLNS	BH	AL	SH
SIE	1.0	0.90	-0.75	-0.91	-0.78	0.69	0.86	-0.63	0.89	0.77	-0.34
SIV		1.0	-0.57	-0.79	-0.61	0.47	0.67	-0.43	0.75	0.65	-0.19
SST			1.0	0.95	0.99	-0.96	-0.97	0.91	-0.88	-0.78	0.73
TS				1.0	0.96	-0.89	-0.97	0.82	-0.94	-0.83	0.60
QS					1.0	-0.95	-0.97	0.88	-0.89	-0.78	0.74
CLDLOW						1.0	0.95	-0.97	0.86	0.79	-0.67
PRECSL							1.0	-0.90	0.93	0.85	-0.60
FLNS								1.0	-0.81	-0.75	0.59
BH									1.0	0.90	-0.43
AL										1.0	-0.26
SH											1.0

Total effect indices measure the total contribution of each parameter to the variance of a QOI, specifically they measure the contributions of all interactions involving a specific parameter. Consequently the total effect index of a single variable will always be at least as large as the main effect index of that variable. Furthermore, the sum of all total effect indices can be greater than 1, because Sobol indices for parameter interactions involving at least two variables can be used to compute the total effects of multiple variables. For example, the Sobol index of  $S_{ij} = \mathbb{V}[\hat{f}_{ij}]$  will contribute to the total effect indices of both the  $i$ th and  $j$ th variables. Comparing main effect and total effect indices can be used to determine the strength of high-order (involving more than two parameter) interactions. For example, in Figure 5-7(a), the main effect of ksno ( $z_1$ ) is less than 1%

of the total variance, yet the total effect of this variable is over 10% of the total variance. This is true for other parameters as well.

Main and total effect indices summarize the contributions of a single parameter to the variance of a QOI. Sobol indices can be used to identify the contribution of specific parameter interactions to the total variance. Sobol indices involving just one parameter are labeled  $(z_k)$  and indices involving two parameters are labeled  $(z_i, z_j)$  with  $i \neq j$ . The slices labeled “other” contain the sum of the contributions by miscellaneous pairs of parameters, each of which was deemed too small to visualize in the chart (the percent contribution was  $< 1\%$ ). There are no strong interactions involving three or more variables.

In the following we summarize the conclusions that can be drawn from Figures 5-6–5-16. The most notable conclusions are emphasized in ***bold italics***.

- ***Overall, we find that the atmospheric parameters related to the Clubb cloud model, clubb\_c1, clubb\_c8, gamma\_coeff and cldrf\_dp1 (see Table 3-4) are the most significant for the QOIs considered. This result is consistent with results obtained in earlier sensitivity studies, namely [58] and [76].***
- ***We also find that the clubb\_c1 atmospheric parameter frequently shows up in the combined effects with other parameters.***
- The sea ice and ocean parameters have little influence on the SST, TS, CLDLow, PRECSL, FLNS and SH parameters.
- The sea ice parameters standardgm\_tracer\_kappa and cvmix\_kpp\_criticalbulkrichardsonnumber have the most influence on the sea level pressure QOIs (BH, AL, SH).
- In general, the total effects mimic the Sobol indices, with a few minor exceptions (e.g., the BH QOI in spring).
- The  $R^2$  values are high ( $> 0.8$ ) for most of the QOIs, indicating that we have enough accuracy to accurately draw qualitative conclusions about parameters with large sensitivities. More data is needed to order all variables reliably.
- For certain QOIs, e.g., SIE, SIV, TS, CLDLow, PRECSL, FLNS, BH, AL, and SH, there are large differences in the most influential parameters across seasons. Some of these make sense intuitively, whereas others do not. For example, one can see from Figure 5-7 that the sea ice parameters ksno is important only in the winter and spring for the SIV QOI. This makes sense: this parameter represents the snow conductivity, which is relevant only when the sea ice is frozen. Interestingly, these trends are reversed for the SIE QOI, despite this QOI being highly correlated with SIV. Additional analysis is needed to understand these apparent contradictions.
- ***Our GSA reveals that there are interactions between a number of parameters that span different E3SM components, e.g., atmosphere-ocean, ocean-sea ice, atmosphere-sea ice. Only a full coupled climate model ESM study is capable of revealing such parameter interactions.***

#### 5.4. Results from Medium-Low (ne11) Resolution (MLR)

Our next objective is to repeat the study described Section 5.3 but at the medium-low resolution of the E3SM v1. This resolution corresponds to an ne11 or  $2.7^\circ$  grid resolution. For this study, we selected a subset of the ne4 sets of parameter values (see Appendix A), namely the subset of parameter values that led simulations at the ne4 resolution that made it to at least year 75. As before, a pre-industrial control forcing was used for these runs. Also, as with the ne4 study, a “spin-up” was performed prior to setting off our ensemble of perturbed runs, so as to ensure the initial climate state is more-or-less in equilibrium. The details of this spin-up are described earlier in Section 3.3. We remark that the ULR and MLR spin-ups gave rise to slightly different spun-up states (see Figure 3-11 vs. Figure 3-12), as discussed earlier. Specifically, one can see by inspecting these figures that the surface temperature for the ne11 spin-up is too low relative to the E3SM DECK 1 degree reference solution, whereas the opposite trend is observed for the ne4 spin-up. Similarly, the Arctic sea ice extent is too high for the ne11 spin-up run, whereas it is too low for the ne4 run. Similar trends are seen for our baseline ne4 and ne11 runs, which use the same parameters used in each of the spin-ups (Figures 5-17 and 5-17). Like the ne4 runs, the ne11 runs were performed on the Skybridge high-capacity cluster at Sandia, but this time utilizing 384 processors, or 68 nodes, of the machine. The ne11 simulations were run to year 75 ( $n_{\text{GSA-yr}} = 75$ ), and the relevant QOIs (Table 5-2) were calculated by performing an average over the last 25 years ( $n_{\text{GSA-spinup}} = 50$ ,  $n_{\text{GSA-avg}} = 75$ ), both annually and seasonally, like for the ne4 case.

Of the 140 perturbed ensemble simulations launched at the ne11 resolution, including the baseline simulation, only 16 completed to year 75. The successful runs corresponds to runs with indices {90, 181, 108, 117, 65, 22, 144, 124, 154, 141, 47, 44, 123, 167, 180, 0} where the index 0 corresponds to the baseline run. The parameters for each of these runs can be found in Appendix A.

The majority of the ne11 runs died with one of two errors:

- (a) a segmentation fault in the radiation transport scheme within the E3SM Atmosphere Model (EAM);
- (b) an `MPI_ABORT` error within MPAS-Ocean occurring due to the presence of NaNs among the MPAS-Ocean states.

Numerical experiments informed by conversations with EAM and MPAS-Ocean developers<sup>12</sup> revealed that these errors can be circumvented for at least some of the runs by reducing various physics timesteps within the two models. Specifically, upon encountering error (a), the general rule of thumb is to increase the `se_nsplrit` parameter from the default value of 4. Increasing this parameter by a factor of  $2\times$  has the effect of reducing the EAM time-step by a factor of 2. Error (b) can be circumvented by reducing the following two time-steps in MPAS-Ocean: `config_btr_dt` and/or `config_dt`. For the successful perturbed runs, we used `se_nsplrit` values as high as 32, `config_btr_dt` values as low as 000\_01:00:00 and `config_dt` values as low as 00:30:00. By tuning these three time-step-related parameters, it may be possible to obtain more successful perturbed runs at the ne11 resolution. This will be explored in a subsequent journal publication [74].

---

<sup>12</sup>Personal correspondence with Oksana Guba, Matt Hoffman and John Wolfe.

Figures 5-19 and 5-20 show “spaghetti plots” for the 16 successful runs that made it to year 75. Here, the baseline run is distinguished from the perturbed runs with the red markers. As with the  $ne4$  runs, significant changes to the QOIs are seen in some of the perturbed runs, including several cases in which a total loss of Arctic sea ice is observed (SIE, SIV  $\rightarrow 0$  with time). It is interesting to compare each QOI for a particular run at the  $ne11$  resolution with the corresponding QOI for the same run at the  $ne4$  resolution. In order to do so, one must normalize the QOIs, which can be done by subtracting out the baseline (Figures 5-17 and 5-17). Figures 5-21 and 5-22 show the result for a representative run, corresponding to the 65th perturbed set of parameters (Appendix A). The reader can observe the following from examining these figures:

- Similar trends are seen for the SIE, SIV, TS, BH, AL and SH QOIs at the  $ne11$  resolution as at the  $ne4$  resolution.
- The SST, QS and FLNS QOIs are substantially lower for the  $ne11$  resolution run than for the  $ne4$  resolution run.
- The CLDLOW and PRECSL QOIs are substantially higher for the  $ne11$  resolution run than for the  $ne4$  resolution run.

Attempts to understand these biases will be made in a subsequent publication [74].

Next, we present box plots for the eleven QOIs of interest for the  $ne11$  resolution simulations (Figures 5-23 and 5-24). Comparing these figures with the analogous figures for the  $ne4$  resolution runs (Figures 5-3 and 5-4), the following observations are noteworthy:

- We again find that the minimum and maximum sea ice extent and volume occur in the autumn and spring seasons, respectively.
- Similar trends are seen in the SIE, SIV, SST, TS, QS and PRECSL QOIs.
- In general, there are fewer outliers for the  $ne11$  dataset. This could be attributed to there being fewer ensemble members in this dataset. The quartile ranges are significantly larger for some QOIs at the  $ne11$  resolution relative to the  $ne4$  resolution (e.g., SIV), but smaller for other QOIs (e.g., QS).
- The trends for the FLNS, CLDLOW, BH, AL and SH QOIs are somewhat different at the  $ne11$  resolution. For instance, FLNS is highest in the spring and lowest in the summer for the  $ne11$  resolution runs, whereas it is the highest in the summer and lowest in the autumn for the  $ne4$  resolution runs. This QOI appears to be resolution dependent, as discussed in more detail in Section 5.5.

Additional analyses may be needed to make definitive conclusions from the perturbed ensemble data.

Finally, we look at the correlation coefficients between the eleven QOIs considered for the  $ne11$  runs that made it to year 75 (Tables 5-8 – 5-12), and compare these to their  $ne4$  analogs (Tables 5-3 – 5-7). In general, the correlations between SIE, SIV, SST, TS and QS are more or less the same for the  $ne11$  runs as for the  $ne4$  runs. For the other QOIs, there are some fundamental differences, however. When considering annually-averaged data (Tables 5-8 vs. 5-3), the reader can observe that there are much stronger correlations between the CLDLOW, PRECSL and FLNS QOIs and

other QOIs for the `ne4` dataset than for the `ne11` dataset. In contrast, there are higher correlations between the sea level pressure QOIs BH and SH for the `ne11` dataset than the `ne4` dataset. Moreover, whereas for the `ne4` dataset, the most significant correlations were observed during the spring season, for the `ne11` dataset, the spring season actually saw the fewest number of significant correlations. Lastly, it is interesting to remark that some of the strong positive correlations in the `ne4` dataset because strong negative correlations in the `ne11` dataset, and vice versa. A concrete example of this occurrence is the correlation between the PRECSL and SIE QOIs in the spring, as well as the BH QOI and pretty much any other QOI, also in the spring. It is difficult to say if these differences are due to resolution alone, given that there are several other differences in the `ne4` and `ne11` datasets, including different initial conditions and a different number of ensemble runs. Reconciling these differences and their effect on the QOI correlations will be the subject of future work.

**Table 5-8 Table of correlation coefficients between the eleven QOIs considered (Table 5-2), averaged annually over the last 25 years, for `ne11` runs that made it to year 75. Large positive correlation coefficients ( $\geq 0.85$ ) are colored blue. Large negative correlation coefficients ( $\leq -0.85$ ) are colored yellow.**

	SIE	SIV	SST	TS	QS	CLDLOW	PRECSL	FLNS	BH	AL	SH
SIE	1.0	0.88	-0.97	-0.98	-1.0	-0.21	0.02	-0.66	0.88	0.67	0.79
SIV		1.0	-0.94	-0.80	-0.88	0.19	0.38	-0.65	0.71	0.55	0.47
SST			1.0	0.90	0.97	-0.007	-0.28	0.63	-0.82	-0.59	-0.70
TS				1.0	0.98	0.34	0.17	0.68	-0.92	-0.69	-0.85
QS					1.0	0.18	-0.03	0.68	-0.90	-0.67	-0.80
CLDLOW						1.0	0.81	-0.11	-0.18	-0.20	-0.42
PRECSL							1.0	-0.006	-0.13	-0.18	-0.27
FLNS								1.0	-0.83	-0.25	-0.62
BH									1.0	0.48	0.87
AL										1.0	0.54
SH											1.0

A GSA study analogous to the one presented in Section 5.3 is beyond the scope of this report, and will be discussed in a separate publication that is currently in preparation, namely [74]. In addition to understanding the effect of the grid resolution on the relative parameter and QOI relationships, we aim to try to uncover to what extent the ULR model is correlated with the MLR model, towards performing a multi-fidelity analysis.

## 5.5. Comparison to Other Work

It is interesting to compare some of our results in Sections 5.3 and 5.4 to results of previous sensitivity studies involving climate models. We note that no comparison will be truly apples-to-apples, as no previous study focused on E3SM v1 at the grid resolutions considered here.

We focus the discussion herein to the closest study existing in present-day literature to the present study, namely the study discussed by Urrego-Blanco *et al.* in [76]. This study is based on the E3SM

**Table 5-9 Table of correlation coefficients between the eleven QOIs considered (Table 5-2), averaged during the winter season (January–March) over the last 25 years, for ne11 runs that made it to year 75. Large positive correlation coefficients ( $\geq 0.85$ ) are colored blue. Large negative correlation coefficients ( $\leq -0.85$ ) are colored yellow.**

	SIE	SIV	SST	TS	QS	CLDLOW	PRECSL	FLNS	BH	AL	SH
SIE	1.0	0.84	-0.90	-0.98	-0.99	-0.73	-0.98	-0.67	0.92	0.71	0.94
SIV		1.0	-0.90	-0.76	-0.86	-0.27	-0.75	-0.42	0.71	0.54	0.69
SST			1.0	0.82	0.92	0.42	0.80	0.37	-0.79	-0.58	-0.80
TS				1.0	0.97	0.82	1.0	0.78	-0.94	-0.70	-0.97
QS					1.0	0.69	0.96	0.65	-0.93	-0.68	-0.95
CLDLOW						1.0	0.83	0.82	-0.73	-0.54	-0.81
PRECSL							1.0	0.77	-0.94	-0.69	-0.96
FLNS								1.0	-0.72	-0.33	-0.75
BH									1.0	0.60	0.97
AL										1.0	0.66
SH											1.0

**Table 5-10 Table of correlation coefficients between the eleven QOIs considered (Table 5-2), averaged during the spring season (April–June) over the last 25 years, for ne11 runs that made it to year 75. Large positive correlation coefficients ( $\geq 0.85$ ) are colored blue. Large negative correlation coefficients ( $\leq -0.85$ ) are colored yellow.**

	SIE	SIV	SST	TS	QS	CLDLOW	PRECSL	FLNS	BH	AL	SH
SIE	1.0	0.83	-0.94	-0.98	-0.99	-0.21	0.47	0.66	0.44	-0.29	-0.76
SIV		1.0	-0.92	-0.76	-0.85	0.22	0.77	0.26	0.22	-0.27	-0.68
SST			1.0	0.87	0.94	-0.04	-0.72	-0.48	-0.36	0.33	0.74
TS				1.0	0.98	0.32	-0.33	-0.67	-0.50	0.28	0.69
QS					1.0	0.16	-0.49	-0.60	-0.48	0.28	0.72
CLDLOW						1.0	0.58	-0.71	-0.18	0.31	0.15
PRECSL							1.0	-0.10	0.07	-0.12	-0.36
FLNS								1.0	0.42	-0.45	-0.65
BH									1.0	-0.40	-0.15
AL										1.0	0.30
SH											1.0

v0, which has different climate components than the model considered herein, E3SM v1. Additionally, a different (finer) ne30 or  $1^\circ$  grid resolution is used in the Urrego-Blanco *et al.* study. Only five parameters were perturbed in this study: an atmospheric stability parameter  $C$  in the POP-CAM5 ocean/atmosphere model, the water vapor threshold for cloud formation parameter  $cldfrc\_rhminl$  in CAM5, the  $micro\_mg\_dcs$  parameter in CAM5 which controls the auto-conversion of ice crystals to snow, the snow thermal conductivity parameter  $ksno$  in CICE5, and the  $rsnw\_mlt$  parameter in CICE5, which controls the maximum snow grain size during sea ice melting. The only parameter that is considered in both the Urrego-Blanco *et al.* study and our study is  $ksno$ . One can loosely relate the  $cldfrc\_rhminl$  parameter in [76] to the four Clubb parameters in our study, as it is also a cloud parameterization parameter. Note that the QOIs considered in [76] are a subset of the QOIs

**Table 5-11 Table of correlation coefficients between the eleven QOIs considered (Table 5-2), averaged during the summer season (July–September) over the last 25 years, for ne11 runs that made it to year 75. Large positive correlation coefficients ( $\geq 0.85$ ) are colored blue. Large negative correlation coefficients ( $\leq -0.85$ ) are colored yellow.**

	SIE	SIV	SST	TS	QS	CLDLOW	PRECSL	FLNS	BH	AL	SH
SIE	1.0	0.91	-0.98	-0.99	-1.00	0.93	1.0	-0.39	-0.07	0.76	-0.60
SIV		1.0	-0.94	-0.88	-0.90	0.90	0.89	-0.57	-0.07	0.71	-0.50
SST			1.0	0.96	0.98	-0.94	-0.97	0.50	0.07	-0.71	0.51
TS				1.0	1.0	-0.93	-1.0	0.34	0.09	-0.79	0.63
QS					1.0	-0.94	-1.0	0.40	0.09	-0.77	0.59
CLDLOW						1.0	0.92	-0.65	0.05	0.69	-0.41
PRECSL							1.0	-0.35	-0.10	0.79	-0.62
FLNS								1.0	-0.34	-0.14	-0.22
BH									1.0	-0.36	0.14
AL										1.0	-0.72
SH											1.0

**Table 5-12 Table of correlation coefficients between the eleven QOIs considered (Table 5-2), averaged during the autumn season (October–December) over the last 25 years, for ne11 runs that made it to year 75. Large positive correlation coefficients ( $\geq 0.85$ ) are colored blue. Large negative correlation coefficients ( $\leq -0.85$ ) are colored yellow.**

	SIE	SIV	SST	TS	QS	CLDLOW	PRECSL	FLNS	BH	AL	SH
SIE	1.0	0.91	-0.98	-0.97	-1.0	-0.92	-0.57	-0.92	0.93	0.50	0.80
SIV		1.0	-0.94	-0.81	-0.90	-0.74	-0.32	-0.78	0.80	0.40	0.52
SST			1.0	0.90	0.97	0.84	0.39	0.84	-0.87	-0.50	-0.70
TS				1.0	0.98	0.98	0.74	0.95	-0.95	-0.51	-0.88
QS					1.0	0.94	0.59	0.92	-0.94	-0.52	-0.81
CLDLOW						1.0	0.80	0.89	-0.90	-0.55	-0.93
PRECSL							1.0	0.70	-0.64	-0.35	-0.80
FLNS								1.0	-0.96	-0.27	-0.79
BH									1.0	0.29	0.78
AL										1.0	0.54
SH											1.0

considered in our study (assuming the Sea Ice Area, or IAREA in [76], is highly correlated with Sea Ice Extent), though they are averaged differently than in our study: in [76], an 80 year spin-up is considered, of which only the last 30 years of the simulations are used to calculate the QOIs. Lastly, recall that the sensitivity analysis methodology in [76] differs from ours: we perform a true GSA, compared to the application of the MOAT method [50] in [76].

First, we compare the box plots from our ULR (Figures 5-3 and 5-4) and MLR (Figures 5-23 and 5-24) simulations to those in [76] (Fig. 1, p. 9576). The reader can observe that we see similar trends in the SIV QOI box plots for both the ULR and MLR runs as those in the Urrego-Blanco

*et al.* study. In contrast, the BH QOI box plot appears significantly different from those in [76] for the ULR simulations: the median value of BH rises from winter to spring and from spring to summer, dropping finally in the autumn in our  $ne_4$  study (Figure 5-3(e)), in contrast to this value dropping or remaining approximately constant across the four seasons from winter to autumn in the Urrego-Blanco *et al.* study. The reader can remark, however, that the  $ne_{11}$  box plots for this QOI (Figure 5-24(e)) show a trend that is much more in line with the trend observed in [76]. It is conceivable that the  $ne_4$  results are fundamentally different due to this resolution not being capable of representing the relevant physical phenomena affecting the BH QOI. A somewhat similar trend is observed from the CLDLLOW QOI box plots. In our  $ne_4$  study, the median value of CLDLLOW decreases from winter to spring and from spring to summer, increasing again in autumn (Figure 5-4(a)). The opposite trend is observed in [76]. While our  $ne_{11}$  box plot for this QOI does not mimic exactly the trends reported in Urrego-Blanco *et al.*, the reader can observe that the relevant box plots are much more similar between these two studies, with similar relative median values for CLDLLOW in all seasons except the autumn season. Again, a reasonable conjecture is that  $ne_4$  study leads to different trends due to the underlying discretization being too coarse to resolve all the relevant physical phenomena. It is important to recognize, however, that there are other differences between the two studies besides resolution: they are based on different version of the E3SM (v1 vs. v0) and different sensitivity analysis methodologies (MOAT vs. GSA).

Turning our attention now to the total effects (Figures 2 and 9 in [76]), it proves to be more difficult to reconcile the results of the present  $ne_4$  study and that of Urrego-Blanco *et al.* A few comparisons can be made nonetheless. Both studies reveal that the SIV QOI is most influenced by the widest range of parameters spanning the three climate components considered. Both studies also suggest that the  $ksno$  parameter affects the FLNS QOI to a significant degree in the winter season. In addition, both studies demonstrate that the cloud microphysics parameterization parameter(s) appear to be the one(s) affecting most of the QOIs across all four seasons. Importantly, this result is consistent with the discussion by Qian *et al.* in [58], where a similar conclusion is reached by performing a sensitivity analysis of the stand-alone atmosphere model in E3SM v1.

Another interesting exercise is to compare the cross-correlations between the various QOIs (Tables 5-4 – 5-7 and Tables 5-9 – 5-12) with the so-called “emerging relationships” uncovered in [76] through the set of ensemble runs performed therein. The analysis of Urrego-Blanco *et al.* leads to the following conclusions:

- there is a positive correlation between FLNS and Arctic sea ice area in the melting season (modulated by clouds), while the opposite correlation is found in the freezing season (dominated by sea ice);
- there is a significant positive correlation between Arctic sea ice area and BH during the autumn season, a significant negative correlation during the summer season, and negligible correlations during the winter and spring seasons (since ice coverage in the central Arctic is almost at its maximum through most of these seasons)

The present results lead to somewhat different conclusions, especially at the coarser  $ne_4$  resolution. While we find moderate negative correlations between FLNS and SIE during the autumn and winter (freezing) seasons, our  $ne_4$  study actually suggests that FLNS and SIE are strongly negatively correlated in the melting (spring, summer) seasons, which is contrary to the result in [76].

It is interesting to remark that our `ne11` resolution results are more consistent with those in [76]: a moderate positive correlation is seen in the spring (melting season). A similar trend is observed for the BH QOI. Unlike the work of Urrego-Blanco *et al.* our `ne4` results show that there is almost no correlation between BH and SIE in the autumn, and that the highest correlation between these QOIs occurs in the spring. As with FLNS, the BH-SIE correlations for the `ne11` runs do share some similarities with those in [76]: one can see that both sets of results reveal a strong positive correlation between these QOIs in the fall. The discrepancies between the `ne4` and Urrego-Blanco *et al.* correlations may be attributed to the large differences in grid resolution. Further data/analysis would be needed to corroborate this hypothesis, however.

## 5.6. Significance

As mentioned earlier, the study described herein is, to the authors' knowledge, the first global sensitivity analysis involving the fully-coupled E3SM Earth system model, and the first sensitivity study involving version 1 of the fully-coupled E3SM. In addition to providing invaluable insight into the relative importance of various E3SM v1 parameters on a number of Arctic-focussed QOIs, the results of the present study can be used also to inform subsequent simulations and spin-ups on the `ne4` and `ne11` resolution grids. Specifically, the results of this study can be used to determine which parameter to tune and to what extent to achieve a desired equilibrium configuration (i.e., to increase/decrease the SST, SIV, etc. to get them to better agree with available observational data). Further studies, including a GSA study using the `ne11` set of ensemble simulations, will be presented in a follow-up publication, [74].

## 6. ENRICHMENT ANALYSIS FOR SIMULATION ERROR STATUS

As discussed in Section 5.3, slightly more than half of our 212 perturbed runs did not make it to 100 years and these runs failed with several different types of errors. We performed an analysis to see if the parameters of the experiment were correlated with whether the simulation completed successfully. To test for correlations in this set of experiments, we perform enrichment analysis [23]. We see if any particular section of variables is enriched for a particular error state or lack of errors. We do this by using the Fisher Exact Test [20].

For these experiments, we used the global sensitivity analysis parameters listed in Table 5-1.

We binned the simulation parameters into a number of bins for labeling purposes. We experimented with using 2 bins and 4 bins to divide each of the parameter values. We then labeled each simulation with a group, or cluster, ID, dependent on the value of its parameters on a per-parameter basis. For instance, if a simulation used a value for *ksno* that was in the lower half of the *ksno* range, it would be given a *ksno* cluster label of 0. If a simulation used a value for *ksno* that was in the upper half of the range, it would be given a *ksno* cluster label of 1. With 4 bins, the per-parameter cluster labels were chosen from the set {0,1,2,3}.

The error statuses are as follows:

- **OK** : No error occurred
- **SEG** : SIGSEGV, segmentation fault occurred
- **SHR** : Unknown error submitted to shr\_abort\_abort. Stack trace terminated abnormally.
- **NEG** : negative layer thickness. timestep or remap time too large
- **CCM** : ccm kohlerc - no real(r8) solution found (quartic)
- **MPI** : MPI\_ABORT

In some cases, all the error statuses are coalesced into one class. In this case the error states are as follows:

- **OK** : No error occurred
- **ERR** : An error occurred during the simulation

Given the structure of the parameter binning and error labels, there are 4 different experimental combinations.

- Parameters are split into lower and upper halves, and error types are treated independently;
- Parameters are split into quartiles, and error types are treated independently;
- Parameters are split into lower and upper halves, and error types are coalesced;
- Parameters are split into quartiles, and error types are coalesced.

The experiments in combination A have 10 variables, split into halves, yielding 20 parameter labels. Simulation error statuses being treated independently (6 outcomes) yields a total of 120 parameter-bin/error pairings. This increases to 240 in combination B. The other experiments coalesce the errors (2 outcomes), yielding only 40 pairings in experimental combination C, and 80 in experimental combination D.

The Fisher Exact test uses these parameter-bin/error pairings to determine whether a particular pairing is over-represented given the distributions by analyzing the resultant contingency tables. For instance, it might analyze whether an MPI error is represented in each half of the *ksno* parameter values more than it should be given the distribution of OK simulations in each half. Because, in this instance, there are other types of errors, it would have to test MPI errors versus the other simulation error statuses as well and determine an aggregate p-value. However, because there are so many tests being performed, one has to adjust the raw p-values to reduce the false discovery rate. The Benjamini-Yekutieli [8] correction is used to adjust the false discovery rate under arbitrary independence assumptions. Pairings with corrected p-values  $\leq .05$  are listed below:

As shown in Table 6-1 the analysis found that two parameters have value ranges in the perturbed distribution that are associated with failures: *gamma\_coeff* and *clubb\_c8*. Both are from cloud parameterizations in the atmosphere model. Results show there is a connection between *clubb\_c8* and the segmentation fault error when the parameter value is in the upper half of the distribution. For the *gamma\_coeff* parameter there is fairly clear pattern indicating that values in the lower half of the distribution more often result in errors, while values in the upper half of the distribution

Simulation Parameter	Error Status	Range	Unadjusted P-Value	Corrected P-Value
clubb_c8	SEG	upper half	1.89E-05	7.03E-04
gamma_coeff	OK	upper half	6.19E-04	1.15E-02
gamma_coeff	SEG	lower half	5.78E-04	1.15E-02
gamma_coeff	ERR (coalesced)	lower half	6.19E-04	2.58E-03
gamma_coeff	OK (coalesced)	upper half	6.19E-04	2.58E-03
gamma_coeff	ERR (coalesced)	first quarter	3.14E-04	6.82E-03
gamma_coeff	OK (coalesced)	fourth quarter	2.7E-03	2.93E-02

**Table 6-1 Parameter-bin error pairing for pairings with corrected p-values less than 0.05.**

more often run to completion. This type of analysis can provide insight into stable parameter ranges for use in future studies. Specifically, we intend to apply an enrichment analysis approach to investigate the numerous failures we encountered for our `ne11` resolution perturbed simulations (Section 5.4).

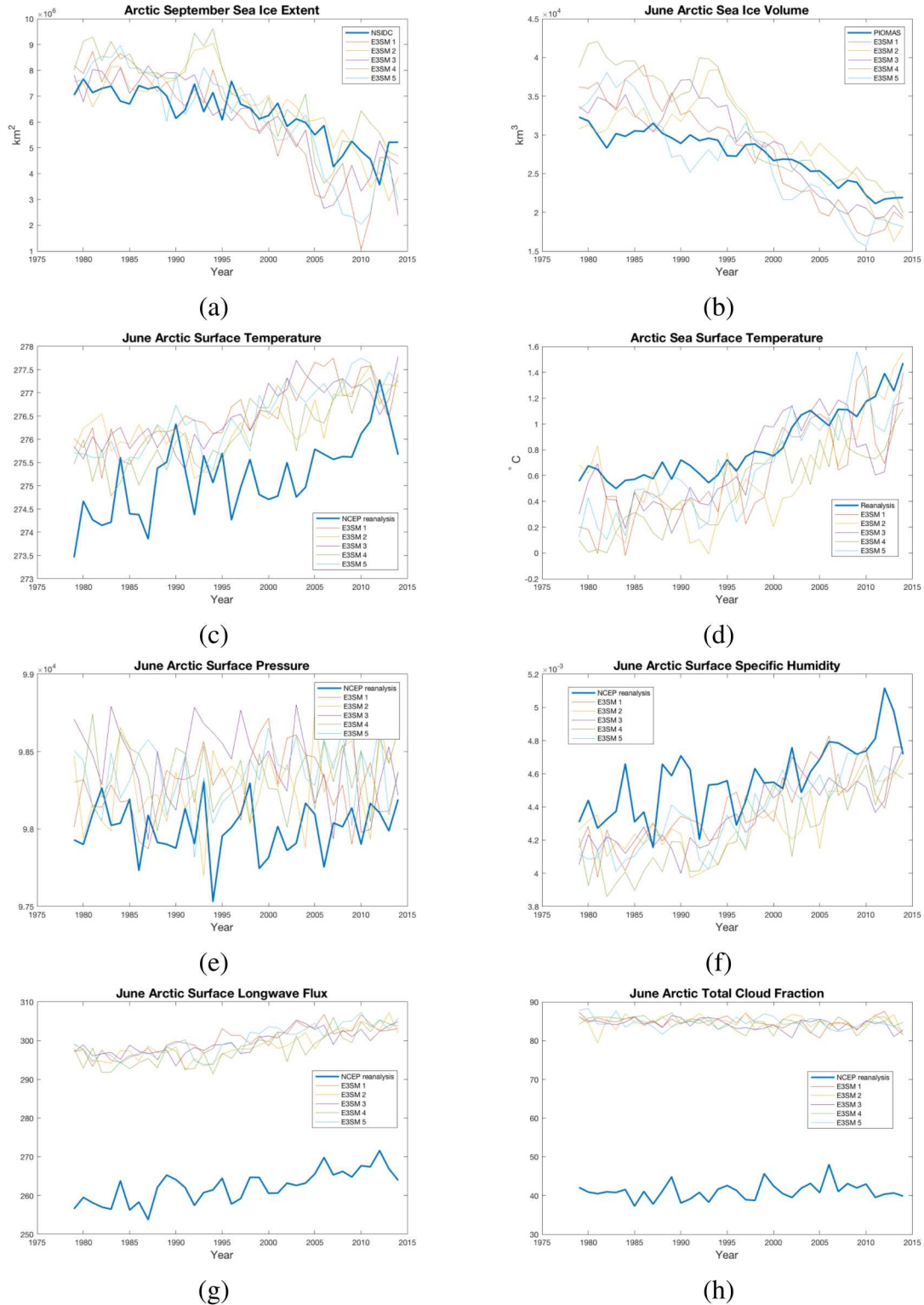
## 7. CONCLUSION

The Arctic is changing rapidly and feedbacks may be accelerating the changes. In this project we focused on analyzing Arctic sea ice in the context of the coupled Earth system. Accurate seasonal predictions of sea ice minimum extent and long-term estimates of timing for a seasonally ice-free Arctic depend on a better understanding of the factors influencing sea ice dynamics and variation. This includes loss due to greenhouse gas forcing and the superimposed internal variability of the complex Earth system.

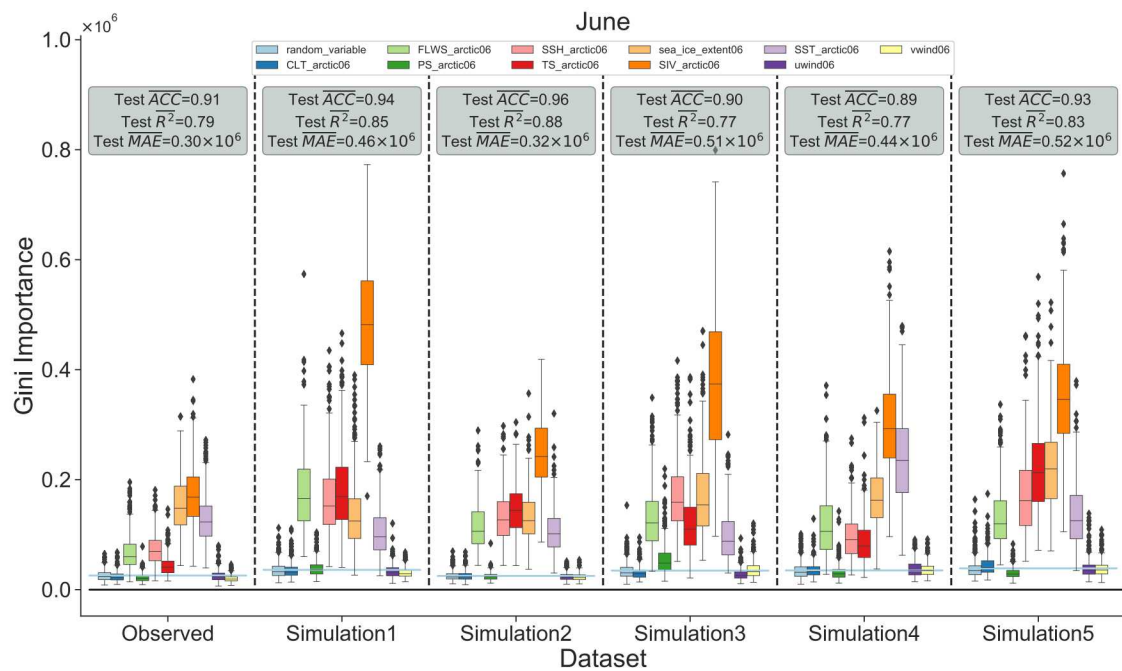
We investigated the most influential features in accurate predictions of September Arctic sea ice extent using machine learning models trained separately on observational data and on simulation data from five E3SM historical ensembles. Monthly averaged data from June, July, and August for sea ice extent (SIE), sea ice volume (SIV), surface air temperature (TS), surface pressure (PS), cloud fraction (CLT), specific humidity (HSS), surface downward long wave flux (FLWS), sea surface temperature (SST), and meridional and zonal winds (UWIND, VWIND) were used to train a random forest regression model. Gini importance measures were computed for each input feature with the testing data. We found that SIE, SIV, TS, SSH and FLWS had importance values greater than a random input for each dataset, but the magnitude of the importance value difference between simulations and observations. When training with June data SIV is most important for all five historical ensemble simulations, but SIE is most important data-model feature when training with August data. Further analysis of how feature importance changes with forecast length could provide insights into multiyear, seasonal, and subseasonal predictability.

We have additionally performed a global sensitivity analysis (GSA) using a fully coupled ultra-low resolution configuration E3SM. To our knowledge, this is the first global sensitivity analysis involving the fully-coupled E3SM Earth system model. We have found that parameter variations show significant impact on the Arctic climate state and atmospheric parameters related to cloud parameterizations are the most significant. We also find significant interactions between parameters

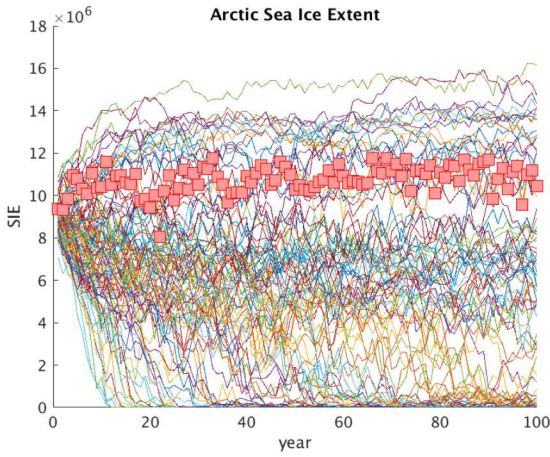
of different components of E3SM. The results of this study provide invaluable insight into the relative importance of various parameters from the sea ice, atmosphere and ocean components of the E3SM (including cross-component parameter interactions) on various Arctic-focused quantities of interest (QOIs). Additionally, our GSA results can inform targeted studies at higher resolutions, as well as help guide model spin-ups at various resolutions.



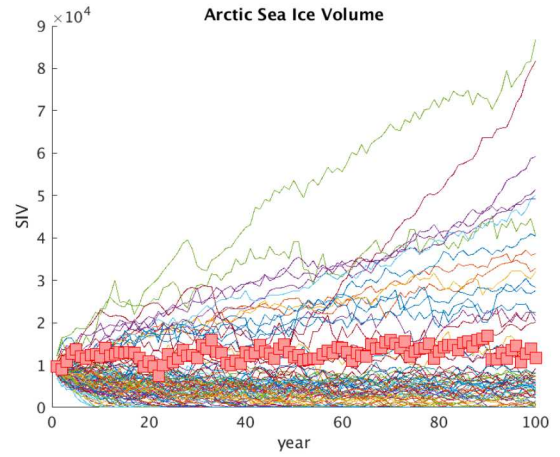
**Figure 4-1 Comparison between E3SM simulations and NCEP observational data used in the data-driven model for (a) September sea ice extent, (b) June sea ice volume, (c) June Arctic surface air temperature, (d) June Arctic sea surface temperature, (e) June Arctic surface pressure, (f) June Arctic specific humidity, (g) June Arctic surface longwave flux, (h) June Arctic total cloud fraction.**



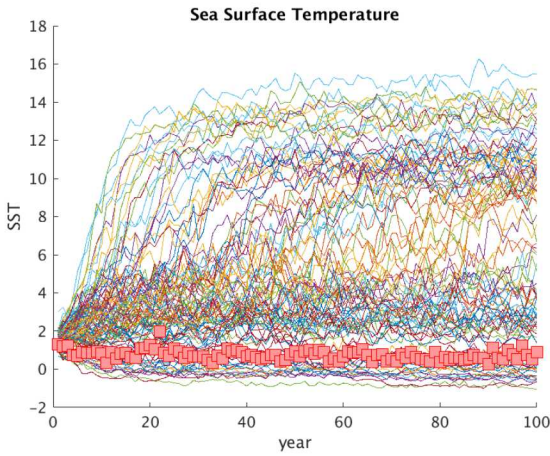
**Figure 4-2** Box plot of June feature importance values for random 13 year predictions generated by 385 models. The  $\overline{R^2}$ , anomaly correlation coefficient ( $\overline{ACC}$ ), and mean absolute error ( $\overline{MAE}$ ) is displayed towards the top. The blue line crossing each dataset is the importance of an included random variable in each feature set.



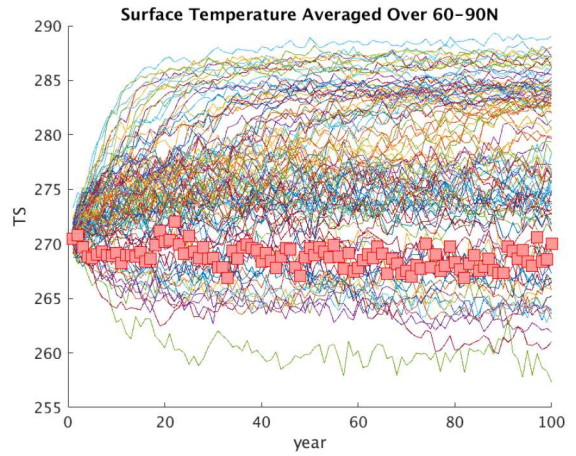
(a) SIE



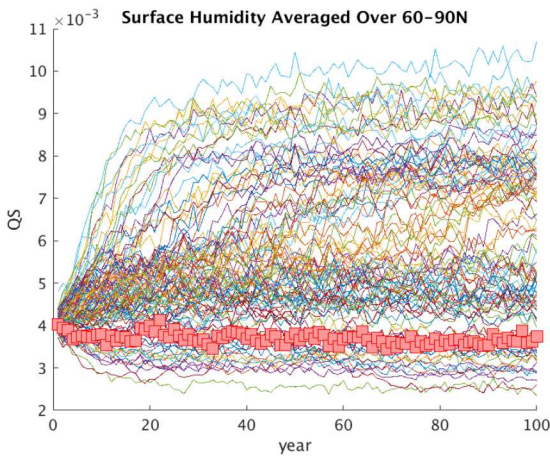
(b) SIV



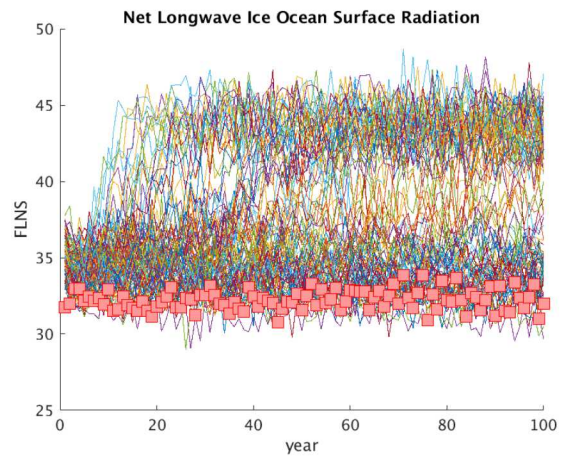
(c) SST



(d) TS

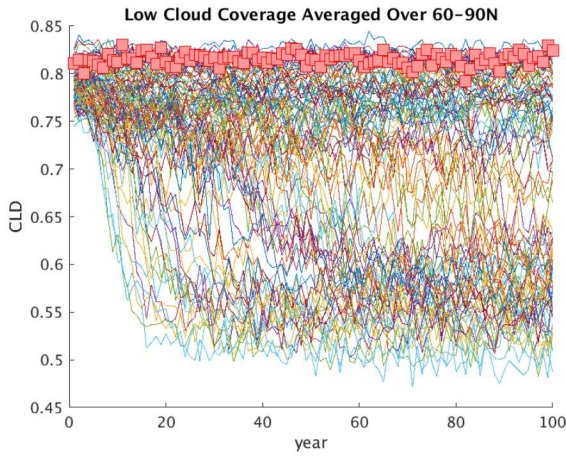


(e) QS

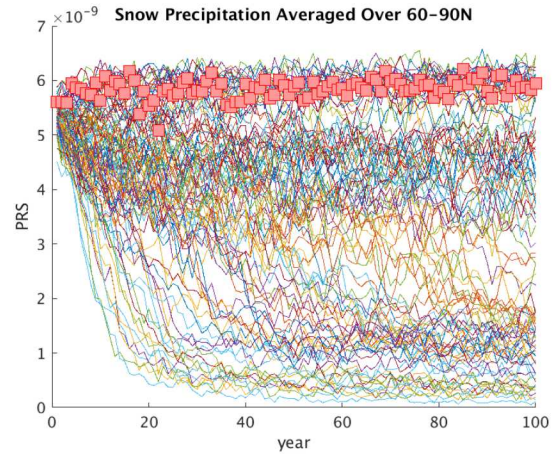


(f) FLNS

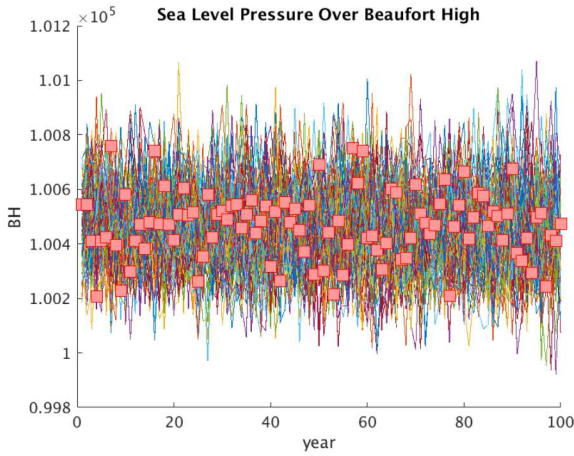
Figure 5-1 “Spaghetti plots” of first six QOIs from Table 5-2 for *ne4* runs that made it to year 100. The baseline run is distinguished from the others by the red markers.



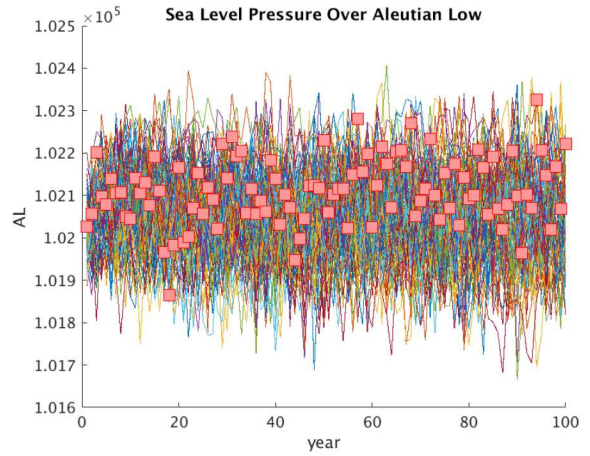
(a) CLDLOW



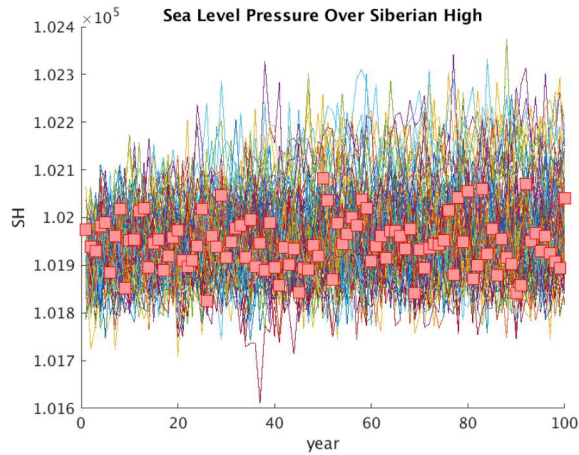
(b) PRECSL



(c) BH

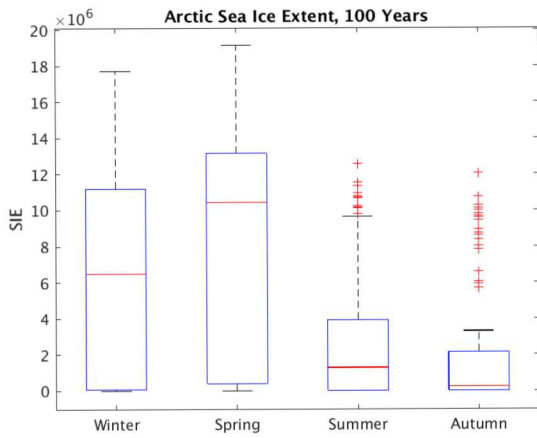


(d) AL

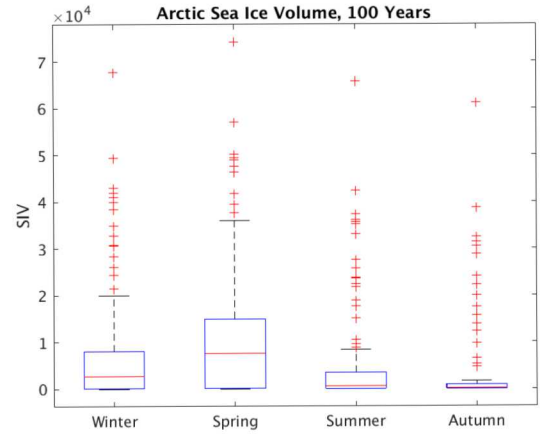


(f) SH

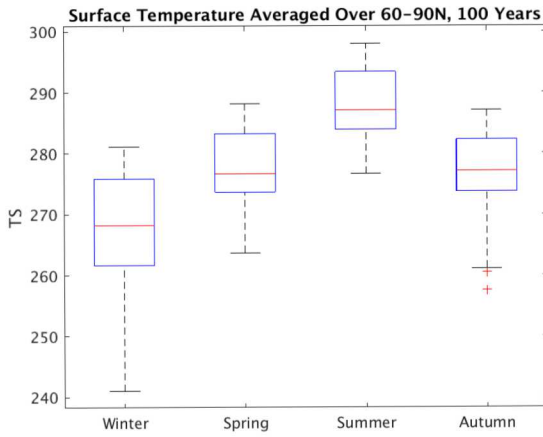
Figure 5-2 “Spaghetti plots” of last five QOIs from Table 5-2 for *ne4* runs that made it to year 100. The baseline run is distinguished from the others by the red markers



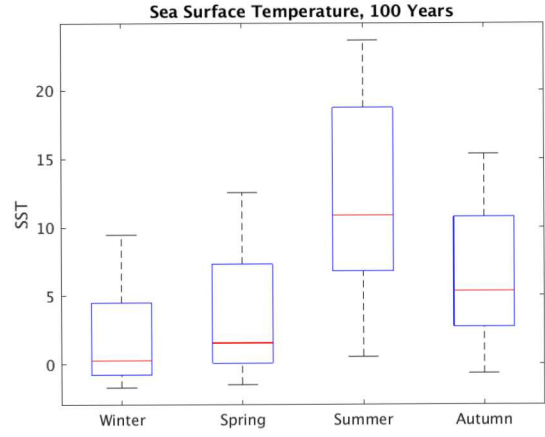
(a) SIE



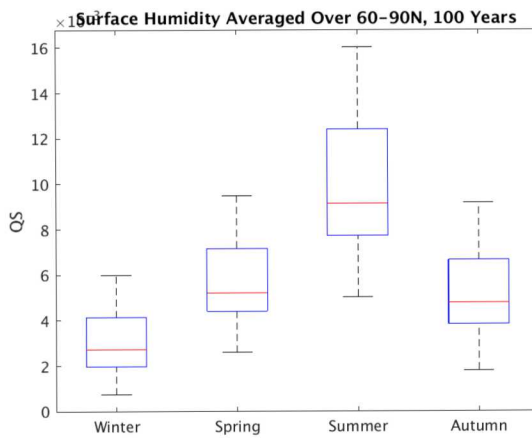
(b) SIV



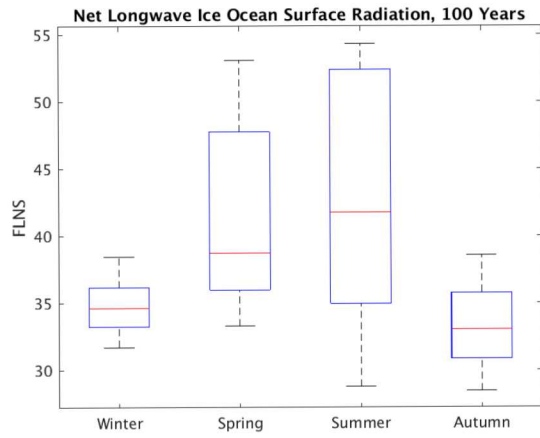
(c) SST



(d) TS

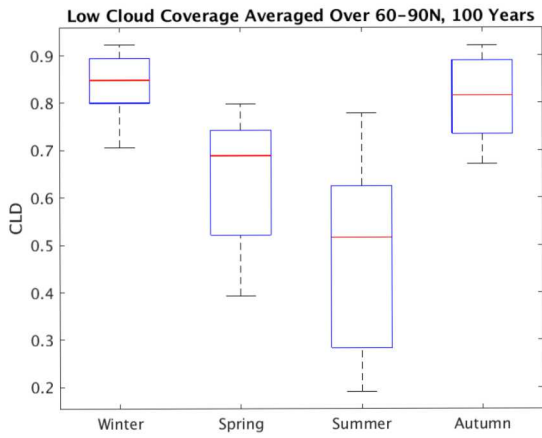


(e) QS

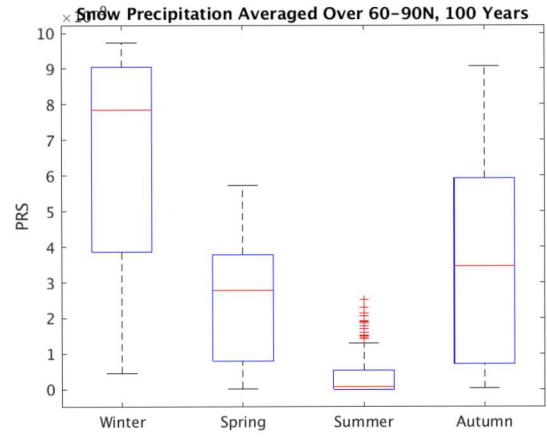


(f) FLNS

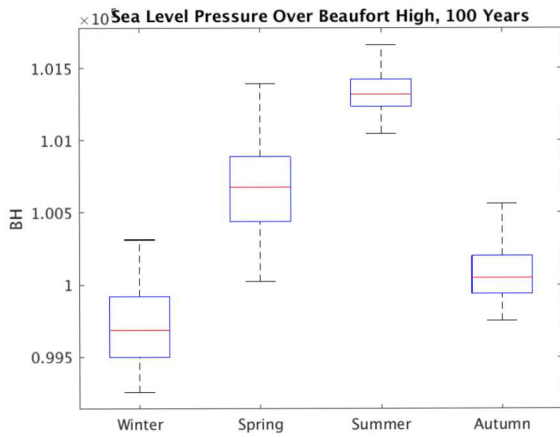
**Figure 5-3** Box plots of first six QOIs for perturbed `ne4` runs of 100 years, averaged over the last 50 years.



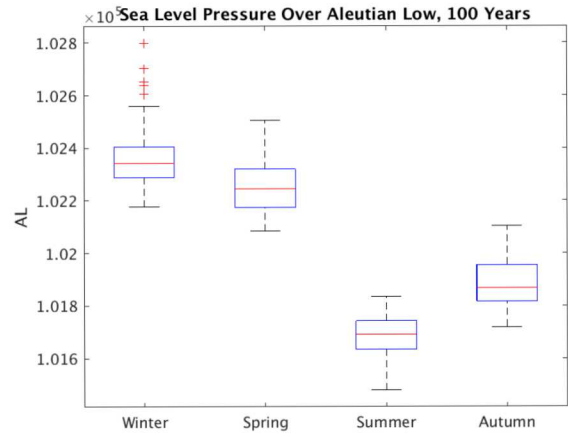
(a) CLDLOW



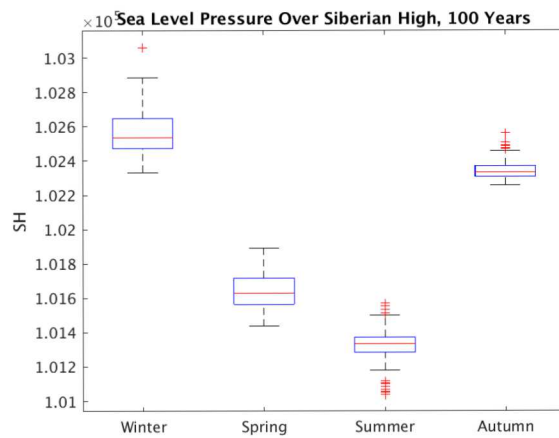
(b) PRECSL



(c) BH

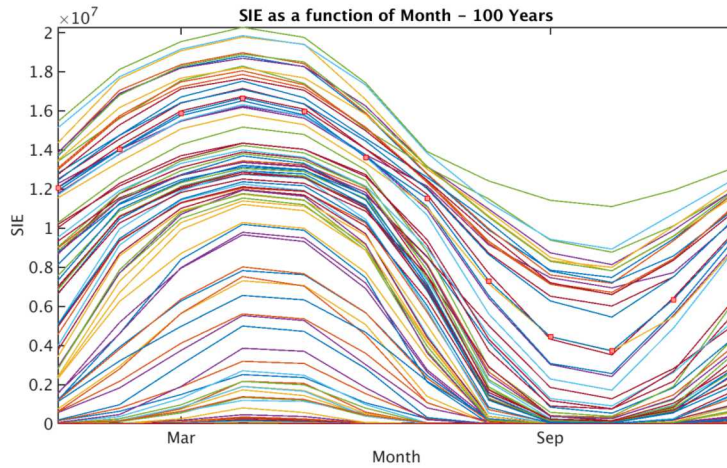


(d) AL

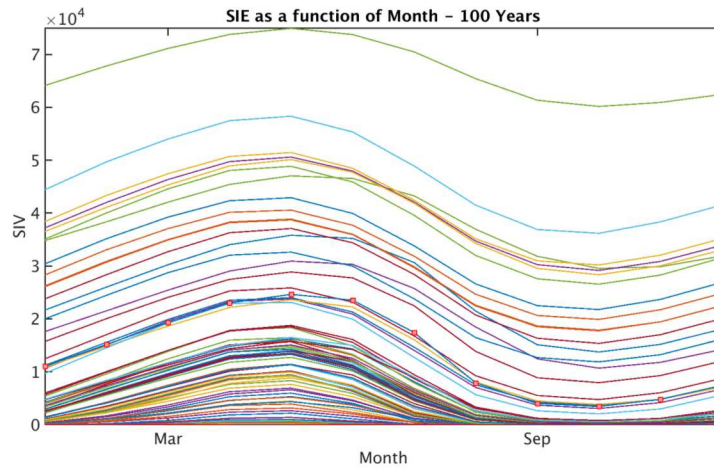


(e) SH

Figure 5-4 Box plots of last five QOIs for perturbed  $ne_4$  runs of 100 years, averaged over the last 50 years.

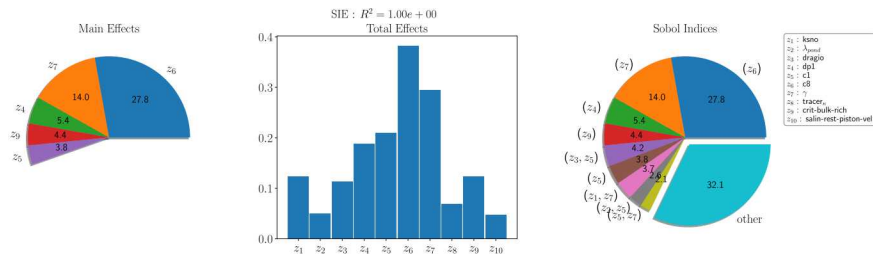


(a) SIE

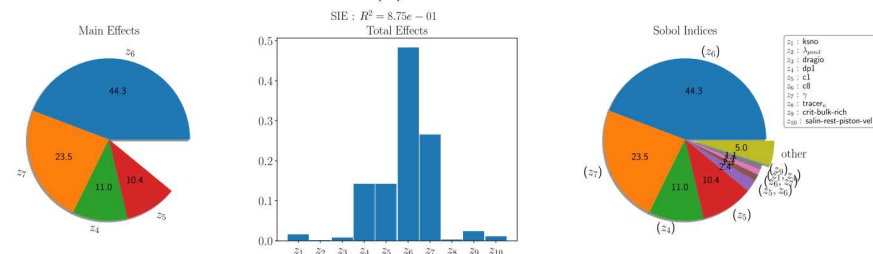


(b) SIV

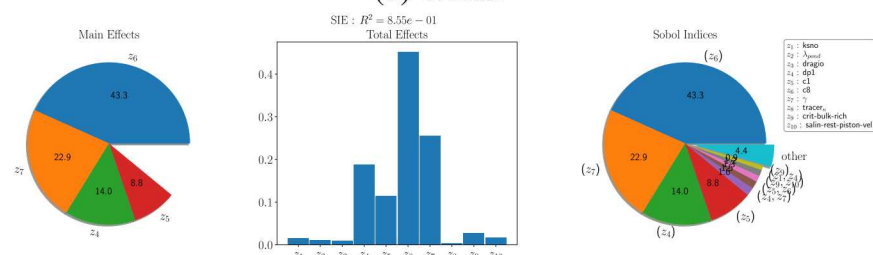
**Figure 5-5 “Spaghetti plots” of Arctic sea ice extent and Arctic sea ice volume for  $ne4$  runs that made it to 100 years as a function of the month. The maximum/minimum Arctic sea ice extent occur in general April/October, respectively, and the maximum/minimum Arctic sea ice volume occur in general in May/October, respectively. The baseline run is distinguished from the others by the red markers.**



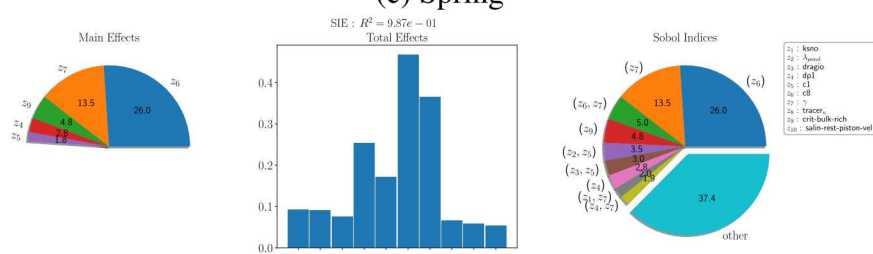
(a) Annual



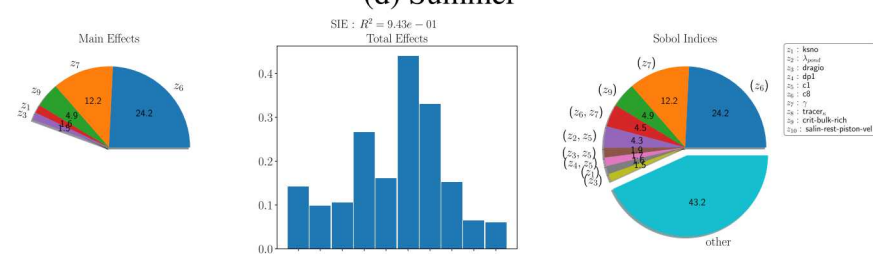
(b) Winter



(c) Spring

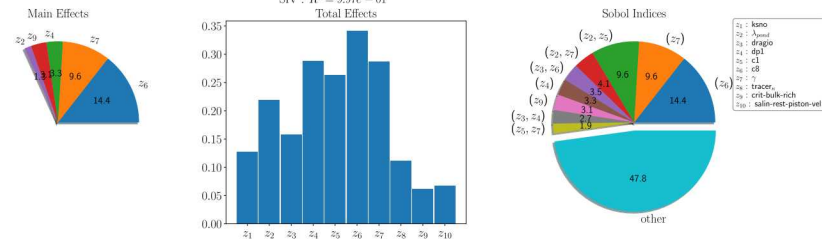


(d) Summer

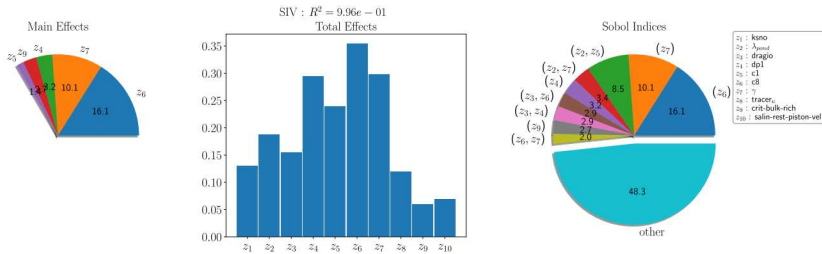


(c) Autumn

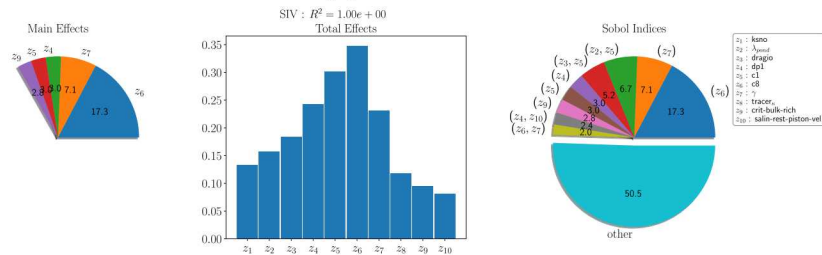
Figure 5-6 GSA results for ne4 resolution: main effects, total effects and Sobol indices for Sea Ice Extent (SIE) QOI calculated annually and by season



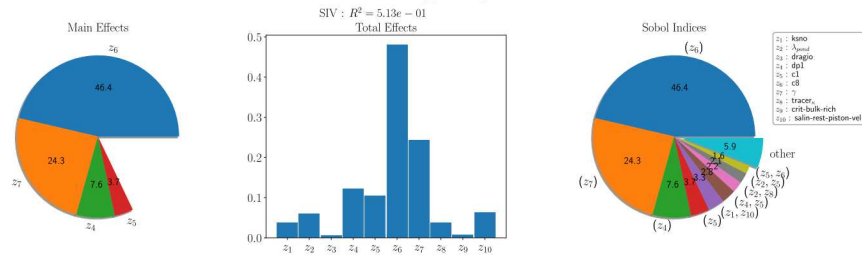
(a) Annual



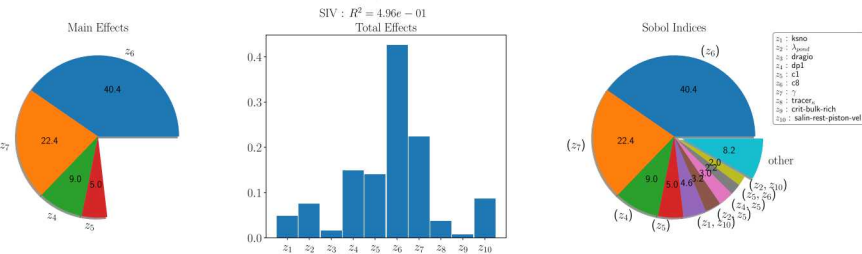
(b) Winter



(c) Spring

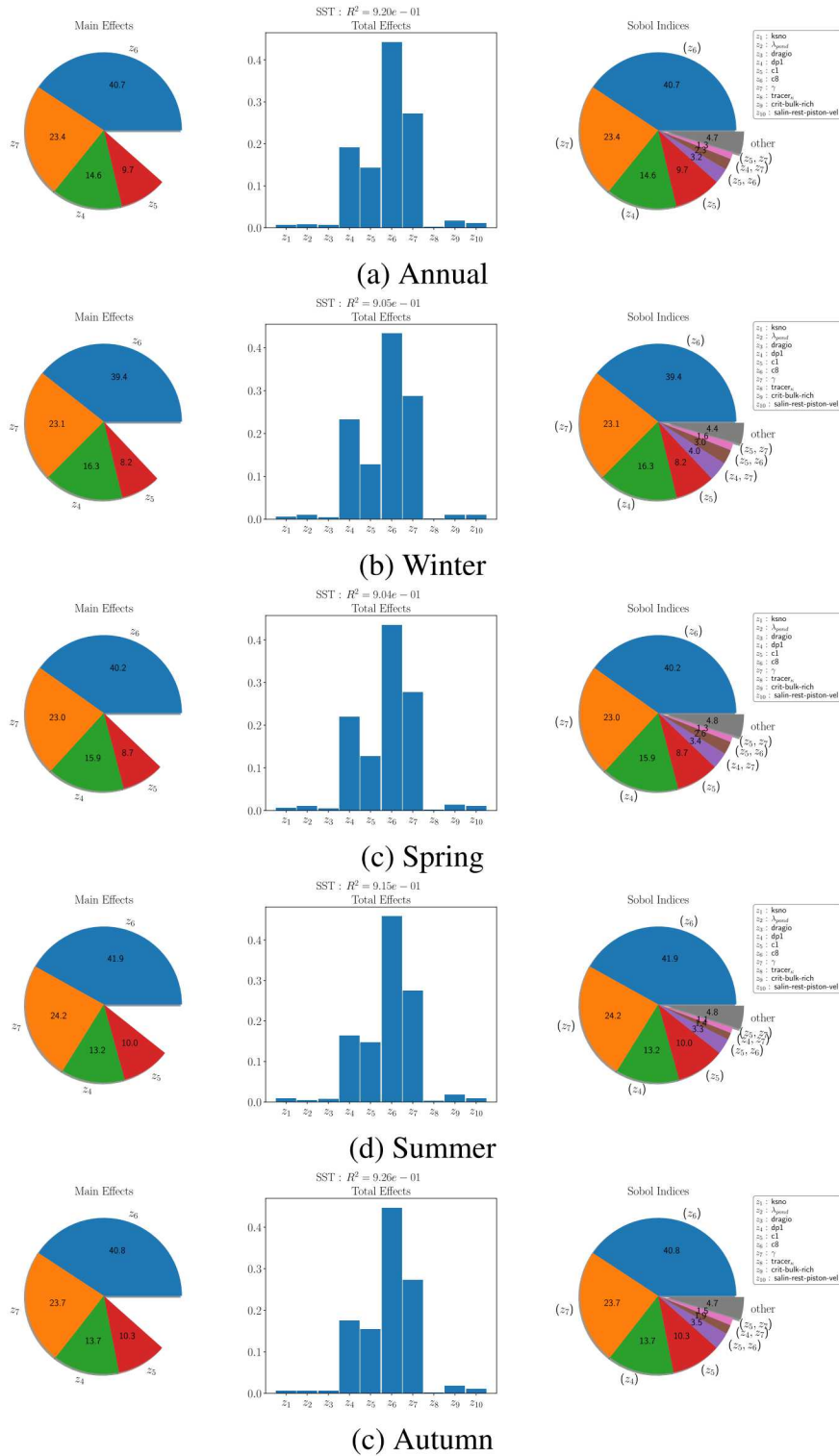


(d) Summer

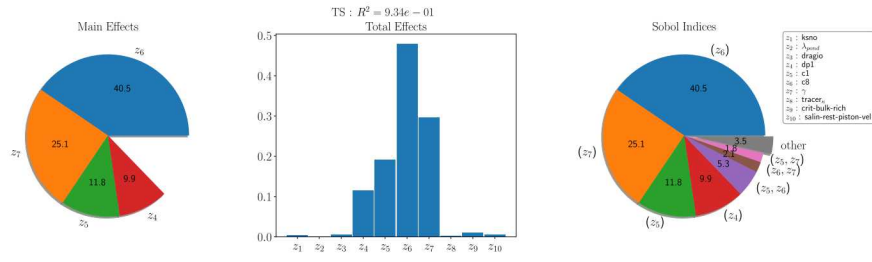


(c) Autumn

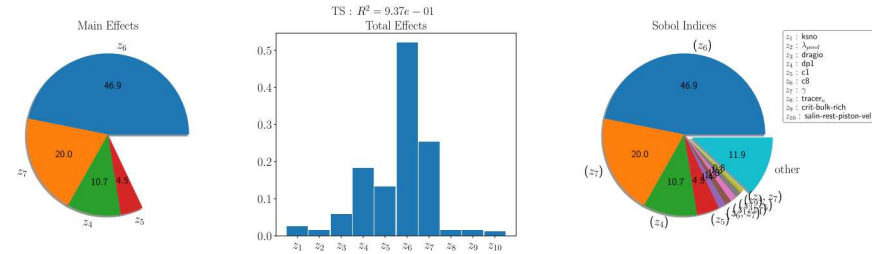
Figure 5-7 GSA results for ne4 resolution: main effects, total effects and Sobol indices for Sea Ice Volume (SIV) QOI calculated annually and by season



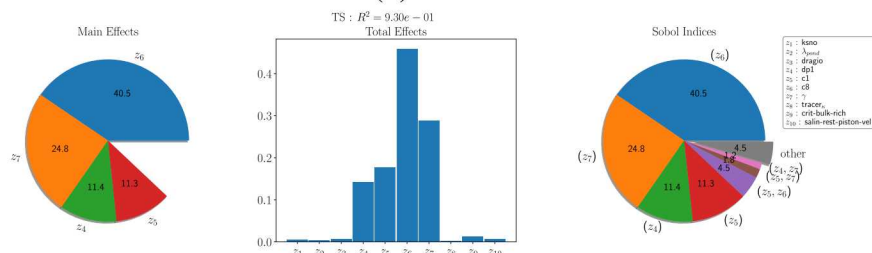
**Figure 5-8 GSA results for  $ne_4$  resolution: main effects, total effects and Sobol indices for Sea Surface Temperature Averaged Over 60-90° (SST) QOI calculated annually and by season**



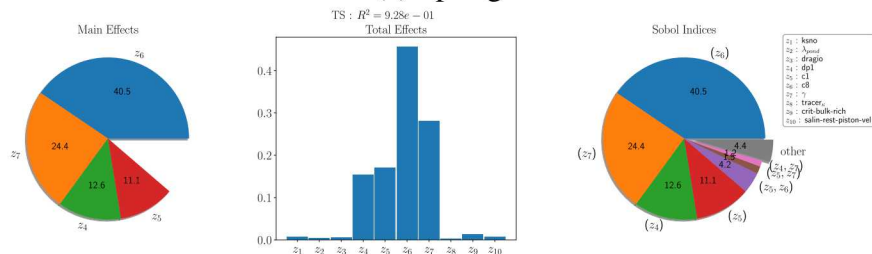
(a) Annual



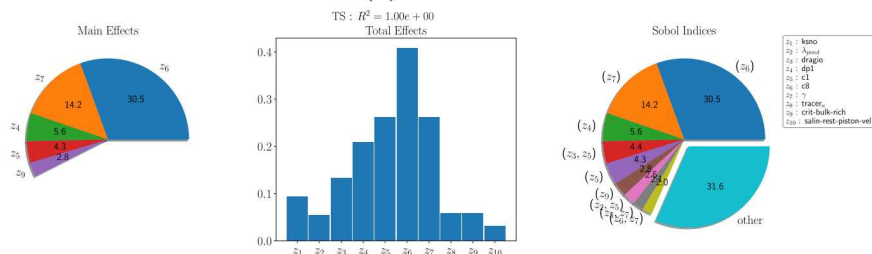
(b) Winter



(c) Spring

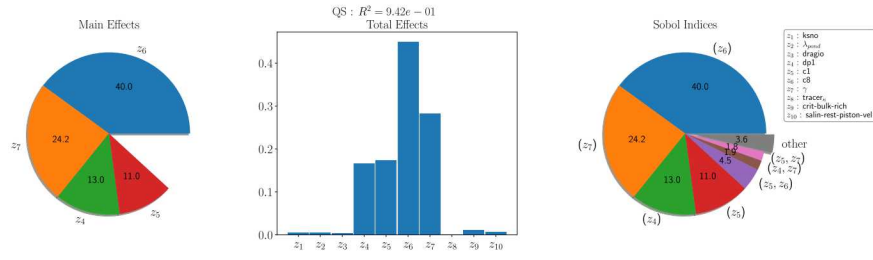


(d) Summer

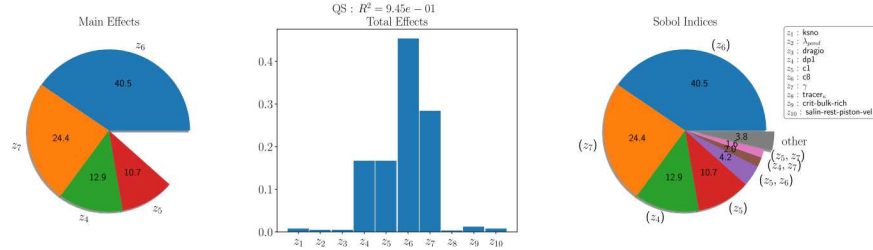


(e) Autumn

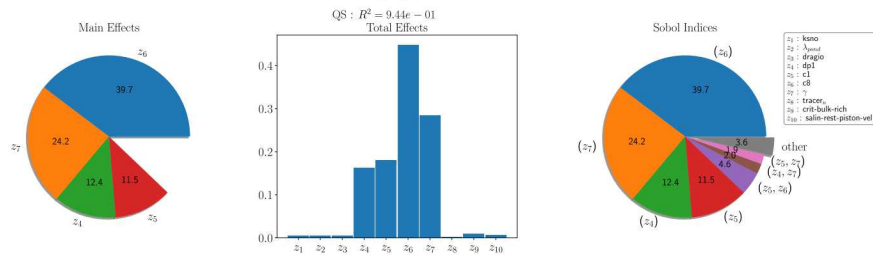
**Figure 5-9 GSA results for ne4 resolution: main effects, total effects and Sobol indices for Surface Temperature Averaged Over 60-90° (TS) QOI calculated annually and by season**



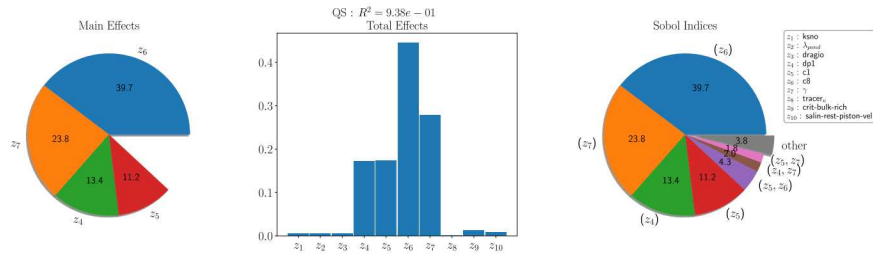
(a) Annual



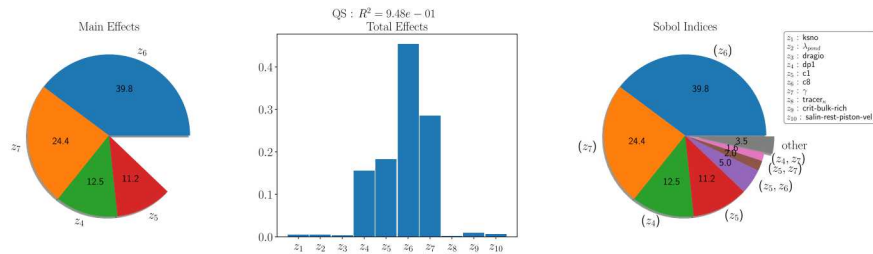
(b) Winter



(c) Spring



(d) Summer



(e) Autumn

**Figure 5-10 GSA results for ne4 resolution: main effects, total effects and Sobol indices for Specific Humidity Averaged Over 60-90° (QS) QOI calculated annually and by season**

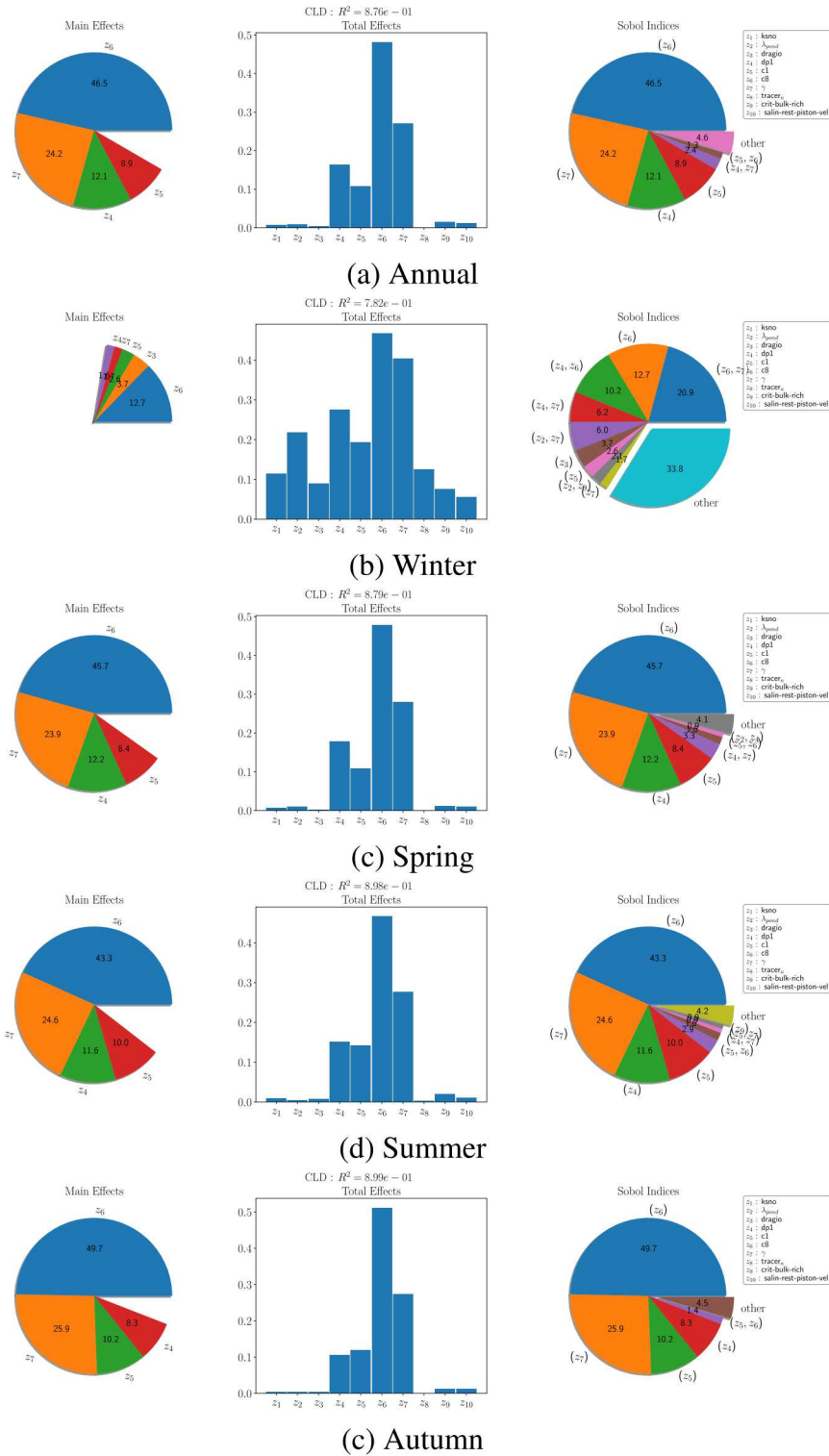
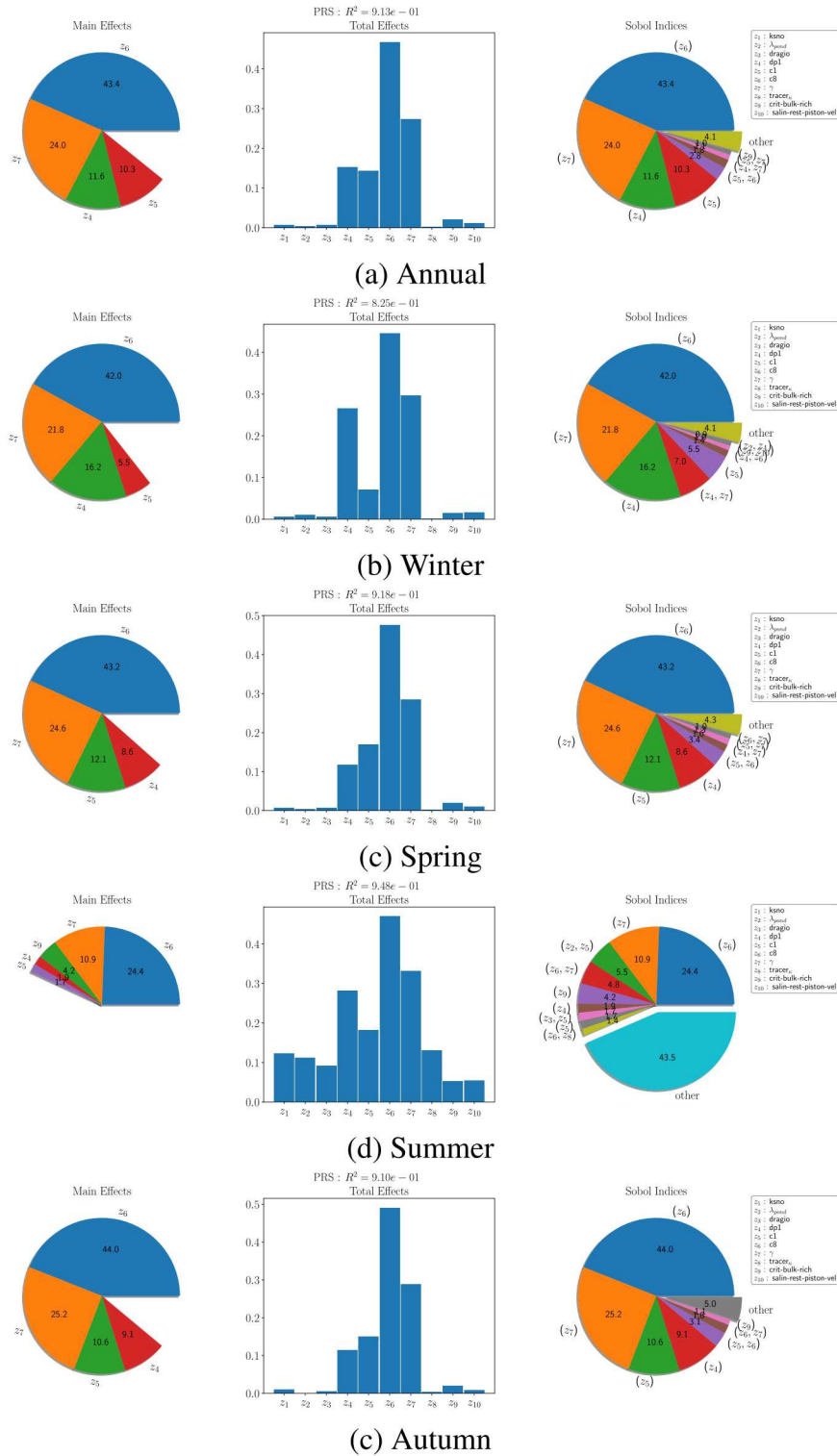
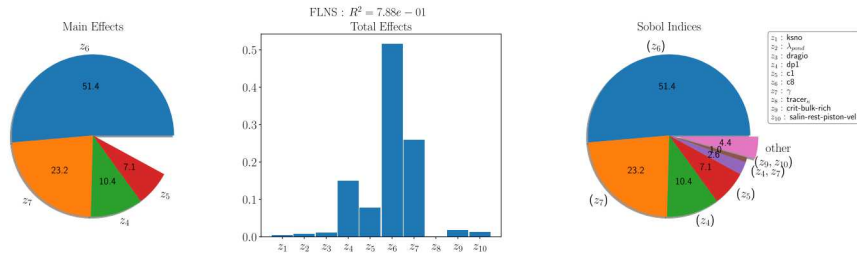


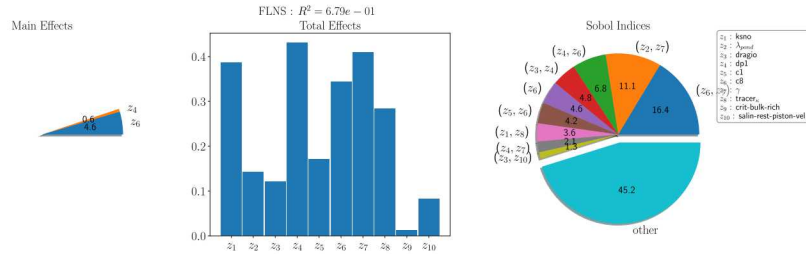
Figure 5-11 GSA results for  $ne4$  resolution: main effects, total effects and Sobol indices for Low Cloud Coverage Averaged Over 60-90° (CLDLow) QOI calculated annually and by season



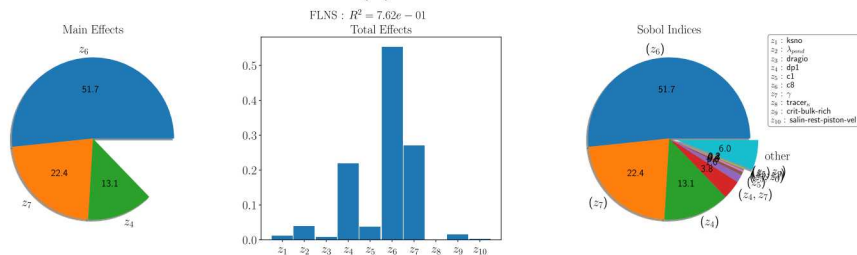
**Figure 5-12 GSA results for ne4 resolution: main effects, total effects and Sobol indices for Snow Precipitation Averaged Over 60-90° (PRECSL) QOI calculated annually and by season**



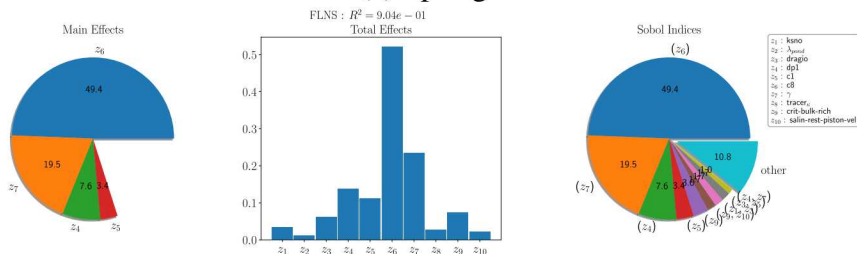
(a) Annual



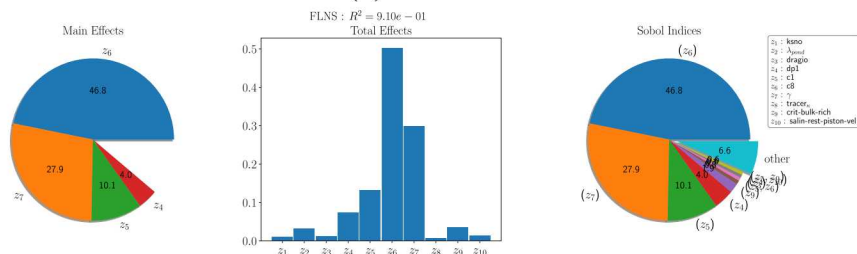
(b) Winter



(c) Spring

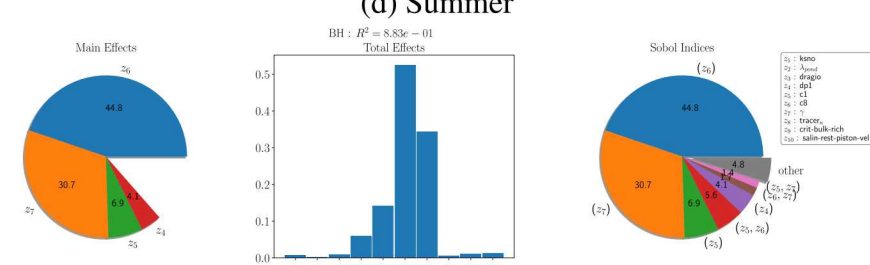
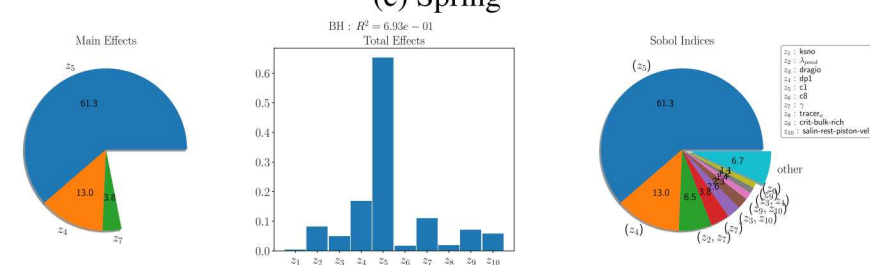
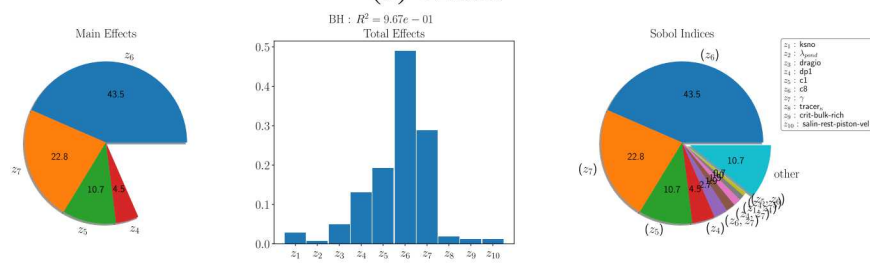
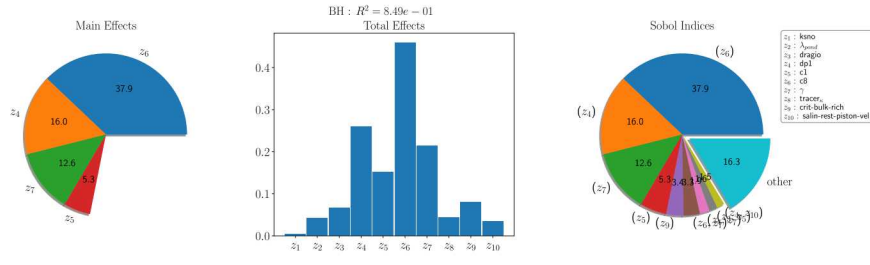
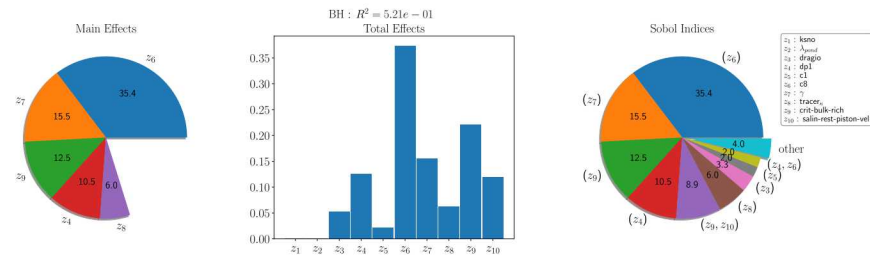


(d) Summer

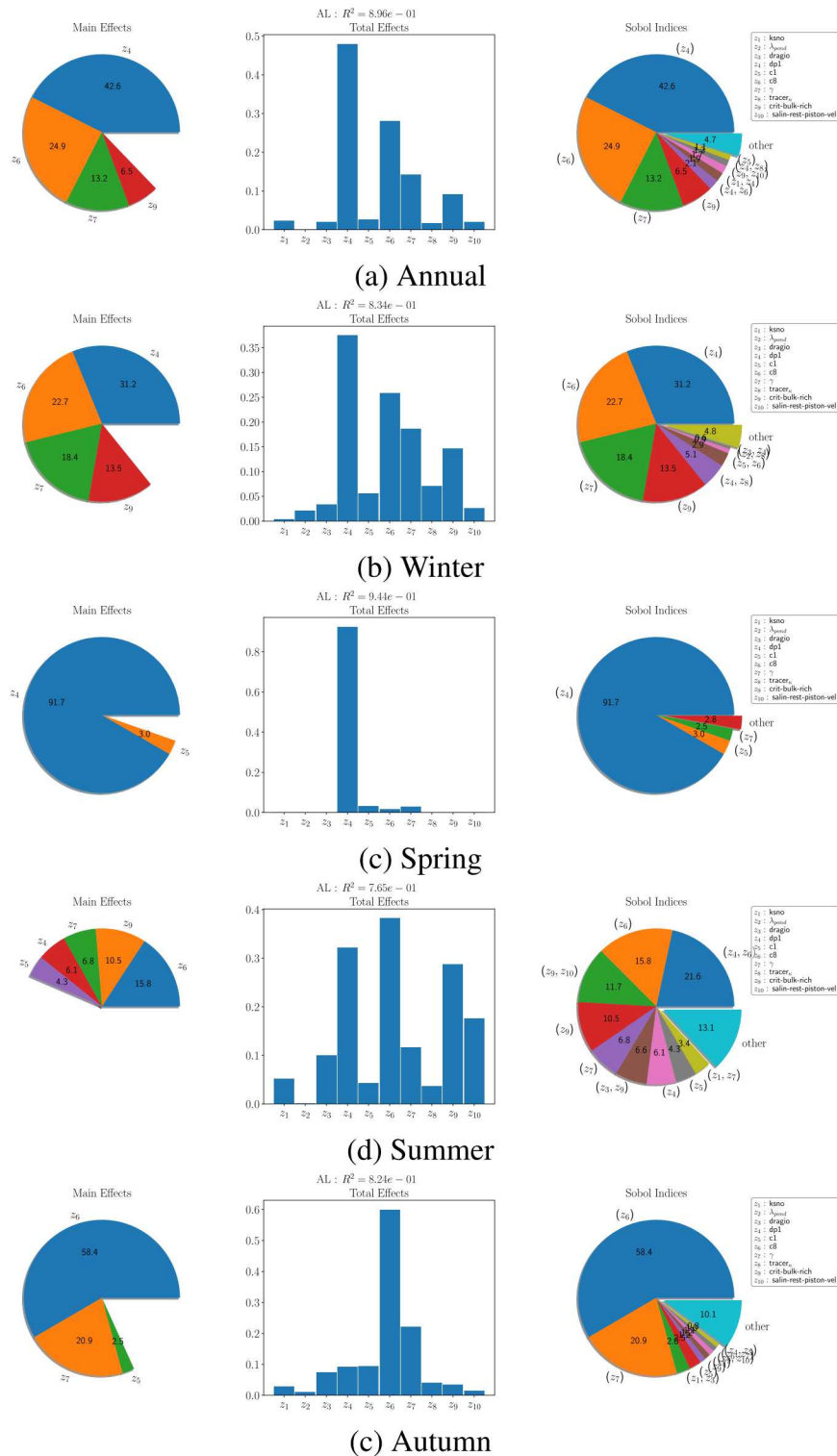


(e) Autumn

Figure 5-13 GSA results for ne4 resolution: main effects, total effects and Sobol indices for Net Longwave Ice Ocean Surface Radiation Averaged Over 60-90° (FLNS) QOI calculated annually and by season



**Figure 5-14 GSA results for ne4 resolution: main effects, total effects and Sobol indices for Sea Level Pressure Over Beaufort High (BH) QOI calculated annually and by season**



**Figure 5-15 GSA results for ne4 resolution: main effects, total effects and Sobol indices for Sea Level Pressure Over Aleutian Low (AL) QOI calculated annually and by season**

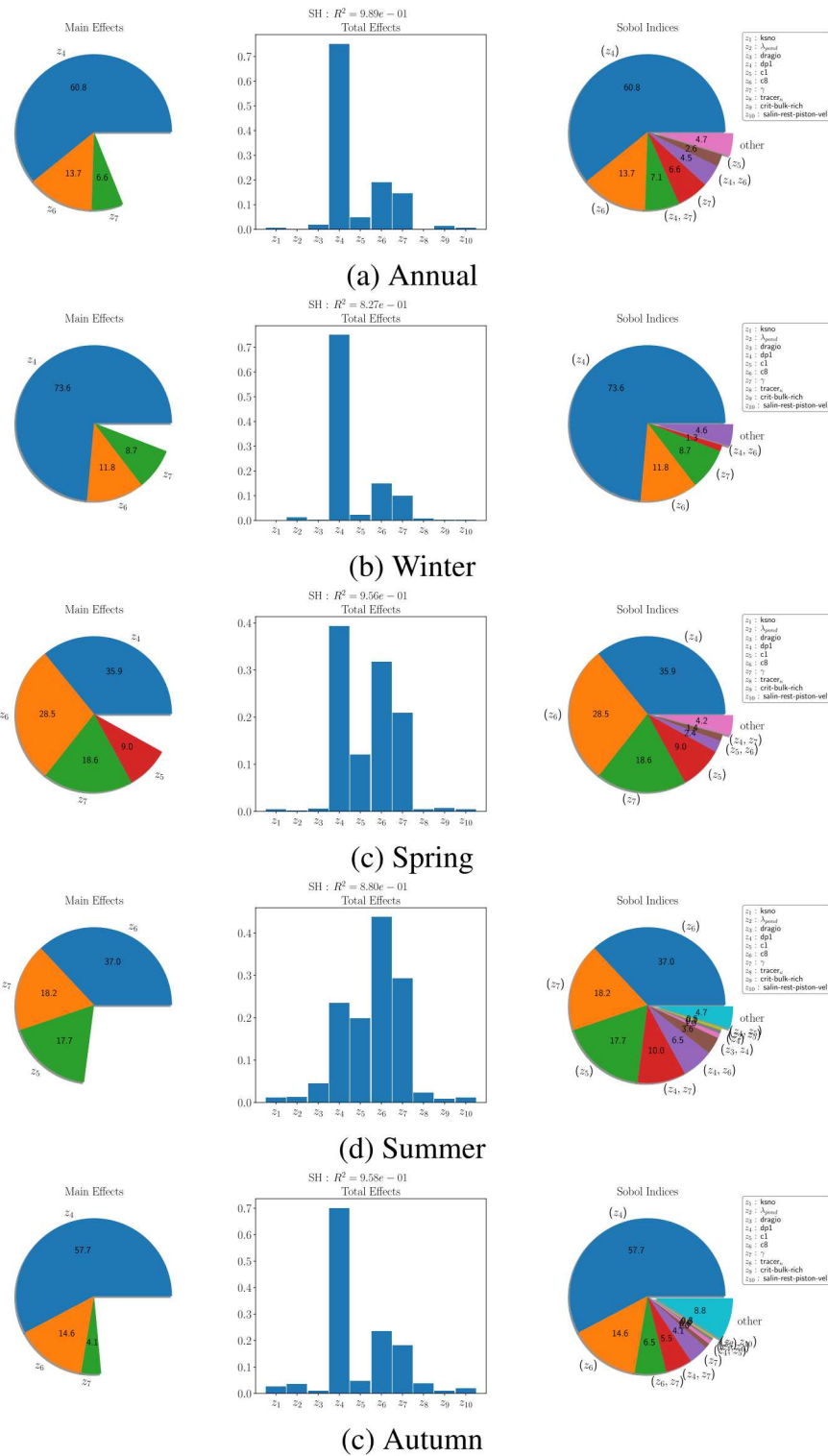
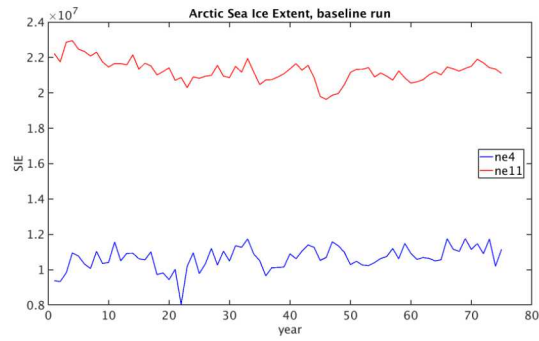
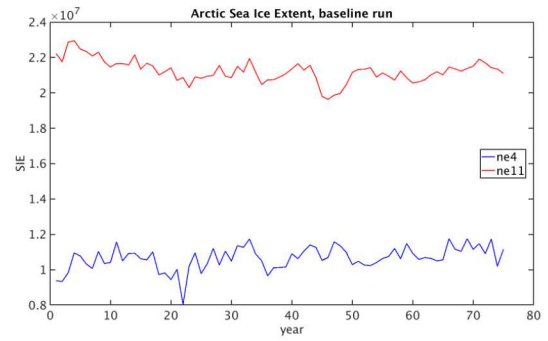


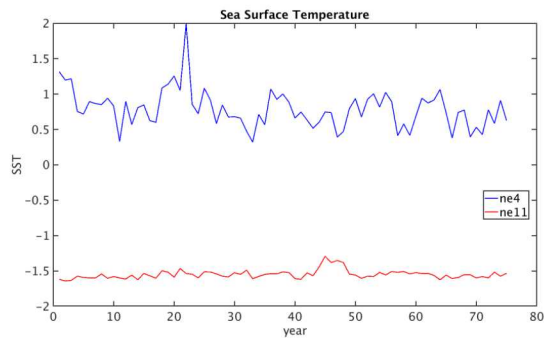
Figure 5-16 GSA results for  $ne4$  resolution: main effects, total effects and Sobol indices for Sea Level Pressure Over Siberian High (SH) QOI calculated annually and by season



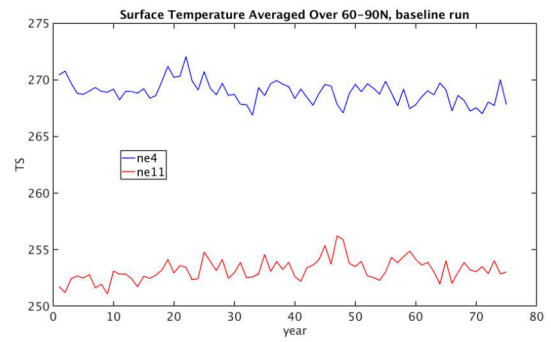
(a) SIE



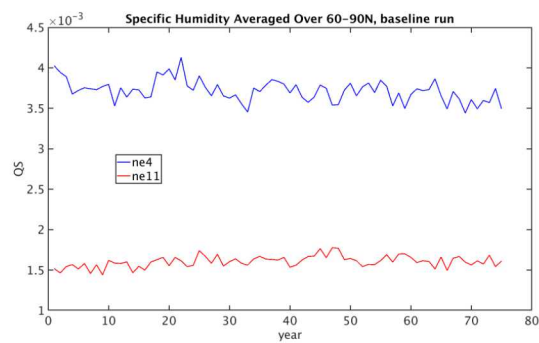
(b) SIV



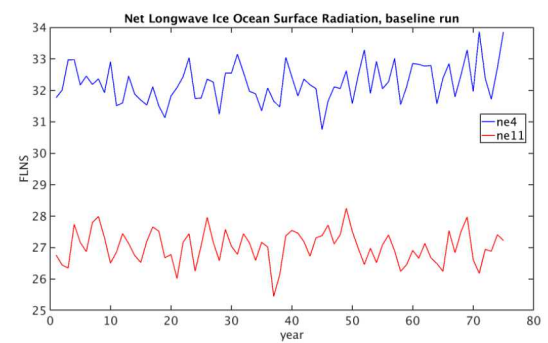
(c) SST



(d) TS

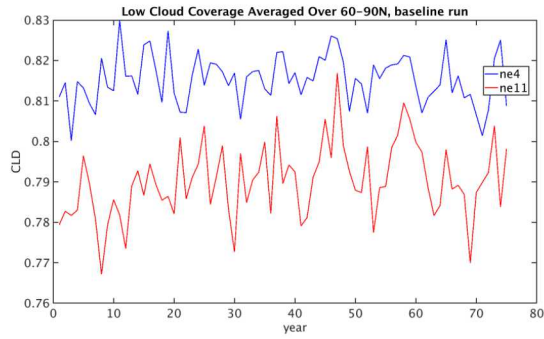


(e) QS

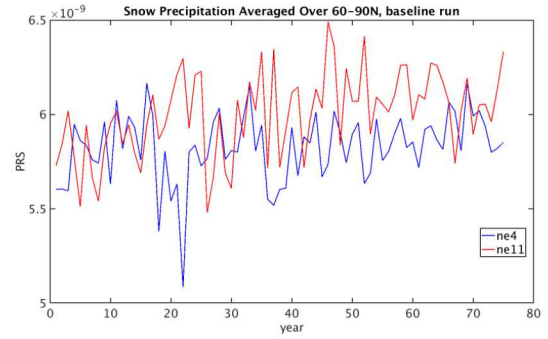


(f) FLNS

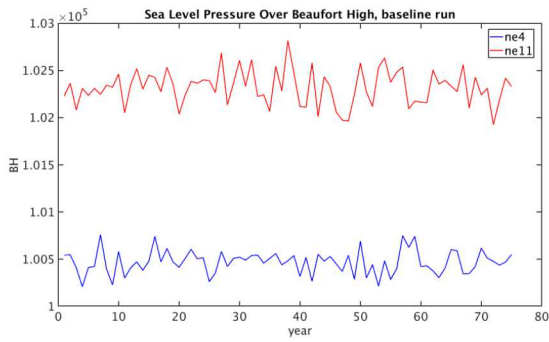
**Figure 5-17 “Spaghetti plots” of first six QOIs from Table 5-2 for ne4 and ne11 baseline runs.**



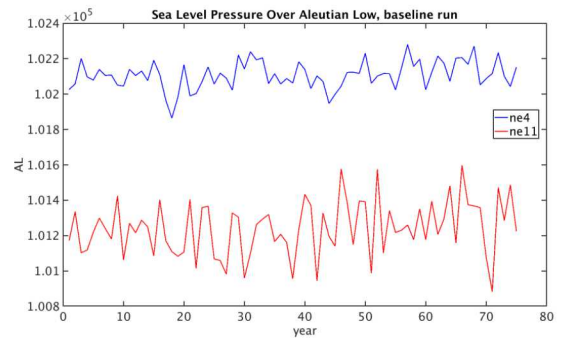
(a) CLDLow



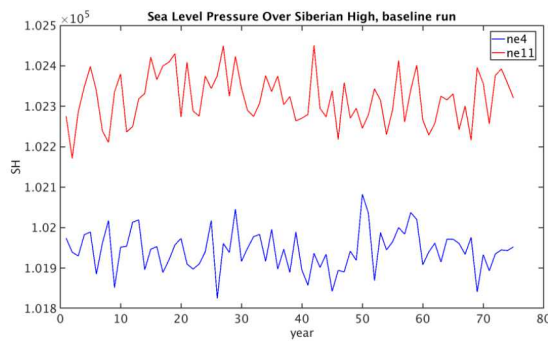
(b) PRECSL



(c) BH

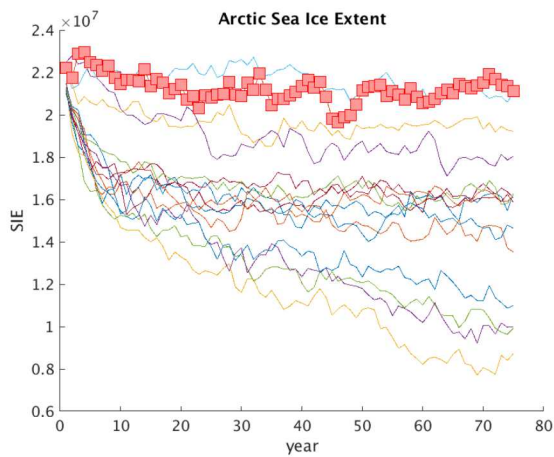


(d) AL

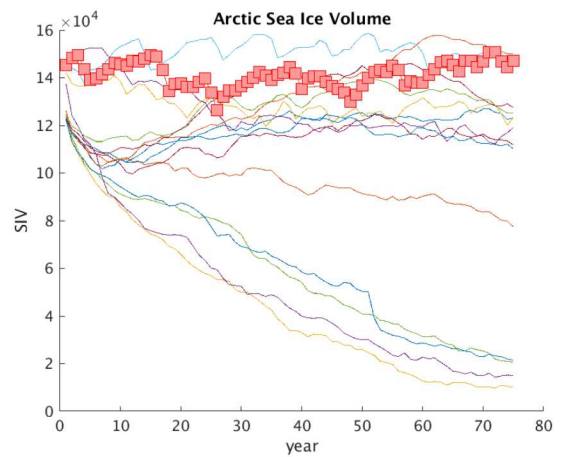


(f) SH

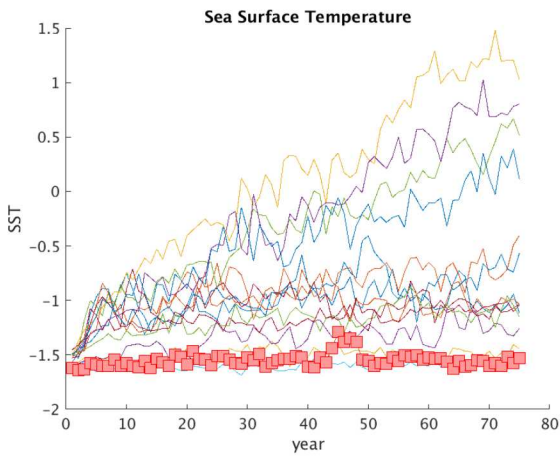
Figure 5-18 “Spaghetti plots” of last five QOIs from Table 5-2 for ne4 and ne11 baseline runs.



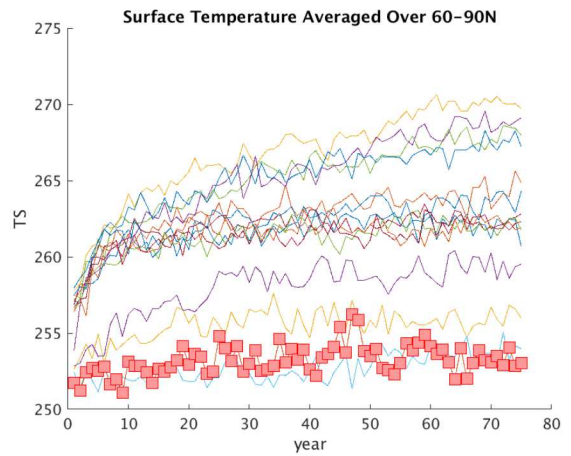
(a) SIE



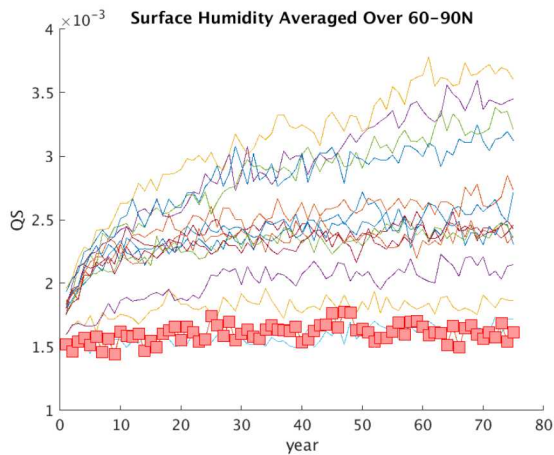
(b) SIV



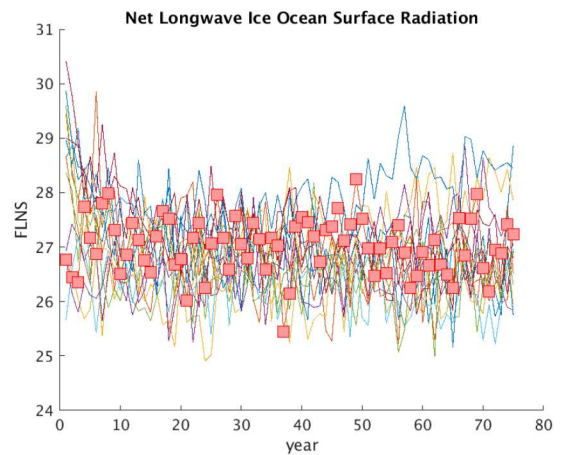
(c) SST



(d) TS

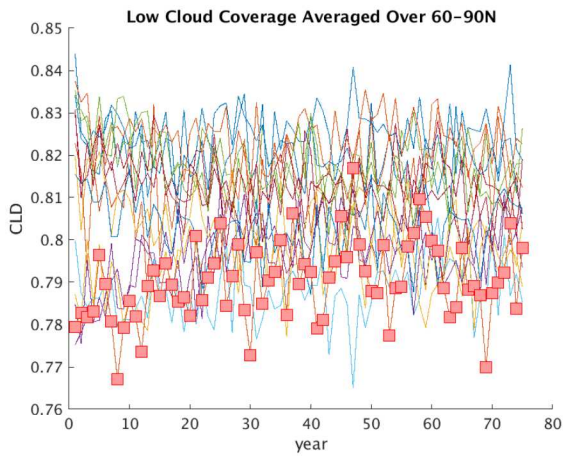


(e) QS

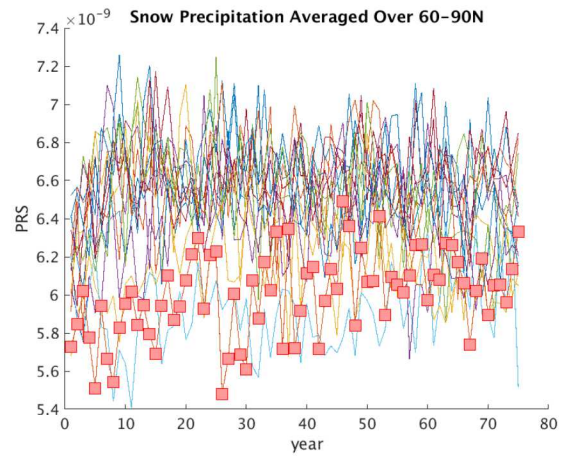


(f) FLNS

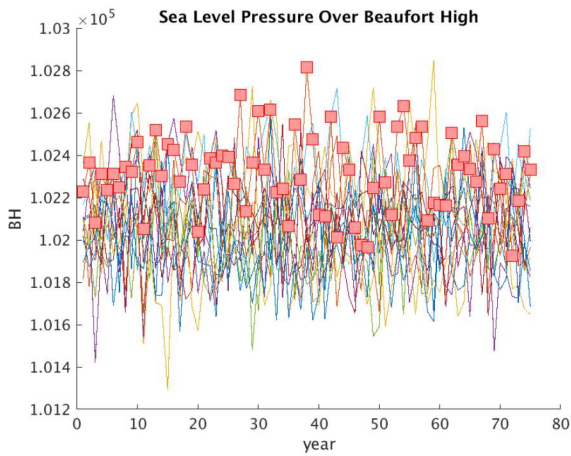
Figure 5-19 “Spaghetti plots” of first six QOIs from Table 5-2 for `ne11` runs that made it to year 75. The baseline run is distinguished from the others by the red markers.



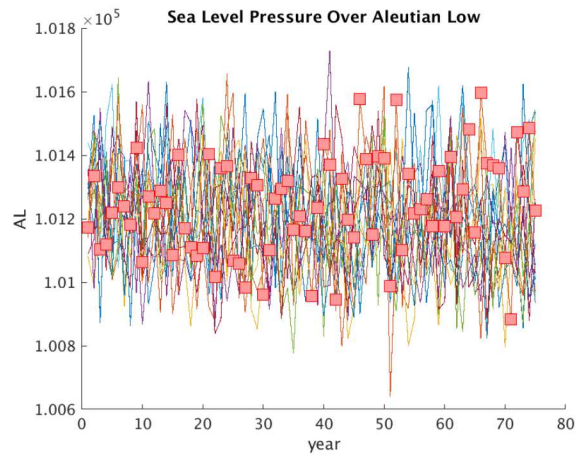
(a) CLDLOW



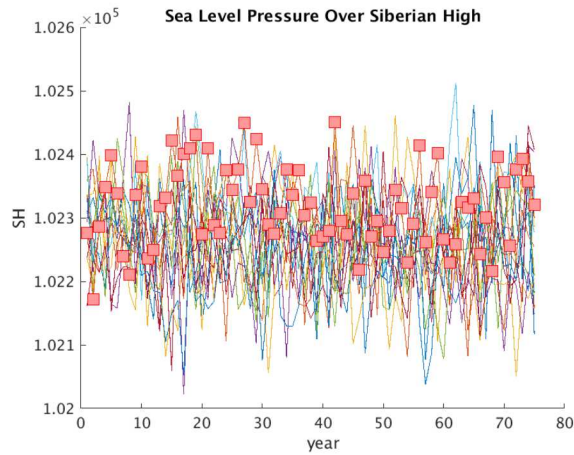
(b) PRECSL



(c) BH

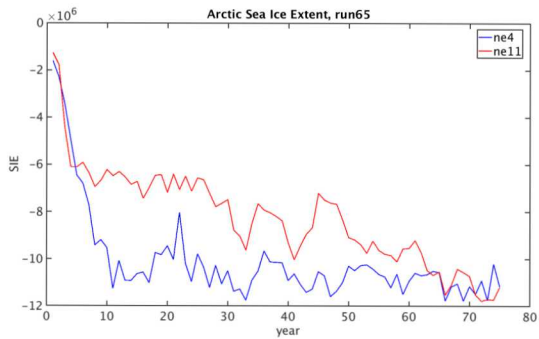


(d) AL

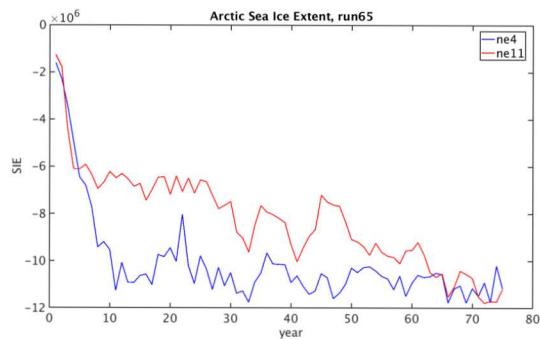


(f) SH

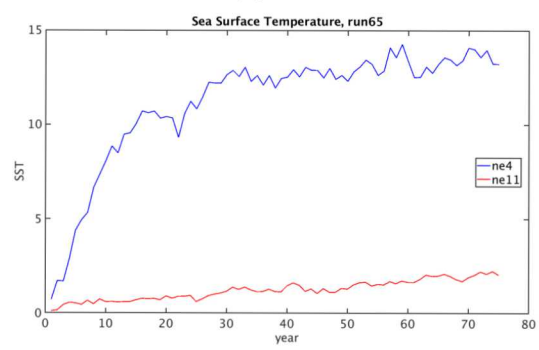
Figure 5-20 “Spaghetti plots” of last five QOIs from Table 5-2 for `ne11` runs that made it to year 75. The baseline run is distinguished from the others by the red markers



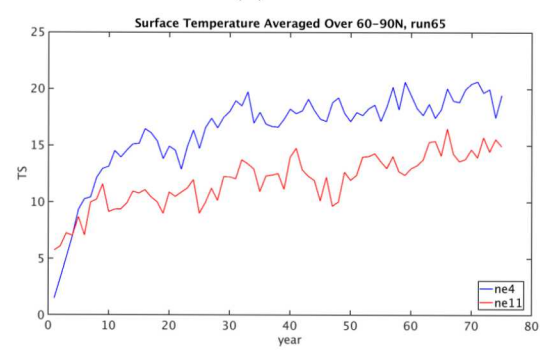
(a) SIE



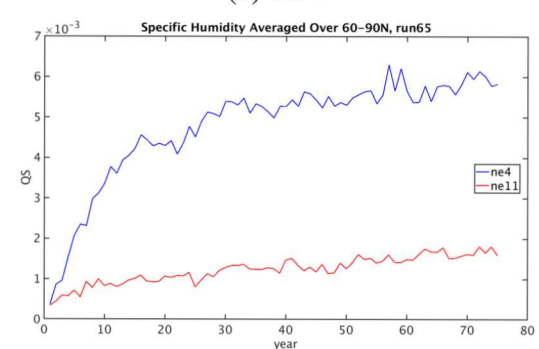
(b) SIV



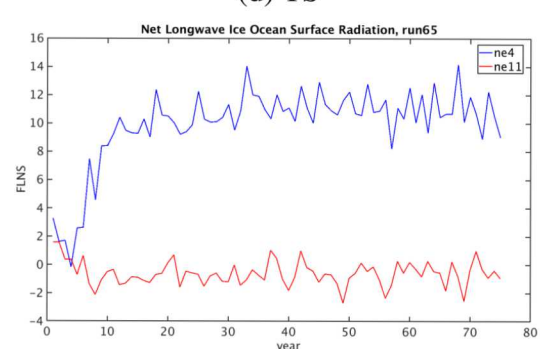
(c) SST



(d) TS

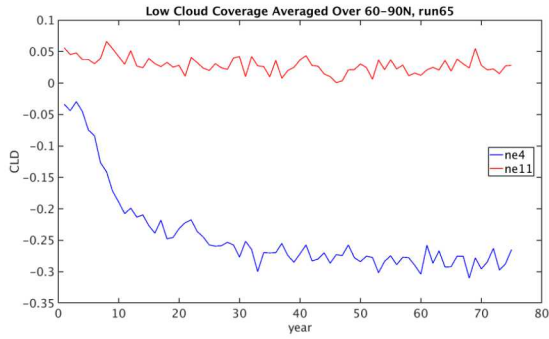


(e) QS

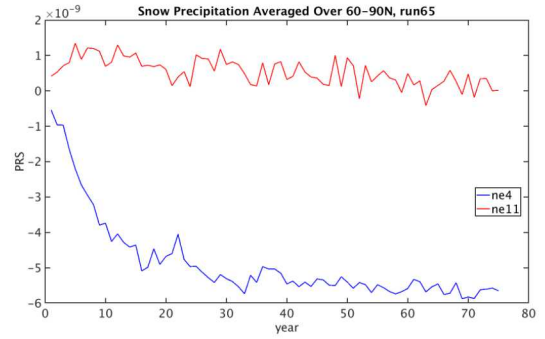


(f) FLNS

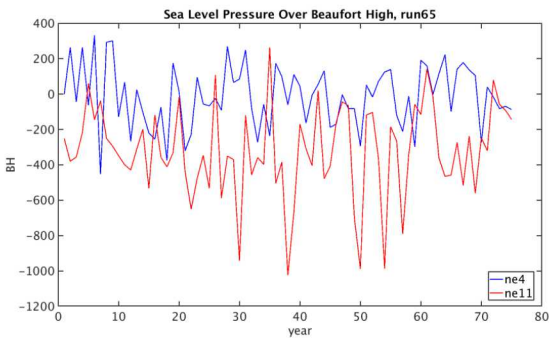
Figure 5-21 “Spaghetti plots” of first six QOIs from Table 5-2 for ne4 and ne11 run #65, with baseline QOIs subtracted.



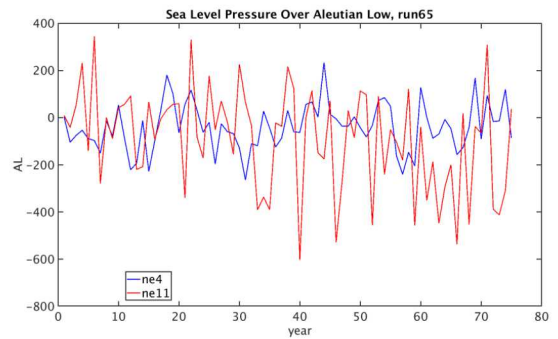
(a) CLDLOW



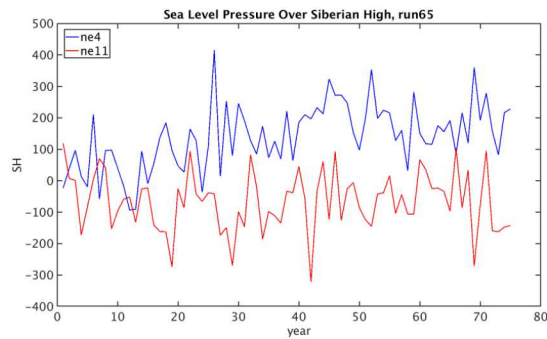
(b) PRECSL



(c) BH

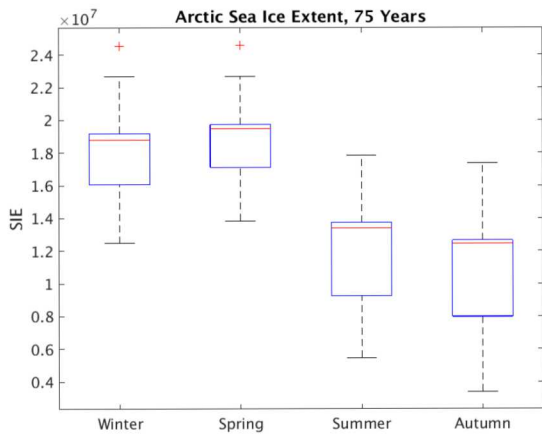


(d) AL

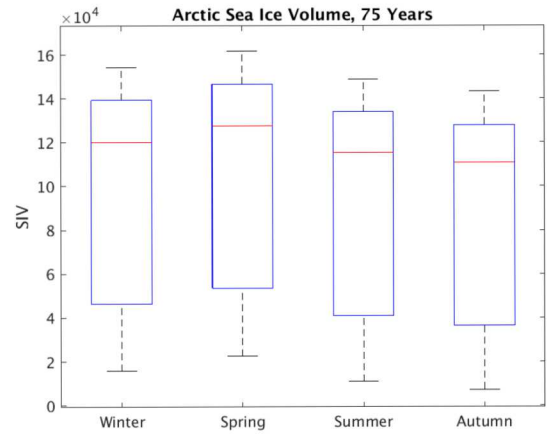


(f) SH

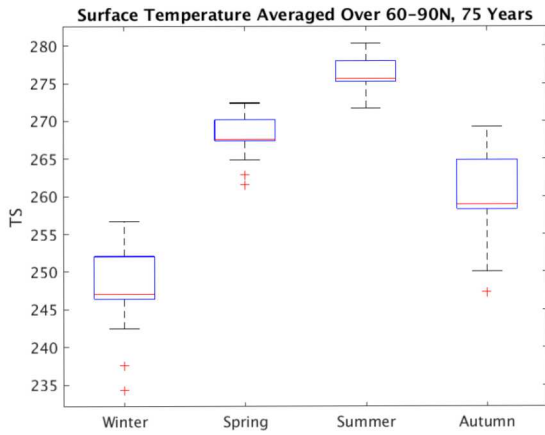
Figure 5-22 “Spaghetti plots” of last five QOIs from Table 5-2 for ne4 and ne11 run #65, with baseline QOIs subtracted.



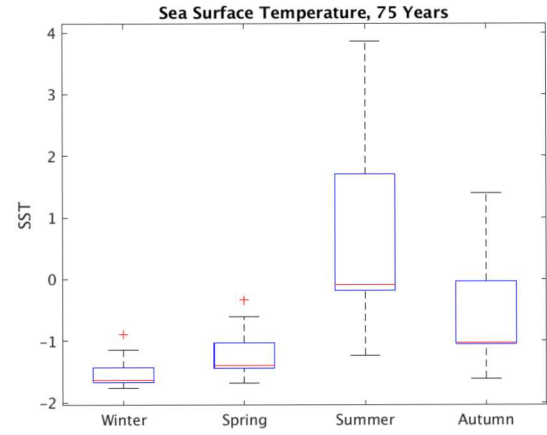
(a) SIE



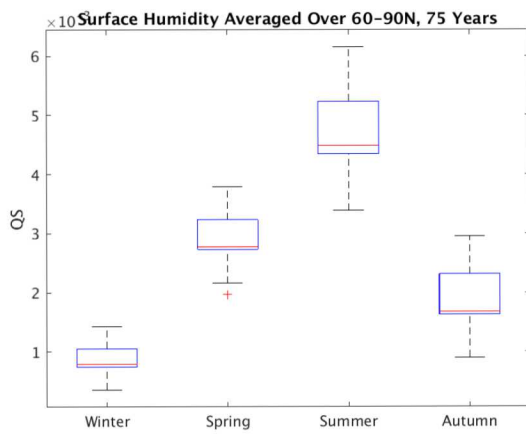
(b) SIV



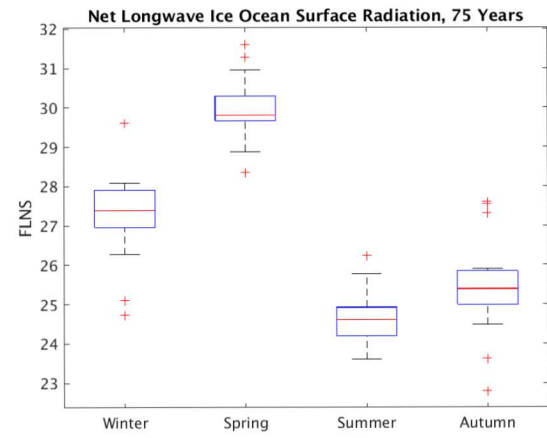
(c) SST



(d) TS

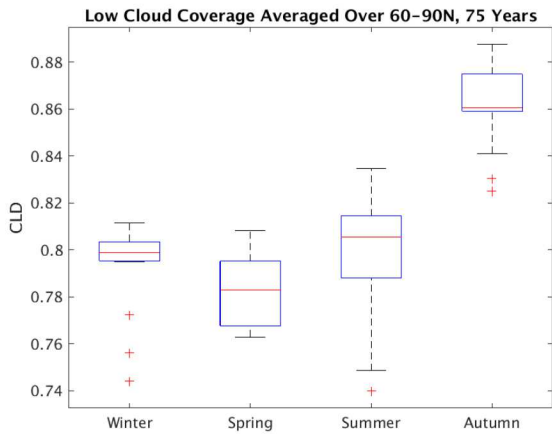


(e) QS

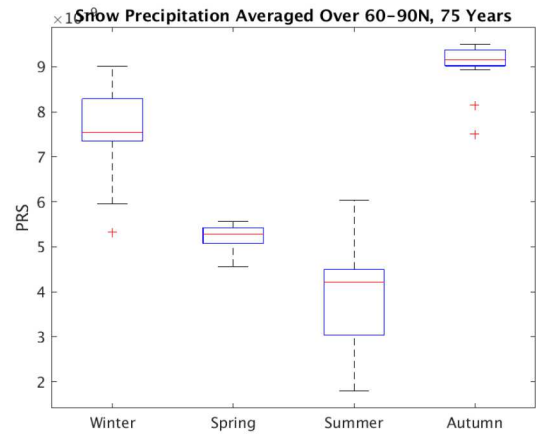


(f) FLNS

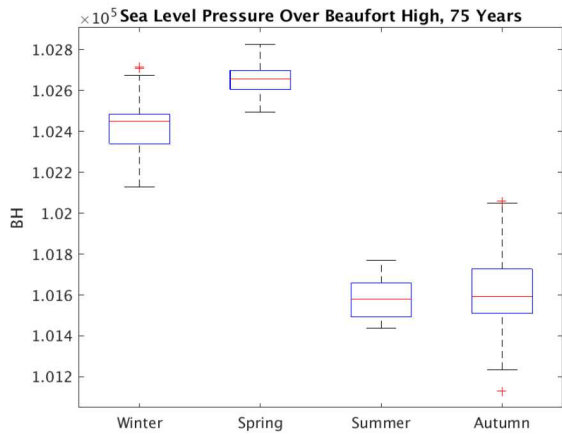
**Figure 5-23** Box plots of first six QOIs for perturbed `ne11` runs of 75 years, averaged over the last 25 years.



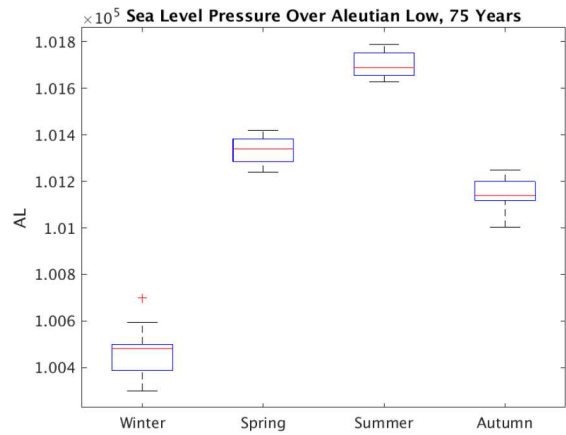
(a) CLDLow



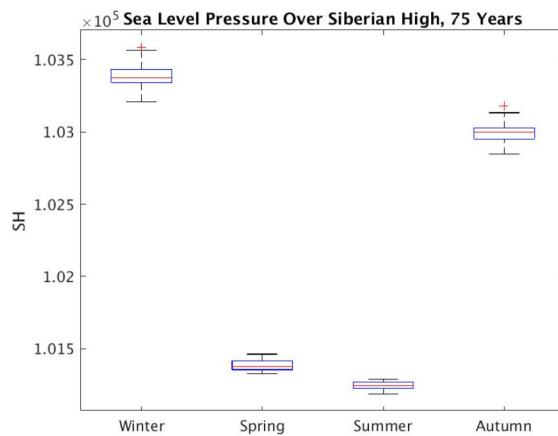
(b) PRECSL



(c) BH



(d) AL



(e) SH

Figure 5-24 Box plots of last five QOIs for perturbed ne11 runs of 75 years, averaged over the last 25 years.

## REFERENCES

- [1] Snow, water, ice, and permafrost in the Arctic (SWIPA). Technical report, Oslo, 2017.
- [2] Arctic report card 2019. Technical report, 2019.
- [3] Dorian S. Abbot, Mary Silber, and Raymond T. Pierrehumbert. Bifurcations leading to summer Arctic sea ice loss. *Journal of Geophysical Research*, 116(D19120), 2011.
- [4] B. Adams, L. Bauman, W. Bohnhoff, K. Dalbey, M. Ebeida, J. Eddy, M. Eldred, P. Hough, K. Hu, J. Jakeman, L. Swiler, and D. Vigil. DAKOTA, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 5.4 User’s Manual. Technical report, 2009 (updated 2013).
- [5] A. Alexanderian, J. Winokur, I. Sraj, A. Srinivasan, M. Iskandarani, W. Thacker, and O. Knio. Global sensitivity analysis in an ocean general circulation model: a sparse spectral projection approach. *Computational Geosciences*, 16:757–778, 2012.
- [6] X. Asay-Davis, D. Comeau, J. gyke, M. Hoffman, M. Petersen, S. Pricec, T. Ringler, A. Turner, L. Van Roekel, and P. Wolfram Jr. Antarctic ice shelf-ocean interactions in high-resolution, global simulations using the Energy Exascale Earth System Model (E3SM) Part 2: Sensitivity studies and model tuning. 2018 Ocean Sciences Meeting, American Geophysical Union, Portland, Oregon, 2018.
- [7] Sebastian Bathiany, Henk Dijkstra, Michel Crucifix, Vasilis Dakos, Victor Brovkin, Mark S. Williamson, Timothy M. Lenton, and Marten Scheffer. Beyond bifurcation: using complex models to understand and predict abrupt climate change. *Dynamics and Statistics of the Climate System*, 1(1), 11 2016.
- [8] Yoav Benjamini and Daniel Yekutieli. The control of the false discovery rate in multiple testing under dependency. *Ann. Statist.*, 29(4):1165–1188, 08 2001.
- [9] B. Bernard, G. Madec, T. Penduff, J.-M. Molines, A.-M. Treguier, J. Le Sommer, A. Beckmann, A. Biastoch, C. Boning, J. Dengg, C. Derval, E. Durand, S. Gulev, E. Remy, C. Talandier, S. Theetten, M. Maltrud, J. McClean, and B. De Cuevas. Impact of partial steps and momentum advection schemes in a global ocean circulation model at eddy-permitting resolution. *Ocean Dynamics*, 56:543–567, 2006.
- [10] Leo Breiman. Random forests. *Machine Learning*, 45:5–32, 2001.
- [11] Mitchell Bushuk and Dimitrios Giannakis. The seasonality and interannual variability of Arctic sea ice reemergence. *Journal of Climate*, 30:4657–4676, 2017.
- [12] J. Cohen, X. Zhang, J. Francis, T. Jung, R. Kwok, J. Overland, and D. Smith. Arctic change and possible influence on mid-latitude climate and weather: A US CLIVAR white paper. Technical Report 2018-1, 2018.
- [13] Judah Cohen, Karl Pfeiffer, and Jennifer A. Francis. Warm Arctic episodes linked with increased frequency of extreme winter weather in the United States. *Nature Communications*, 9(869), 2018.

- [14] Curt Covey, Donald D. Lucas, John Tannahill, Xabier Garaizar, and Richard Klein. Efficient screening of climate model sensitivity to a large number of perturbed input parameters. *Journal of Advances in Modeling Earth Systems*, 5(3):598–610, 2013.
- [15] Ivana Cvijanovic, Benjamin D. Santer, Celine Bonfils, Donald D. Lucas, John C. H. Chiang, and Susan Zimmerman. Future loss of Arctic sea-ice cover could drive a substantial decrease in California’s rainfall. *Nature Communications*, 8(1947), 2017.
- [16] Vasilis Dakos, Stephen R Carpenter, Egbert H van Nes, and Marten Scheffer. Resilience indicators: prospects and limitations for early warnings of regime shifts. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 370(1659):20130263, 2015.
- [17] Qinghua Ding, Axel Schweiger, Michelle L’Heureux, Eric J. Steig, David S. Battisti, Nathaniel C. Johnson, Eduardo Blanchard-Wrigglesworth, Stephen Po-Chedley, Qin Zhang, Kirstin Harnos, Mitchell Bushuk, Bradley Markle, and Ian Baxter. Fingerprints of internal drivers of Arctic sea ice loss in observations and model simulations. *Nature Geoscience*, 12(1):28–33, 2019.
- [18] V. Eyring, S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, 9(5):1937–1958, 2016.
- [19] F. Fetterer, K. Knowles, W. N. Meier, M. Savoie, and A. K. Windnagel. Sea ice index, version 3. Technical report, Boulder, Colorado USA, 2017.
- [20] R.A. Fisher. *Statistical Methods for Research Workers*. Biological monographs and manuals. Oliver and Boyd, 1934.
- [21] R.G. Ghanem and P.D. Spanos. *Stochastic Finite Elements: A Spectral Approach*. Springer-Verlag New York, Inc., New York, NY, USA, 1991.
- [22] Jean-Christophe Golaz, Peter M. Caldwell, Luke P. Van Roekel, Mark R. Petersen, Qi Tang, Jonathan D. Wolfe, Guta Abeshu, Valentine Anantharaj, Xylar S. Asay-Davis, David C. Bader, Sterling A. Baldwin, Gautam Bisht, Peter A. Bogenschutz, Marcia Branstetter, Michael A. Brunke, Steven R. Brus, Susannah M. Burrows, Philip J. Cameron-Smith, Aaron S. Donahue, Michael Deakin, Richard C. Easter, Katherine J. Evans, Yan Feng, Mark Flanner, James G. Foucar, Jeremy G. Fyke, Brian M. Griffin, Cecile Hannay, Bryce E. Harrop, Elizabeth C. Hunke, Robert L. Jacob, Douglas W. Jacobsen, Nicole Jeffery, Philip W. Jones, Noel D. Keen, Stephen A. Klein, Vincent E. Larson, L. Ruby Leung, Hong-Yi Li, Wuyin Lin, William H. Lipscomb, Po-Lun Ma, Salil Mahajan, Mathew E. Maltrud, Azamat Mamatjanov, Julie L. McClean, Renata B. McCoy, Richard B. Neale, Stephen F. Price, Yun Qian, Philip J. Rasch, J.E. Jack Reeves Eyre, William J. Riley, Todd D. Ringler, Andrew F. Roberts, Erika L. Roesler, Andrew G. Salinger, Zeshawn Shaheen, Xiaoying Shi, Balwinder Singh, Jinyun Tang, Mark A. Taylor, Peter E. Thornton, Adrian K. Turner, Milena Veneziani, Hui Wan, Hailong Wang, Shanlin Wang, Dean N. Williams, Philip J. Wolfram, Patrick H. Worley, Shaocheng Xie, Yang Yang, Jin-Ho Yoon, Mark D. Zelinka, Charles S. Zender, Xubin Zeng, Chengzhu Zhang, Kai Zhang, Yuying Zhang, Xue Zheng, Tian Zhou, and Qing Zhu. The DOE E3SM coupled model version 1: Overview and evaluation at standard resolution. *Journal of Advances in Modeling Earth Systems*, 11(7):2089–2129, 2019.

- [23] David L. Gold, Kevin R. Coombes, Jing Wang, and Bani Mallick. Enrichment analysis in high-throughput genomics-accounting for dependency in the NULL. *Briefings in Bioinformatics*, 8(2):71–77, 10 2006.
- [24] Hugues Goosse, Jennifer E. Kay, Kyle C. Armour, Alejandro Bodas-Salcedo, Helene Chepfer, David Docquier, Alexandra Jonko, Paul J. Kushner, Olivier Lecomte, Francois Massonnet, Hyo-Seok Park, Felix Pithan, Gunilla Svensson, and Martin Vancoppenolle. Quantifying climate feedbacks in polar regions. *Nature Communications*, 9(1919), 2018.
- [25] K. A. Graeter, E. C. Osterberg, D. G. Ferris, R. L. Hawley, H. P. Marshall, G. Lewis, T. Meehan, F. McCarthy, T. Overly, and S. D. Birkel. Ice core records of west greenland melt and climate forcing. *GRL*, 45:3164–3172, 2018.
- [26] Virginie Guemas, Edward Blanchard-Wrigglesworth, Matthieu Chevallier, Jonathan J. Day, Michel Deque, Francisco J. Doblas-Reyes, Neven S. Fuckar, Agathe Germe, Ed Hawkins, Sarah Keeley, Torben Koenigk, David Salas y Melia, and Steffen Tietsche. A review on arctic sea-ice predictability and prediction on seasonal to decadal time-scales. *Q. J. R. Meteorol. Soc.*, 142:546–561, 2016.
- [27] M. Hecht and R. Smith. Towards a physical understanding of the North Atlantic: A review of mode studies. In M. Hecht and H. Hasumi (Eds.), *Ocean Modeling in an Eddyding Regime Geophysical Monograph Series (Section 3, pp. 339–352)*. Washington, DC: American Geophysical Union, 2008.
- [28] M. W. Hecht, E. Hunke, M. E. Maltrud, M. R. Petersen, and B. A. Wingate. Lateral mixing in the eddyding regime and a new broad-ranging formulation. In M. Hecht and H. Hasumi (Eds.), *Ocean Modeling in an Eddyding Regime Geophysical Monograph Series, (Section 2, pp. 213–240)*. Washington, DC: American Geophysical Union, 2008.
- [29] Kaitlin Hill, Dorian S. Abbott, and Mary Silber. Analysis of an Arctic sea ice loss model in the limit of a discontinuous albedo. *SIAM Applied Dynamical Systems*, 15(2):1163–1192, 2016.
- [30] C. Hobbs, P. Ashwin, S. Wioeczorek, R. Vitolo, and P. Cox. Tipping points in open systems: bifurcation, noise-included and rate-dependent examples in the climate system. *Philosophical Transactions of the Royal Society A*, 370:1166–1184, 2012.
- [31] Boyin Huang, Viva F. Banzon, Eric Freeman, Jay Lawrimore, Wei Liu, Thomas C. Peterson, Thomas M. Smith, Peter W. Thorne, Scott D. Woodruff, and Huai-Min Zhang. Extended Reconstructed Sea Surface Temperature Version 4 (ERSST.v4). Part I: Upgrades and Inter-comparisons. *Journal of Climate*, 28(3):911–930, 2015.
- [32] Harley E. Hurlburt and Patrick J. Hogan. Impact of 1/8 degree to 1/64 degree resolution on gulf stream model-data comparisons in basin-scale subtropical Atlantic Ocean models. *Dynamics of Atmospheres and Oceans*, 32(3):283 – 329, 2000.
- [33] Monica Ionita, Klaus Grosfeld, Patrick Scholz, Renate Treffeisen, and Gerrit Lohmann. September Arctic sea ice minimum prediction - a new skillful statistical approach. *Earth System Dynamics*, 10:189–203, 2019.

- [34] J. Jakeman. Pyapprox: Release 1.0.0. In preparation, 2020.
- [35] J.D. Jakeman, M.S. Eldred, and K. Sargsyan. Enhancing  $\ell_1$ -minimization estimates of polynomial chaos expansions using basis selection. *Journal of Computational Physics*, 289(0):18 – 34, 2015.
- [36] J.D. Jakeman, F. Franzelin, A. Narayan, M. Eldred, and D. Pflüger. Polynomial chaos expansions for dependent random variables. *Computer Methods in Applied Mechanics and Engineering*, 2018. submitted.
- [37] M. Kanamitsu, W. Ebisuzaki, J. Woollen, S-K. Yang, J. J. Hnilo, M. Fiorino, and G.L. Potter. NCEP-DOE AMIP-II reanalysis (r-2). *Bulletin of the American Meteorological Society*, pages 1631–1643, 2002.
- [38] Jiwon Kim, Kwangjin Kim, Jaeil Cho, Yong Q. Kang, Hong-Joo Yoon, and Yang-Won Lee. Satellite-based prediction of Arctic sea ice concentration using a deep neural network with multi-model ensemble. *Remote Sensing*, 11(19), 2019.
- [39] Jong G. Kim, Elizabeth C. Hunke, and William H. Lipscomb. Sensitivity analysis and parameter tuning scheme for global sea-ice modeling. *Ocean Modelling*, 14(1):61 – 80, 2006.
- [40] Y. J. Kim, H.-C. Kim, D. Han, S. Lee, and J. Im. Prediction of monthly arctic sea ice concentrations using satellite and reanalysis data based on convolutional neural networks. *The Cryosphere*, 14:1083–1104, 2020.
- [41] Robert E Kopp, Rachael L Shwom, Gernot Wagner, and Jiacan Yuan. Tipping elements and climate–economic shocks: Pathways toward integrated assessment. *Earth’s Future*, 4(8):346–372, 2016.
- [42] Christian Kuehn. A mathematical framework for critical transitions: Bifurcations, fast–slow systems and stochastic dynamics. *Physica D: Nonlinear Phenomena*, 240(12):1020–1035, 2011.
- [43] Timothy M. Lenton. Tipping elements in the Earth climate system. *PNAS*, 105(6):1786–1793, 2008.
- [44] Timothy M. Lenton. Arctic climate tipping points. *Ambio*, 41(1):10–22, 2012.
- [45] Anders Levermann, Jonathan L Bamber, Sybren Drijfhout, Andrey Ganopolski, Winfried Haerberli, Neil RP Harris, Matthias Huss, Kirstin Krüger, Timothy M Lenton, Ronald W Lindsay, et al. Potential climatic transitions with profound impact on Europe. *Climatic Change*, 110(3-4):845–878, 2012.
- [46] Mathew E Maltrud and Julie L McClean. An eddy resolving global 1/10 degree ocean simulation. *Ocean Modelling*, 8(1):31 – 54, 2005.
- [47] Gerald A. Meehl, Christine T. Y. Chung, Julie M. ARblaster, Marika M. Holland, and Cecilia M. Bitz. Tropical decadal variability and the rate of Arctic sea ice decrease. *GRL*, 45:11326–11333, 2018.

- [48] W. Meier, F. Fetterer, M. Savoie, S. Mallory, R. Duerr, and J. Stroeve. NOAA/NSIDC climate data record of passive microwave sea ice concentration, version 3. Technical report, Boulder, Colorado USA, 2017.
- [49] John R. Mioduszewski, Stephen Vavrus, Muyin Want, Marika Holland, and Laura Landrum. Past and future interannual variability in Arctic sea ice in coupled climate models. *The Cryosphere*, 13:113–124, 2019.
- [50] Max D. Morris. Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33(2):161–174, 1991.
- [51] J. Nichol, M. Peterson, K. Peterson, D. Stracuzzi, M. Fricke, and M. Moses. Machine learning feature analysis explains disparity between climate models and observed climate change. *In preparation*, 2020.
- [52] Dirk Notz and Julienne Stroeve. The trajectory towards a seasonally ice-free Arctic ocean. *Current Climate Change Reports*, 4:407–416, 2018.
- [53] Dirk Olonscheck, Thorsten Mauritsen, and Dirk Notz. Arctic sea-ice variability is primarily driven by atmospheric temperature fluctuations. *Nature geoscience*, 12:430–434, 2019.
- [54] Ana C. Ordonez, Cecilia M. Bitz, and Edward Blanchard-Wrigglesworth. Processes controlling Arctic and Antarctic sea ice predictability in the Community Earth System Model. *Journal of Climate*, 31:9771–9786, 2018.
- [55] N. C. Parazoo, C. D. Koven, D. M. Lawrence, V. Romanovsky, and C. E. Miller. Detecting the permafrost carbon feedback: talik formation and increased cold-season respiration as precursors to sink-to-source transitions. *The Cryosphere*, 12(1):123–144, 2018.
- [56] G. Peng, W. N. Meier, D. Scott, and M. Savoie. A long-term and reproducible passive microwave sea ice concentration data record for climate studies and monitoring. *Earth Syst. Sci. Data*, 5:311–318, 2013.
- [57] K. Peterson, P. Bochev, and B. Paskaleva. Development, sensitivity analysis, and uncertainty quantification of high-fidelity arctic sea ice models. Sandia National Laboratories Report, SAND2010–6218, Albuquerque, NM and Livermore, CA, 2010.
- [58] Yun Qian, Hui Wan, Ben Yang, Jean-Christophe Golaz, Bryce Harrop, Zhangshuan Hou, Vincent Larson, L. Ruby Leung, Guanxing Lin, Po-Lun Ma, Hsi-Yen Ma, Phil Rasch, Balwinder Singh, Hailong Wang, Shaocheng Xie, and Kai Zhang. Parametric sensitivity and uncertainty quantification in the version 1 of E3SM atmosphere model based on short perturbed parameter ensemble simulations. *Journal of Geophysical Research Atmospheres*, 123:13046–13073, 2018.
- [59] J. G. I. Rae, H. T. Hewitt, A. B. Keen, J. K. Ridley, J. M. Edwards, and C.M. Harris. A sensitivity study of the sea ice simulation in the global coupled climate model, HadGEM3. *Ocean Modeling*, 74:60–76, 2014.

- [60] P. J. Rasch, S. Xie, P. L. Ma, W. Lin, H. Wang, Q. Tang, S. M. Burrows, P. Caldwell, K. Zhang, R. C. Easter, P. Cameron-Smith, B. Singh, H. Wan, J.-C. Golaz, B. E. Harrop, E. Roesler, J. Bacmeister, V. E. Larson, K. J. Evans, Y. Qian, M. Taylor, L. R. Leung, Y. Zhang, L. Brent, M. Branstetter, C. Hannay, S. Mahajan, A. Mametjanov, R. Neale, J. H. Richter, J.-H. Yoon, C. S. Zender, D. Bader, M. Flanner, J. G. Foucar, R. Jacob, N. Keen, S. A. Klein, X. Liu, A.G. Salinger, M. Shrivastava, and Y. Yang. An overview of the atmospheric component of the Energy Exascale Earth System Model. *Journal of Advances in Modeling Earth Systems*, 11:2377–2411, 2019.
- [61] Shanon M. Reckinger, Mark R. Petersen, and Scott J. Reckinger. A study of overflow simulations using MPAS-Ocean: vertical grids, resolution, and viscosity. *Ocean Modeling*, 96:291–313, 2015.
- [62] Erica Rosenblum and Ian Eisenman. Sea ice trends in climate models only accurate in runs with biased global warming. *Journal of Climate*, 30:6265–6277, 2017.
- [63] Marten Scheffer, Jordi Bascompte, William A Brock, Victor Brovkin, Stephen R Carpenter, Vasilis Dakos, Hermann Held, Egbert H Van Nes, Max Rietkerk, and George Sugihara. Early-warning signals for critical transitions. *Nature*, 461(7260):53–59, 2009.
- [64] Marten Scheffer, Stephen R. Carpenter, Timothy M. Lenton, Jordi Bascompte, William Brock, Vasilis Dakos, Johan Van de Koppel, Ingrid A. Van de Leemput, Simon A. Levin, Egbert H. Van Nes, et al. Anticipating critical transitions. *Science*, 338(6105):344–348, 2012.
- [65] E. Schuur, A. McGuire, C. Schadel, and et al. Climate change and the permafrost carbon feedback. *Nature*, 520:171–179, 2015.
- [66] A. Schweiger, R. Lindsay, J. Zhang, M. Steele, and H. Stern. Uncertainty in modeled Arctic sea ice volume. *J. Geophys. Res.*, 2011.
- [67] J. A. Screen and C. Deser. Pacific ocean variability influences the time of emergence of a seasonally ice-free Arctic Ocean. *Geophysical Research Letters*, 46:2222–2231, 2019.
- [68] Florian Sevellec, Alexey V. Fedorov, and Wei Liu. Arctic sea-ice decline weakens the atlantic meridional overturning circulation. *Nature Climate Change*, 7:604–610, 2017.
- [69] L. C. Smith and S. R. Stephenson. New trans-Arctic shipping routes navigable by midcentury. *PNAS*, 110(13):4871–4872, 2013.
- [70] I.M Sobol. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, 55(1):271 – 280, 2001. The Second IMACS Seminar on Monte Carlo Methods.
- [71] Julienne Stroeve and Dirk Notz. Changing state of Arctic sea ice across all seasons. *Environmental Research Letters*, 13:103001, 2018.
- [72] Julienne C. Stroeve, Vladimir Kattsov, Andrew Barrett, Mark Sereze, Tatiana Pavlova, Marika Holland, and Walter N. Meier. Trends in Arctic sea ice extent from CMIP5, CMIP3, and observations. *GRL*, 39:L16502, 2012.

- [73] Bruno Sudret. Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering and System Safety*, 93(7):964 – 979, 2008. Bayesian Networks in Dependability.
- [74] I. Tezaur, E. Roesler, A. Powell, K. Peterson, and J. Jakeman. Global sensitivity analysis of ultra-low resolution E3SM. JAMES (in preparation), 2020.
- [75] P. Uotila, S. O’Farrell, S.J. Marsland, and D. Bi. A sea-ice sensitivity study with a global ocean-ice model. *Ocean Modelling*, 51:1 – 18, 2012.
- [76] Jorge R. Urrego-Blanco, Elizabeth C. Hunke, and Nathan Urban. Emergent relationships among sea ice, longwave radiation, and the Beaufort high circulation exposed through parameter uncertainty analysis. *JGR Oceans*, 124:9572–9589, 2019.
- [77] Jorge R. Urrego-Blanco, Nathan M. Urban, Elizabeth C. Hunke, Adrian K. Turner, and Nicole Jeffery. Emergent relationships among sea ice, longwave radiation, and the Beaufort high circulation exposed through parameter uncertainty analysis. *JGR Oceans*, 121:2709–2732, 2016.
- [78] Till J. W. Wagner and Ian Eisenman. How climate model complexity influences sea ice stability. *Journal of Climate*, 28:3998–4013, 2015.
- [79] M. Winton. Do climate models underestimate the sensitivity of Northern hemisphere sea ice cover? *J. Climate*, 24:3924–3934, 2011.
- [80] D. Xiu and G.E. Karniadakis. The Wiener-Askey Polynomial Chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, 24(2):619–644, 2002.
- [81] C. Zhao, X. Liu, Y. Qian, J. Yoon, Z. Hou, G. Lin, S. McFarlane, H. Wang, B. Yang, P.-L. Ma, H. Yan, and J. Bao. A sensitivity study of radiative fluxes at the top of atmosphere to cloud-microphysics and aerosol parameters in the community atmosphere model CAM5. *Atmospheric Chemistry and Physics*, 13(21):10969–10987, 2013.







Figure A-4 Parameter values and number of simulation years for ultra-low resolution E3SM sensitivity analysis, part 4.

Run	keno	lambda_pond	draglo	dp1	c1	c8	gamma	tracer_kappa	crit_bulk_rich	salin_rest_piston_wel	Years completed
166	2.309702139347900E-01	8.5918727262132400E-05	2.8673141832416800E-02	5.8624331001192300E-02	3.1170030171051600E+00	2.8759751808829600E+00	2.1878909366205300E-01	1.1846923800185300E+03	3.6153879314661000E-01	1.9746008488815200E-06	100
167	4.6957145687192700E-01	6.4620050140125900E-05	1.54583231771000E-01	2.8011130653321700E-02	2.0703870477154900E+00	4.8147830519658900E+00	3.8113948153331900E-01	1.7816455096006400E+03	6.390942332693900E-01	1.0013672136468800E-06	81
168	3.9758860971778600E-01	3.7093246956833100E-05	7.2755345802847300E-02	2.5300970133394000E-02	4.8697276614773000E+00	2.6045140717178600E+00	1.2614369578659500E-01	1.0361135998740800E+03	8.4251646716147700E-01	1.6493755385745300E-06	100
169	2.191558622330400E-01	7.4031943944477100E-05	6.52323229970270700E-02	7.1652524459022200E-02	1.1355112147013000E+00	5.28339259457555300E+00	4.436932289972900E+00	1.1732993949204700E+03	9.0793189220130400E+00	1.4845494863344300E-06	45
170	4.045496899634600E-01	5.3026470462853700E-05	5.1160007674805800E-02	7.6544163133949000E-02	3.5772070921957500E+00	4.4507333274759500E+00	4.2950146095827200E-01	7.8497378043830400E+02	8.8189250063151100E-01	1.8182324280031000E-06	100
171	4.0243432838469700E-01	1.1410468972318100E-04	8.1856409292016100E-02	7.6033179555088200E-02	4.5343591300770600E+00	3.9077162584289900E+00	4.0861849253997200E-01	1.2280828244984100E+03	5.2403776217252000E-01	1.2770436450839000E-06	100
172	4.8311363300308600E-01	8.9270624232693900E-05	3.4691835517715600E-02	4.6805336829274800E-02	3.9769496405497200E+00	7.3803434250876300E+00	4.5696228034794300E-01	1.7730379885062600E+03	4.43101015295833400E-01	1.9800192654515000E-06	100
173	5.4209397509694100E-01	1.7136588009078800E-05	1.2655814353032000E-01	6.4893565103411600E-02	2.1191565683115000E+00	6.3942298511594500E+00	4.398072076961400E-01	6.4287974154576700E+02	2.1650080718100100E+00	1.9099823269147800E-06	100
174	2.666387747451700E-01	4.1814432949831800E-05	1.53235157866751800E-01	4.7174275200814000E-02	2.5793736223134000E+00	3.442679381571800E+00	4.0160513697192100E-01	8.3833136390894700E+02	2.6270392406731800E-01	1.6274254855699800E-06	48
175	5.833747858115900E-01	5.7630156429822900E-05	1.4780465785148600E-01	4.5672231446951600E-02	3.7661276679652900E+00	2.2971734129823700E+00	3.2976461434736800E-01	1.5062004178762400E+03	9.3028910160064700E-01	2.8353837630502100E-06	77
176	4.3024115630000000E-01	5.4933665130000000E-05	3.5848797890000000E-02	9.3816934280000000E-02	2.4131250010000000E+00	5.0940834770000000E+00	4.7415892430000000E-01	1.6985495070000000E+03	6.8040586690000000E-01	1.4765137310000000E-06	100
177	4.3024115627631500E-01	5.4933665130397500E-05	3.584879788985100E-02	9.3816934283822800E-02	2.4131250008940700E+00	5.0940834768116500E+00	4.7415892425924500E-01	1.6985495072789500E+03	6.8040586691349700E-01	1.4765137306647400E-06	124
178	5.8073857121439200E-01	4.9169662301393700E-07	1.5057331461282400E-01	7.8606865797191900E-02	1.7508352156728500E+00	2.4580919868312800E+00	1.340064082271500E-01	1.2120364042930300E+03	8.7801276050674900E-01	2.2760220885742500E-06	90
179	4.638665388382700E-01	5.5077636406779400E-05	1.5701840515937700E-01	8.2591047976166000E-02	4.885358637201800E+00	7.8103538318537200E+00	2.7669578511267900E-01	1.6400398724712400E+03	6.9865107219666200E-01	1.2883235960733100E-06	0
180	4.7210121694952300E-01	8.1092099322454900E-05	5.5838200872251800E-02	9.2337098568677900E-02	1.5553323710337300E+00	2.6279838057234900E+00	4.8141612475737900E-01	1.3564006749540600E+03	2.434284331396600E-01	2.4776736483909200E-06	96
181	4.2545746183022900E-01	1.6041218761489600E-05	2.0921002048114300E-02	6.3276691399514700E-02	2.7943333107978100E+00	5.4301098803989600E+00	4.1281257709488300E-01	6.4238348929211500E+02	2.4120276495814300E-01	2.4551838757935900E-06	104
182	3.8668197216466100E-01	8.3016157804396000E-05	1.1346446903717700E-01	4.1621464472264100E-02	1.1662989249453000E+00	7.2500295182690000E+00	2.9686608109623200E-01	9.3369241906330000E+02	3.2408646270632800E-01	2.3642228602897400E-06	12
183	2.1613511452451400E-01	9.8615156163144200E-05	1.2840410348949900E-01	5.5938940905304500E-02	3.2990078916773200E+00	3.7029740926809600E+00	1.6086885593831500E-01	1.0113452285528200E+03	6.38081215880811100E-01	2.939546946648700E-06	22
184	3.9660719875246300E-01	1.9874969549821500E-05	1.4979565603667900E-01	3.3786232899874500E-02	1.1368096096632900E+00	6.5109318234026400E+00	4.0150422854349000E-01	9.8116346178576400E+02	2.3031776603311300E-01	1.3525951949646700E-06	92
185	4.2156834695488200E-01	3.2360018055582100E-05	7.6485668576648500E-02	2.4366636108607100E-02	4.7830387949943500E+00	5.092293621332000E+00	4.2156861949324600E-01	1.4164850866422100E+03	6.2070359904319100E-01	2.5640503712697000E-06	124
186	2.8125199118766400E-01	8.0723584557567500E-05	9.7414418866513900E-02	3.4343082606792400E-02	3.0951264081528000E+00	5.817615252478000E+00	3.171276172796500E-01	1.5383168594911700E+03	2.3844751212745900E-01	1.678329117147000E-06	124
187	2.670148378238000E-01	1.093013963077100E-04	1.571777824568600E-02	6.826803555712100E-02	1.6177696726620500E+00	5.811059392988600E+00	6.625736680813100E-01	1.365831616396700E+03	9.906209457665800E-01	2.1077120811911300E-06	45
188	4.0791141800180100E-01	7.9412578531711700E-05	1.929865217349000E-02	3.9769106190651700E-02	3.4265314377844300E+00	2.0364070761256700E+00	1.6189279444515700E-01	7.8612850997597000E+02	7.6711918208748100E-01	2.0261873462004600E-06	100
189	2.793057671748100E-01	2.6612583611362400E-05	1.3675947158518300E-01	3.6656787823885700E-02	3.6829276271164400E+00	3.6101269368082300E+00	3.0536233661696300E-01	7.2949467998860600E+02	8.3187941312789900E-01	2.9612988396547700E-06	124
190	5.5425502704456500E-01	8.7846752292809100E-05	3.4704879315290600E-02	4.9377944264560900E-02	1.6545083411037900E+00	3.7160324663851500E+00	3.979560217820100E-01	1.4375814769417000E+03	3.6375286504626300E-01	6.375286504626300E-06	50
191	4.4676194097846700E-01	9.7404168270992300E-05	1.4649182849768500E-01	3.827331792563200E-02	4.9288343898952000E+00	6.8241178477266500E+00	2.6133273812010900E-01	8.509327724576000E+02	5.6302730478346400E-01	1.4998270273441400E-06	4
192	5.451910650357600E-01	4.7776859830216700E-05	1.6562955687381300E-02	5.0561130858957800E-02	2.3036936791613700E+00	5.1276736604049800E+00	2.4560039732605200E-01	1.4844233572483100E+03	9.6342615857720400E-01	1.45728382810300E-06	20
193	3.3502435693517300E-01	6.1507583185631100E-06	1.1847176701836300E-01	5.6944061163812900E-02	3.8723055716582200E+00	3.2820977014489500E+00	3.539715405553800E-01	1.0924849050119500E+03	3.3091719206422600E-01	2.9425168870016900E-06	100
194	3.248843019828200E-01	2.1980731699632400E-05	4.3904890527296800E-02	8.8654081366071300E-02	4.6362838773056900E+00	5.8144402559846600E+00	1.459253244102000E-01	8.52326235885477100E+02	7.3364266026765100E-01	9.4689849503338300E-07	12
195	5.0135618867352600E-01	5.4252531280157000E-05	5.4206730794254700E-02	2.8151483573019500E-02	3.5220405133441100E+00	7.8738156948238600E+00	2.7066962802782700E-01	1.7240723744966100E+03	6.1792059075087300E-01	1.4629281039582600E-06	99
196	2.0777502301005100E-01	6.1235203764765100E-05	1.4441980510535700E-01	5.8699780990752700E-02	2.2181620784103900E+00	2.0940087200142400E+00	2.0635248422622700E-01	1.5767546975053800E+03	7.4069267958402600E-01	1.9827974660089200E-06	29
197	5.7774155810475300E-01	9.1932614301925900E-05	8.4275252111442400E-02	5.869984101504100E-02	4.908931834623200E+00	2.2284284839406600E+00	2.0111035788431800E-01	1.5753912345506300E+03	7.1364767029881500E-01	1.0904725428903500E-06	11
198	2.8359232113005800E-01	7.0057953907469600E-05	7.6466690256726000E-02	2.5622593462467200E-02	3.1580107854679200E+00	7.5559925762936500E+00	3.4792152345180500E-01	1.6807260305620700E+03	9.55247354050744600E-01	2.4182661495637100E-06	99
199	4.6920622947738900E-01	4.485701268575200E-05	1.1922911231122000E-01	2.652185123413800E-02	2.5459908070042700E+00	4.9264968062557300E+00	1.4601968154311200E-01	1.3177915728651000E+03	3.6670832559466400E-01	2.2433652594384200E-06	15
200	2.0811064047738900E-01	1.0353819351543900E-04	5.1704062618780900E-02	9.8655551411211500E-02	3.9795090835541500E+00	7.259406474884600E+00	2.2656816892311000E-01	9.2233489416539700E+02	1.2119478432068100E-01	6.1361650079488700E-06	17
201	5.375393263942600E-01	7.1234231982181200E-05	1.3346450488413700E-01	6.812901301394700E-02	2.5423801734660300E+00	3.7073212270624900E+00	3.0014913436025400E-01	1.782224050723000E+03	7.645423439145100E-01	1.463176368235000E-06	99
202	5.853460787268100E-01	8.7952785589015000E-05	2.6902397310873500E-02	3.3548750877380400E-02	4.9153002062812400E+00	2.1150141619145900E+00	3.691481282404100E-01	1.0650357029400800E+03	7.7598501015454500E-01	2.1828724674639300E-06	91
203	3.1803072197362800E-01	2.8924508619664800E-05	7.8393721670517700E-02	3.373316700905600E-02	4.4197262283414600E+00	2.7396102117516500E+00	1.0342284440994300E-01	1.5427170921117100E+03	2.1877051834017000E-01	2.6624482376966600E-06	51
204	5.3072773376479700E-01	2.242608319524000E-05	9.3183891644422000E-02	4.2431769631803000E-02	2.7129595056116700E+00	4.7056231070309900E+00	2.2131559364497700E-01	8.4124727714806800E+02	8.9282684866339000E-01	1.5748624511528800E-06	93
205	3.8764130072668200E-01	4.4802329307481800E-05	8.1668824615236400E-01	5.6539704594761100E-02	4.329631143785700E+00	2.5818043751642100E+00	4.7017719680443400E-01	1.6321622882969700E+03	4.555404126644100E-01	1.8593697197735300E-06	45
206	3.273210650309900E-01	2.4742907554127500E-05	1.4682813750668400E-01	7.9186131991446000E-02	4.9274726845324000E+00	6.690451709505800E+00	4.883175174019300E-01	1.109694891126300E+03	7.4371043369174000E-01	2.5517158748349200E-06	99
207	5.8220937438309200E-01	9.817824076485700E-05	1.1952305628755100E-01	9.3698543477803500E-02	4.1876467978581800E+00	6.7679161508567600E+00	2.4201723113656100E-01	8.477306584020900E+02	4.1374580264091500E-01	2.9636001033010000E-06	99
208	5.9019658425822900E-01										

## APPENDIX B. INSTRUCTIONS FOR BUILDING AND SUBMITTING ULTRA-LOW SENSITIVITY SIMULATIONS ON SKYBRIDGE

In this section we provide instructions for running perturbed simulations from the ULR sensitivity analyses using the modified E3SM branch.

1. Clone E3SM fork.

```
git clone --recursive git@github.com:/karapeterson/E3SM.git
cd E3SM
```

2. Checkout the branch

```
git checkout add_namelist_params
```

3. Update the submodules.

```
git submodule update
```

4. Remove the existing version of MPAS-source.

```
cd components
rm -rf mpas-source
```

5. Clone branch of MPAS-source.

```
git clone git@github.com:/karapeterson/MPAS-Model.git mpas-source
```

6. Checkout MPAS-source branch:

```
cd mpas-source
git checkout add_namelist_params
```

7. Move to top E3SM directory and check the submodules.

```
cd ../../
git submodule status
```

Results should be:

```
1a3be630bbd8f8ddc2e8020731a62ad0489c4430 components/clm/src/external_models/fates (sci.1.0.0_api.1.0.0-283-g1a3be63)
38e0b0461ce10da7de809b752bacc70860ae8bbd components/clm/src/external_models/mpp (v0.0.1-74-g38e0b04)
f2cbe748addb922c3df55835ae75d7204be9ffda components/clm/src/external_models/sbetr (v.10-44-gf2cbe74)
+51489af3dcf68a8766cf2d8a0db82d241453ff51 components/mpas-source (v6.0-17-g51489af)
758cb47c83c84eb94575cebb81bf353f975f9d2d externals/scorpio (scorpio-v1.0.1)
1e538b56fa7cfa209c119e866a95669875eb200d externals/scorpio_classic (scorpio-classic-v1.0.4)
```

8. Create a Branch Run. From the E3SM/cime/scripts directory. For ne4 run:

```
./create_newcase --res ne4_oQU240 --compset A_WCYCL1850 --project FY180068
--case /nscratch/$USER/acme_scratch/cases/master.A_WCYCL1850.ne4_oQU240.sensitivity_1 --walltime 48:00:00 --pecount 96
```

For ne11 run:

```
./create_newcase --res ne11_oQU240 --compset A_WCYCL1850 --project FY180068
--case /$SCRATCH_TYPE/$USER/acme_scratch_ne11/cases/master.A_WCYCL1850.ne11_oQU240.sensitivity_$case_num
--walltime 48:00:00 --pecount 384
```

9. Move to the case directory. For ne4 run:

```
cd /nscratch/$USER/acme_scratch/cases/master.A_WCYCL1850.ne4_oQU240.sensitivity_1
```

For ne11 run:

```
cd /nscratch/$USER/acme_scratch_ne11/cases/master.A_WCYCL1850.ne4_oQU240.sensitivity_1
```

10. Run case setup.

```
./case.setup
```

11. Modify env\_run.xml. For ne4 run:

```
STOP_OPTION=nyears
STOP_N=25
RESUBMIT=3
RESUBMIT_SETS_CONTINUE_RUN=TRUE
RUN_TYPE=branch
RUN_REFDIR=/gscratch/arctic_tipping_point_results/ELR_E3SM_output_stash
                /master.A_WCYCL1850.ne4_oQU240.branch.tune.Golaz2019/run/
RUN_REFCASE=master.A_WCYCL1850.ne4_oQU240.branch.tune.Golaz2019
RUN_REFDATE=0676-01-01
RUNDIR=/nscratch/$USER/acme_scratch/sandiatoss3/$CASE/run
```

For ne11 run:

```
STOP_OPTION=nyears
STOP_N=25
RESUBMIT=3
RESUBMIT_SETS_CONTINUE_RUN=TRUE
RUN_TYPE=branch
RUN_REFDIR=/nscratch/arctic_tipping_points_results/ne11_spin_up_restart_rpointer_files/
RUN_REFCASE=master.A_WCYCL1850.ne11_oQU240.Golaz_Tuning_May31_2020_pecount_480
RUN_REFDATE=0201-01-01
RUNDIR=/nscratch/$USER/acme_scratch_ne11/sandiatoss3/$CASE/run
```

The above setup is for a 100 year run, submitted in 25 year chunks.

12. Create symbolic links to restart files and copy rpointer files into run directory. For ne4 run:

```
cd /nscratch/$USER/acme_scratch/sandiatoss3/master.A_WCYCL1850.ne4_oQU240.sensitivity_1/run/
ln -s /gscratch/arctic_tipping_point_results/ELR_E3SM_output_stash
                /master.A_WCYCL1850.ne4_oQU240.branch.tune.Golaz2019/run/*676*.nc .
cp /gscratch/arctic_tipping_point_results/ELR_E3SM_output_stash
                /master.A_WCYCL1850.ne4_oQU240.branch.tune.Golaz2019/run/rpointer* .
```

For ne11 run:

```
cd /nscratch/$USER/acme_scratch/sandiatoss3/master.A_WCYCL1850.ne4_oQU240.sensitivity_1/run/
cp /ascldap/users/ikalash/E3SM_Kara_scratches_ne11/rpointer_files/* .
ln -s /nscratch/arctic_tipping_points_results/ne11_spin_up_restart_rpointer_files/*201*nc .
```

13. Move back to case directory. For ne4 run:

```
cd /nscratch/$USER/acme_scratch/cases/master.A_WCYCL1850.ne4_oQU240.sensitivity_1
For ne11 run:
\begin{verbatim}
cd /nscratch/$USER/acme_scratch_ne11/cases/master.A_WCYCL1850.ne4_oQU240.sensitivity_1
```

14. Add user namelists (user\_nl\_cam, user\_nl\_mpaso, user\_nl\_mpscice) from DAKOTA sample output to case directory

15. Build and submit in case directory

```
./case.build
./case.submit
```

## APPENDIX C. OUTCOMES AND IMPACTS

### C.1. Publications

- J. Nichol, M. Peterson, K. Peterson, G. Fricke, M. Moses, Using Machine Learning to Compare Features of Simulated and Observed Arctic Sea Ice Extent, in preparation.
- I. Terza, E. Roesler, A. Powell, K. Peterson, J. Jakeman, Global Sensitivity Analysis of Ultra-low resolution E3SM, in preparation.

### C.2. Presentations

- J. Nichol, Machine Learning to Compare Arctic Simulations with Observed Data, ESCO 2020, June 2020, (online). Won award for best student presentation.
- A. Powell, Using Ultra-Low Resolution E3SM Simulations to Predict Sea Ice Free Summers, and to Elucidate the Role of Arctic Sea Ice in Polar Amplification, ESCO 2020, June 2020 (online)
- M. Peterson, Predicting Arctic sea ice concentration with data-driven models, ICIAM, Valencia, Spain, July 2019.
- K. Peterson, Sea Ice Modeling and Arctic Change, UNM Women in Computing Seminar, March 2019.
- D. Bull, National security implications from tipping events centered in Arctic waters, 2018 International Symposium: Climate Change Effects on the World's Oceans, 4-8 June 2018, Washington DC.
- D. Bull, Methodologies to Optimize Changing National Security Preparedness Demands Arising with Increasing Arctic Access, DoD Arctic S&T Synchronization Workshop, 16-18 May 2018, CRREL.

### C.3. Posters

- J. Nichol, Comparing Simulated and Observed Data with Random Forest Feature Importance, CoDA, February 2020. 3rd Place in student poster competition.
- J. Nichol, Using Machine Learning to Compare Simulated and Observational Sea Ice Extent Data, AGU Fall Meeting, December 2019.
- K. Peterson, Arctic Sea Ice Internal Variability in E3SM and Its Response to Anthropogenic Forcing, AGU Fall Meeting, December 2019.
- A. Powell, Exploring the Use of Ultra-Low Resolution E3SM Simulations to Predict Sea Ice Free Summers, and to Elucidate the Role of Arctic Sea Ice in Polar Amplification, AGU Fall Meeting, December 2019
- M. Peterson, Predicting Minimum Arctic Sea Ice Extent, Sandia Machine Learning R&D Workshop, September 2019

## DISTRIBUTION

### Hardcopy—Internal

Number of Copies	Name	Org.	Mailstop
1	1395	08880	Paul Shoemaker
1	0750	08162	Tracy Woolever
1	0701	06160	Erik Webb
1	1320	01442	Mike Parks
1	1318	01446	Andy Salinger
1	0750	08863	Lori Parrott
1	9159	08754	Jerry McNeish
1	D. Chavez, LDRD Office	1911	0359

### Email—Internal

Name	Org.	Sandia Email Address
Technical Library	01177	libref@sandia.gov







Sandia  
National  
Laboratories

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.